AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Correlated Non-Classical Measurement Errors, 'Second Best' Policy Inference and the Inverse Size-Productivity Relationship in Agriculture

C. Barrett[1]; K. Abay[2]; G. Abate[3]; T. Bernard[4]

1: Cornell University, Charles H Dyson School of Applied Economics & Management, United States of America, 2: University of Copenhagen, Economics, Denmark, 3: IFPRI, , Ethiopia, 4: IFPRI/universite de bordeaux, , France

Corresponding author email: cbb2@cornell.edu

Abstract:

We show analytically and empirically that non-classical measurement errors in the two key variables in a hypothesized relationship can bias the estimated relationship between them in any direction. Furthermore, if the errors are correlated, correcting for either one alone can aggravate bias in the parameter estimate of interest relative to ignoring mismeasurement in both variables, a 'second best' result with implications for a broad class of economic phenomena of policy interest. We illustrate these results empirically by demonstrating the implications of mismeasured agricultural output and plot size for the long-debated (inverse) relationship between size and productivity. Our data from Ethiopia show large discrepancies between farmer self-reported and directly measured values of crop output and plot size; these errors are strongly, positively correlated with one another. In these data, correlated non-classical measurement errors generate a strong but largely spurious estimated inverse size-productivity relationship. We demonstrate empirically our analytical result that correcting for just one measurement problem may aggravate the bias in the parameter estimate of interest.

Acknowledgement: This paper benefited from comments by Marc Bellemare, Leah Bevis, Chris Boone, Brian Dillon, John Gibson, Kalle Hirvonen and seminar participants at the African Development Bank. Any remaining errors are the authors' sole responsibility.

JEL Codes: O12, C12

#478

Correlated Non-Classical Measurement Errors, 'Second Best' Policy Inference and the Inverse Size-Productivity Relationship in Agriculture

December 2017 version for comments

Abstract: We show analytically and empirically that non-classical measurement errors in the two key variables in a hypothesized relationship can bias the estimated relationship between them in any direction. Furthermore, if the errors are correlated, correcting for either one alone can aggravate bias in the parameter estimate of interest relative to ignoring mismeasurement in both variables, a 'second best' result with implications for a broad class of economic phenomena of policy interest. We illustrate these results empirically by demonstrating the implications of mismeasured agricultural output and plot size for the long-debated (inverse) relationship between size and productivity. Our data from Ethiopia show large discrepancies between farmer self-reported and directly measured values of crop output and plot size; these errors are strongly, positively correlated with one another. In these data, correlated non-classical measurement errors generate a strong but largely spurious estimated inverse size-productivity relationship. We demonstrate empirically our analytical result that correcting for just one measurement problem may aggravate the bias in the parameter estimate of interest.

Keywords: Agricultural development, correlated errors, non-classical measurement error, smallholder agriculture.

1. Introduction

Measurement drives analysis. The quality of descriptive and predictive evidence is only as good as the underlying measures used to test key hypotheses. In recent years, empirical researchers have begun to devote considerably more effort to careful measurement and to explore the consequence of different measurement methods for key variables of direct policy relevance.[1] Of particular concern is non-classical measurement error (NCME), which occurs when the error in measuring a variable of interest is correlated with the true value of that variable, with the true values of other variables in the model, or with the errors in measuring those values (Bound et al. 2001). Many papers have clearly demonstrated the widespread prevalence of NCME and its relevance for policy inference in a range of fields, especially labor (e.g., Borjas, 1980; Bound and Krueger, 1991; Bound et al., 1994; French, 2004; Kim and Solon, 2005; Arthi et al., 2018), consumer behavior (Gibson and Kim, 2010; Gibson et al. 2015), development (Baird and Özler, 2012; Beegle et al., 2012; Chao et al., 2012; Desiere and Jolliffe 2018), health (Das et al., 2012; Larsen et al. 2017), and agriculture (De Groote and Traoré, 2005; Carletto et al., 2013; Carletto et al., 2015; Gourlay et al., 2017). The sensible guidance provided by that literature is to employ better measurement methods so as to reduce error. The rise of improved techniques based on high resolution remote sensing, mobile phone, imagery, global positioning system (GPS) and biomarker data, along with electronic survey data entry, steadily opens up new possibilities for reducing policy-relevant measurement error.

Yet in many domains, multiple variables fall prey to NCME. Moreover, mismeasurements may often be correlated between variables, for any of several reasons. For example, survey respondents might consciously and systematically underreport assets and earnings in order to reduce prospective tax liabilities or to increase the likelihood of being deemed eligible for some benefit. Or unconscious error may arise from rounding (sometimes known as 'focal point bias') so that variables that naturally exhibit positive skewness, such as asset holdings or earnings, will commonly exhibit upwardly biased and positively correlated measurement error as a result. Or one

[1] The special issue of *Journal of Development Economics* on measurement and survey design, introduced by McKenzie and Rosenzweig (2012), was a watershed event pushing more careful measurement in development economics. Ozler (2013)'s Development Impact blog entry helped call the development community's attention to these important issues more broadly.

mismeasured variable might be used by a respondent to generate an optimal prediction of another variable (Hyslop and Imbens, 2001), resulting in correlated measurement errors.

If multiple variables are measured with error but only some are amenable to correction, does correction for just one, but not both, otherwise-mismeasured variables reduce bias and improve inference, especially if those measurement errors are correlated? To the best of our knowledge, this important question has not yet been explored in the literature. Yet correlated NCME matters for the same reason that omitted relevant variables matter. When one estimates a parameter omitting multiple relevant variables, each correlated with the regressor of interest, and then controls for only one of them, the resulting parameter estimate does not necessarily move closer to the true parameter value. Indeed, depending on the correlations between the omitted and included regressors and the omitted and dependent variables, the more inclusive regression could generate increased bias. The problem of correlated NCME is conceptually quite similar. Hence our central analytical finding that correcting for one source of measurement error need not reduce bias in a parameter estimate of interest and its corollary, that if one cannot correct for both sources of measurement error, a 'second best' estimate may be preferable in the sense of reduced bias.

This problem arises for a wide range of economic questions. For example, estimates of the wage elasticity of labor supply may be subject to error in measures of earnings and hours worked, the latter of which serves as both the dependent variable and the denominator of the standard wage measure, leading to 'division bias' (Borjas, 1980). Correlated errors in nominal output and price measures may similarly bias the estimated relationship between real output or total factor productivity and inflation (Diewert and Fox, 1999). And measurement error in children's ages, which are likely correlated with errors in height or weight measures used jointly to construct standard anthropometric indicators such as height-for-age, can significantly bias estimates of the determinants of child health (Larsen et al., 2017).

In this paper we explore the consequences of correlated NCME in the key variables of interest in the long-studied size-productivity relationship (SPR) in agriculture.[2] The analytical findings are quite general, but framed so as to lead directly to the empirical part of the paper. The SPR has been studied extensively because of its considerable implications for agricultural

[2] As is well known, classical measurement error is just a special case of the more general NCME form we study. Classical measurement error generates attenuation bias in parameter estimates and artificially inflates variance that may provide misleading description of, for example, income inequality or mobility (Gottschalk and Huynh, 2010). Such bias declines as panel survey intervals increase (Naschold and Barrett, 2011).

development policy. For decades, findings of an inverse relationship were widely invoked to support land reform programs and to support claims of widespread factor market failures that justify interventions. Earlier studies typically found an inverse relationship between farm size and crop output per unit cultivated area (i.e., yield, a partial productivity indicator), attributing this empirical regularity to factor market imperfections (e.g. Sen, 1966; Feder, 1985; Barrett, 1996) or omitted land attributes, including soil quality (Benjamin, 1995; Assuncao and Braido, 2007; Barrett et al., 2010).[3] Recently, improvements in agricultural data collection have allowed researchers to explore the implication of measurement errors in self-reported production and farm or plot size.[4] Some have examined the implication of improved area measurement in explaining the inverse relationships between farm size and productivity using GPS measures of the surface area of plots (Carletto et al., 2013; Holden and Fisher, 2013; Carletto et al., 2015). Most recently, a few papers have explored the implication of measurement errors in farmer self-reported crop output on the estimated SPR using crop-cuts as a more objective measure of production (Gourlay et al., 2017; Desiere and Jolliffe, 2018). They find that, in their data, the inverse relationship is essentially driven by measurement errors associated with self-reported production. The relationship disappears upon using crop-cuts in place of self-reported production.

While these few, recent studies explore the implication of measurement error associated with either area or production, no study has yet considered both measurement problems in a unified framework. This is particularly crucial if both size and production suffer NCME and these measurement errors are correlated. As we demonstrate, when both production and farm size are inaccurately measured, and these errors are correlated, correcting for measurement error in just one variable is not sufficient to generate a consistent and unbiased estimate of the SPR. Furthermore, while previous studies show similar features of measurement errors in self-reported area and production, we know little as to why they generate conflicting empirical implications as to the effects on the estimated SPR.[5] By studying correlated NCME in a more general setting, we

[3] Bevis and Barrett (2017) provide another new behavioral explanation for the inverse relationship between farm size and productivity. They argue that productivity may be higher around the periphery of plots, partly for biophysical 'edge effect' reasons (e.g., improved access to sunlight) but mainly for behavioral reasons (e.g., greater observability of edges). As smaller plots have a greater ratio of edges to interior area, this can explain the commonly observed inverse relationship between productivity and farm size.

[4] We will use the terms 'output' and 'production' synonymously and similarly 'area' and 'size'.

[5] Carletto et al. (2013) find that inaccuracies in land area measurement lead to underestimation of the inverse relationship between plot size and productivity while Carletto et al. (2015) show the opposite.

can reconcile these findings. More importantly, we analytically and empirically establish that correcting just one variable can aggravate rather than attenuate bias in the SPR estimate.

In what follows, we analytically and empirically characterize the implication of various forms of measurement errors in self-reported crop production and cultivated area. We first set up a general framework that allows measurement errors in both output and area as well as potential correlations in these errors. We then analytically characterize the implication of alternative features of measurement errors in output and land area on the SPR. We then empirically demonstrate our analytical findings, employing both self-reported and objectives measures of output and area from an agricultural household survey in Ethiopia. For production, we compare farmers' self-reported production measures and production estimates based on crop-cuts, which are widely considered the gold standard for measuring agricultural output. We similarly rely on both farmer-reported land area as well as measurements based on compass-and-rope method, also known as Polygon method.[6] Compass-and-rope is similarly considered the most reliable method to accurately measure land area (Keita and Carfagna, 2009; Fermont and Benson, 2011; Carletto et al., 2015; Carletto et al., 2016).[7] By employing these four different measures of farm size and production, we illustrate empirically the patterns our analytical results predict regarding the long-debated SPR.

We make three contributions to the literature. First, by characterizing both types of measurement errors in a unified framework, we show that when both size and production measurements suffer from correlated measurement errors, the implication of these inaccuracies on the estimated SPR are analytically ambiguous, depending on several parameters characterizing these measurement errors. To the best of our knowledge, this is the first paper to provide a general analytical framework for understanding the implications of correlated NCME, and of incomplete correction for correlated NCME, for inference. In our data, we find that measurement errors in self-reported area and production are strongly correlated. As a result, correcting for either problem alone may not ensure unbiased estimation of the SPR. Indeed, our analytical and empirical exercises show that correcting for either measurement problem alone may even aggravate bias in

[6] Also, known as traverse measurement, the method involves measuring the length of each side and the angle of each corner using a measuring rope and a compass and the surface area of the measured plot can then be calculated using trigonometry (De Groote and Traoré 2005; Schøning et al., 2005; Casley and Kumar, 1988). Although the method is cumbersome and time consuming, it remains the approach of choice for specialized data collection due mainly to its accuracy compared to GPS or self-reported measures (Carletto et al., 2015).

[7] For example, Fermont and Benson (2011) compare farm size measurement using GPS and compass-and-rope, and show that GPS estimates significantly underestimate smaller farm sizes while both methods perform comparably for larger plots (those greater than 0.5 ha).

the SPR estimate relative to ignoring both measurement problems. This is essentially a 'theory of the second best' (Lancaster and Lipsey 1956) result and serves as a useful caution against overconfidence in the gains from improved measurement of single, key variables.

Second, we empirically corroborate in a new data set the core findings of recent studies (Holden and Fisher, 2013; Gourlay et al., 2017; Desiere and Jolliffe, 2018) that claim that measurement error explains the inverse relationship observed in farmer self-reported area and productivity data. Our results refine these prior findings by identifying features of measurement errors that can generate a spurious inverse SPR. That result is not automatically a byproduct of measurement error in area and output, particularly if these errors are correlated.

Third, our analytical framework and data allow us to compare the relative impact of the measurement errors in self-reported production and area on the estimated SPR. We analytically show and empirically find that when both variables suffer from similar measurement errors, inaccuracies associated with production are relatively more consequential. We also document that measurement errors in self-reported production and area may also affect parameter estimates relating productivity to other covariates of interest (e.g., soil characteristics).

Despite our emphasis on the estimated relationship between agricultural productivity and cultivated area, our analytical results have far more general implications. Not only do these findings reinforce previous concerns about recall-based and self-reported agricultural data, they also reveal the existence of an empirical equivalent to the theory of the second best (Lipsey and Lancaster, 1956), which holds that when one market failure in an economy cannot be corrected, efficiency may – counterintuitively – be maximized by introducing an offsetting market distortion. Two market failures may cancel each other out. We similarly demonstrate that when there exist NCME in both the dependent and independent variables of interest, and especially if those errors are correlated, then correcting for just one source of measurement error may, paradoxically, exacerbate the bias in the resulting parameter estimate of interest.

2. Measurement Errors in Household Surveys

Most agricultural research relies on self-reported, recall-based data. Due to cost and logistical considerations, most data are collected through single visit household surveys – perhaps repeated over time to generate longitudinal (i.e., panel) data – using extensive multi-topic instruments. Respondents are asked to recall and aggregate information often over many months and, in the case of agriculture, sometimes across two or more separate harvests of multiple crop types. While recall and aggregation errors can affect many metrics, they can have especially pronounced consequences for measuring area cultivated and production (i.e., harvested output).

Some such error may be ‘classical’, meaning the error itself is mean zero and uncorrelated with the true value of either the dependent variable or any independent variables of interest. For example, farmers in developing countries may lack the level of literacy and numeracy needed to accurately estimate and aggregate land area and crop production measurements, leading to significant, but random and symmetric (around the true value) measurement error (De Groote and Traoré, 2005). In a regression context, it is well known that classical measurement error will under-estimate relationships: either in absolute magnitude, in case the error lies with the independent variable (through attenuation bias); or in statistical significance, if the error lies with the dependent variable (through increase in the estimator’s variance). In the context of the SPR, classical measurement error will naturally bias estimates towards zero, i.e., toward failure to reject the constant returns to scale null hypothesis.

Non-classical measurement error, in which the error is correlated with the true variable(s) of interest, is of considerably greater concern. Multiple mechanisms might introduce NCME in self-reported land area and crop production. First, farmers may intentionally misreport their land area and crop production so as to conceal wealth and thereby avoid taxes or be found eligible for proxy means tested benefits of various types (Diskin, 1997). This type of misreporting can vary systematically with the true value of farm size, since those with little land or output have little or nothing to hide. Second, farmers may not accurately recall information related to much earlier events; in particular, extended recall periods may cause them to forget details of past events (Beegle et al., 2012; Arthi et al., 2018) or season-specific harvests (Ali et al., 2009; Howard et al., 1995).[8] Third, precise and universally applied measurement units may not be widely employed in low-income rural areas where imprecise local measures are commonplace. Traditional units can vary between locations and farming systems, implying that measurement and conversion into standardized units can introduce systematic errors. Finally, respondents may tend to round off values around focal points (e.g., one hectare or one day), a problem that may be more consequential, in percentage terms, for smaller plots and harvests than for larger ones.[9]

[8] Such recall bias affects many other agricultural metrics, including labor use (Arthi et al., 2018).

[9] While most of the above reasons apply to farm size and production measurements, there are additional problems that may affect measurement of production. For example, farmers may have forgotten season-specific harvests (Ali et al., 2009; Howard et al.,

While the inverse SPR was long observed in survey data, an emerging literature now argues that measurement errors in either land area or production may generate spurious estimates. On the land measurement side, recent studies relying on GPS devices consistently find evidence that farmers overestimate area for smaller plots and underestimate for larger ones (e.g., De Groote and Traoré, 2005; Carletto et al., 2013; Holden and Fisher, 2013; Carletto et al., 2015). However, the implication of area measurement error on estimating the SPR varies and sometimes contradicts each other. For example, Carletto et al. (2013) document that error in land area measurement underestimates the inverse relationship between farm size and productivity, while Carletto et al. (2015) find that it leads to overestimation of the inverse relationship.

On the production side, two recent studies find that the inverse relationship disappears when using crop-cuts instead of self-reported production. They conclude that the estimated inverse relationship is simply driven by measurement errors associated with production measurement (Gourlay et al., 2017; Desiere and Jolliffe, 2018).[10]

While the above few studies explore the implication of measurement error associated with either production or size, no study has analyzed the implication of measurement errors in both metrics. In many situations both area and output are measured with errors and this may have varying implications relative to the measurement errors in either one alone. This is particularly crucial if measurement errors in crop production and farm size are correlated. Intuitively, measurement errors in self-reported production and land area will often be correlated. For example, if households engage in strategic misreporting of land size, they may be more likely to do so as well for their harvests. Similarly, if rounding appears to be the main source of measurement error, rounding in both measures will naturally generate some correlation in measurement errors. For strictly positive-valued variables such as production and land area, upward rounding of production and area generates a potential positive correlation between measurement errors across both variables. The same will be true for positively skewed variables subjected to rounding around focal points, as the density in ranges beneath the focal point will typically exceed the density in the range above it. Below we analytically characterize alternative forms of measurement errors in land and production measurements.

1995) and portion of their production given as gifts and/or in-kind payments (David, 1978).

[10] Gourlay et al. (2017) also used high-resolution remote sensing-based measurements for crop yield estimation.

3. Analytical Framework

Following the standard econometric representation of measurement error (e.g., Bound et al., 2001), we consider the following relationship between a true outcome of interest Y^* (which represents the logarithmic transformation of crop production in our case) and the true value of a single explanatory variable, X^* (representing the logarithmic transformation of area cultivated in our case):

$$Y^* = \theta X^* + \varepsilon \tag{1}$$

We assume that the regression error term, ε, in equation (1) is mean zero and uncorrelated with the explanatory variable. For convenience, we will be working with logarithmic transformation of all values (mismeasured, correctly measured and measurement errors). Let us next assume that we do not observe the true measures of production and land area, rather we observe error-ridden self-reported measures, Y (logarithmic transformation of mismeasured production) and X (logarithmic transformation of mismeasured land), respectively. These (logarithmic transformed) self-reported measures can be represented as combinations of (logarithmic transformed) true production (land area) and logarithmic transformed measurement errors follow:[11]

$$Y = Y^* + u\,, \;\; X = X^* + v \tag{2}$$

In what follows, we demonstrate how the nature of the relationship between the measurement errors, u and v, and X^* affect estimates of the size-productivity relationship (SPR) parameter.

3.1 Size-Productivity Relationship

The above equation is not the regression that researchers commonly estimate in quantifying the relationship between plot size and productivity. Thus, we need to transform the standard regression in equation (1) into an estimable equation that can capture the relationship between productivity (yield) and land area cultivated. Recalling that both dependent and independent variables in equations (1) and (2) are logarithmic transformations of production and size (i.e.,

[11] This specification implies that measurement errors are assumed to be additive in their logarithmic transformed values and hence multiplicative in their original form.

$Y^{*} = \ln(production), X^{*} = \ln(area)$), we can transform equation (1) into the estimable equation of interest as follows:

$$\ln(yield) = \ln(production/area) = Y^{*} - X^{*} = (\theta - 1)X^{*} + \varepsilon \quad (3)$$

$$\ln(yield) = Y^{*} - X^{*} = \beta X^{*} + \varepsilon, \text{where } \beta = \theta \text{ -1} \quad (4)$$

We note that equation (4) is the workhorse estimable equation used in the SPR literature. Thus, we need to analytically show the consequences of alternative forms of measurement errors in either production or plot size on the β estimate in equation (4). The expression in equation (4) shows that land area enters both in the right and left-hand side of the regression, and hence associated measurement error in land area affects both the dependent and independent variables. However, given the relationship between equation (1) and (4), our representation and implications remain general. Following the empirical characterization of measurement errors in our data (see Sections 4 and 5) we analytically examine generic measurement problems that include four cases of NCME, cases where measurement error is correlated with either the true outcome or explanatory variables.

Case 1*: Non-classical measurement error in dependent variable (production), error correlated with dependent variable*

Assume that self-reported production suffers from measurement error that can be expressed as:

$$u = \delta Y^{*} + \omega \Rightarrow Y = (1+\delta)Y^{*} + \omega \quad (5)$$

Where ω is a random term uncorrelated with land area and the error term in equation (1). By substitution, we can show that OLS estimation of equation (4) using self-reported production would result in:

$$\beta^{OLS} = \frac{\text{cov}(Y - X^{*}, X^{*})}{\text{var}(X^{*})} = \frac{\text{cov}((1+\delta)(\beta+1)X^{*} + (1+\delta)\varepsilon + \omega - X^{*}, X^{*}))}{\text{var}(X^{*})} = (1+\delta)\beta \quad (6)$$

If we assume, as we find empirically, a negative correlation between measurement error in crop production and true production (i.e., those with the lowest harvest over-estimate output the most), then OLS using self-reported production weakens the estimated (inverse) relationship between land area and productivity. The degree of underestimation increases with higher correlation between measurement error and true production. Given the definition in equation (5), the proportional bias in equation (6) is equivalent to the correlation between the error-ridden production and measured production (Bound et al., 2001; Gibson and Kim, 2010).

Of course, a large share of the correlation between measurement error in self-reported production and true production may be driven by the correlation between measurement error in self-reported production and true land area.[12] This type of measurement error is more relevant and consequential in our context, at least in generating correlation across measurement errors in production and land area. Thus, the next case analyzes the implication of non-classical measurement error in self-reported production caused by correlation between measurement error in self-reported production and true land area, as will occur if smaller farms are more likely to overestimate output, as appears true in our data.

Case 2: *Non-classical measurement error in dependent variable (production), error correlated with independent variable*

Now let us assume that measurement error in self-reported production is correlated with true farm size and hence can be expressed as:

$$u = \lambda X^{*} + \zeta \tag{7}$$

Where ζ is random noise uncorrelated with the true value of farm size and the error term in equation (1). Using similar substitutions, we get the following expression:

$$\beta^{OLS} = \frac{\operatorname{cov}(Y - X^{*}, X^{*})}{\operatorname{var}(X^{*})} = \frac{\operatorname{cov}((\beta+1)X^{*} + \varepsilon + \lambda X^{*} + \xi - X^{*}, X^{*})}{\operatorname{var}(X^{*})} = \beta + \lambda \tag{8}$$

This implies that a negative correlation, which prevails in our dataset, between measurement error in self-reported production and farm size induces overestimation of the inverse relationship. Along this line of reasoning, Desiere and Jolliffe (2018) and Gourlay et al. (2017) provide empirical evidence showing that self-reported production measures can generate a spurious inverse relationship even when productivity is invariant with respect to area.

So far, we have seen two cases of measurement errors in production that may result in conflicting implications on the inverse relationship. The first case attenuates the inverse relationship while the second case amplifies the relationship. The overall net effect and implication

[12] This is always the case if production is a deterministic function of land area. If production is a probabilistic function of land area, as usually specified in regression production functions, we may theoretically disentangle the correlations between measurement error in self-reported production and measured production caused by land area as well as other (unobservable) factors.

potentially depends on the relative sizes and sources of these correlations between the measurement error and the true measures of production (land area). Since both patterns actually exist in our data, this implies that measurement error in output does not immediately generate a spuriously negative SPR, as recent studies claim based on their empirical findings.

Considering similar levels of correlations, the overestimation caused by Case (2) is expected to dominate the underestimation associated with Case (1). For example, in the absence of any relationship between farm size and productivity, non-classical measurement error driven by correlation between measurement error in production and true value of farm size can generate inverse relationship while the correlation between measurement error in self-reported production and true value of production cannot generate a relationship.

Case 3*: Non-classical measurement error in explanatory variables (plot size)*

Now consider measurement error in land area measurement and assume that we have a precise measure of production. We assume non-classical measurement error, with the size measurement error a function of the true area cultivated, which can be expressed as:

$$v = \alpha X^{*} + \ell \Rightarrow X = (1+\alpha) X^{*} + \ell \tag{9}$$

Where ℓ is uncorrelated with the error-free explanatory variable (true value of farm size) and the error term in equation (1). Let variance of $X^{*} = \rho x_{*}^{2}$ and variance of $\ell = \rho \ell^{2}$. Following similar substitutions and reformulations, we get the following identity:

$$\beta^{OLS} = \frac{\operatorname{cov}(Y^{*} - X, X)}{\operatorname{var}(X)} = \frac{\operatorname{cov}((\beta+1)X^{*} + \varepsilon - ((1+\alpha)X^{*} + \ell), ((1+\alpha)X^{*} + \ell))}{\operatorname{var}((1+\alpha)X^{*} + \ell)} \tag{10}$$

$$\beta^{OLS} = \frac{(\beta-\alpha)(1+\alpha)\rho x_{*}^{2}}{(1+\alpha)^{2}\rho x_{*}^{2} + \rho\ell^{2}} \tag{11}$$

$$\beta^{OLS} = \frac{\beta(1+\alpha)\rho x_{*}^{2}}{(1+\alpha)^{2}\rho x_{*}^{2} + \rho\ell^{2}} - \frac{\alpha(1+\alpha)\rho x_{*}^{2}}{(1+\alpha)^{2}\rho x_{*}^{2} + \rho\ell^{2}} \tag{12}$$

The expression in equation (12) is a generic representation of the consequences of NCME associated with explanatory variables, including those which can also appear in the left-hand side of regressions (dependent variable). The first term in equation (12) stands for cases where the explanatory variable only appears in the right-hand side of the equation, and this expression boils down to the usual attenuation bias if the measurement error associated with the explanatory variable is classical ($\alpha = 0$). Given the expression in equation (9), the first term in equation (12) is a function of the attenuation bias and the correlation between the correct and error-ridden measures of plot size. The second term in equation (12) arises if the explanatory variable (plot size in our context) also appears in the left-hand side (dependent variable) of the regression, as is true in the SPR literature because yield (i.e,. output per unit area) is the dependent partial productivity variable of interest. This whole term disappears if the measurement error behaves classically ($\alpha = 0$). This is consistent with the fact that classical measurement errors in dependent variables are innocuous.

Importantly, we cannot know *a priori* the direction of bias associated with self-reported land area measurement in equation (12). Indeed, we cannot even determine the direction of bias associated with the first term in equation (12), even when land area only appears in the right-hand side of the equation (Gibson and Kim, 2010). The direction of bias in the first term of equation (12) mainly depends on the relationship between the variances of self-reported and true area measurements as well as on the size (and sign) of the correlation between the measurement error and true area of land. Intuitively, there are cases where self-reported land measurement can be expected to have lower variance than the true area measure, for example, if rounding is the main source of measurement error.[13] In these cases, OLS estimation using self-reported farm size will overestimate the inverse relationship if the difference between the two variances is large enough relative to the negative correlation between the measurement error and true area of land. However, the second term in equation (12) renders ambiguous the overall effect of inaccurate land area measurement.

***Case 4**: Correlated non-classical measurement errors*

So far, we considered the implication of measurement error in one of the variables at a time. Now we relax this assumption and analytically show the implication of measurement errors both in production and area. Following our descriptive characterization of measurement errors in our data, we also assume correlations in measurement errors across both measures ($\text{cov}(u,v) = \pi$). Following analogous substitutions and reformulation of some relationships, we can show that OLS

[13] Our data (Table 2) show that the variance of the self-reported area measure is smaller than that of the true area measurement (also reflected through the negative correlation between measurement error associated with farm size and true farm size).

estimation of the size-productivity relationship using both self-reported measures yields the following identity:

$$\beta^{OLS} = \frac{\text{cov}(Y-X,X)}{\text{var}(X)} = \frac{\text{cov}((\beta+1)X^* + \varepsilon + u - ((1+\alpha)X^* + \ell), X^* + v)}{\text{var}((1+\alpha)X^* + \ell)} = \frac{(\beta-\alpha)(1+\alpha)\rho x_*^2 + \lambda\rho x_*^2 + \pi}{(1+\alpha)^2\rho x_*^2 + \rho\ell^2} \quad (13)$$

$$\beta^{OLS} = \frac{\beta(1+\alpha)\rho x_*^2}{(1+\alpha)^2\rho x_*^2 + \rho\ell^2} - \frac{\alpha(1+\alpha)\rho x_*^2}{(1+\alpha)^2\rho x_*^2 + \rho\ell^2} + \frac{\lambda\rho x_*^2}{(1+\alpha)^2\rho x_*^2 + \rho\ell^2} + \frac{\pi}{(1+\alpha)^2\rho x_*^2 + \rho\ell^2} \quad (14)$$

We note that equation (14) represents a general representation encompassing various types of measurement errors, classical and non-classical as well as those affecting the dependent and independent variables of interest. For example, the standard attenuation bias associated with classical measurement errors in the explanatory variable of interest (size in our case) can be shown by setting $\lambda = \alpha = \pi = 0$. Similarly, we can show that ignoring measurement error in self-reported production $(\lambda = 0)$ and correlations between both types of measurement errors ($\pi = 0$) results in the special case of equation (12).

Again, we are not able to identify the direction of bias from equation (14). However, we can intuitively extract the following insights. First, even in the absence of correlation in measurement errors ($\pi = 0$), the fact that both size and productivity suffer from non-classical measurement error ($\alpha \neq 0, \lambda \neq 0$) implies that correcting for measurement errors in one of the variables would not ensure unbiased estimates of the SPR.

Second, to understand the implication of correlations in measurement errors in self-reported production and area, we can apply simplifying assumptions to equation (14). If we momentarily assume that measurement errors in size and production are classical ($\alpha = 0, \lambda = 0$), then classical measurement error associated with farm size attenuates any relationship between area and productivity while the positive correlation between measurement errors leads to upward bias (and hence weakens the inverse relationship).[14]

Third, if we correct for measurement errors in one of our metrics, for example for plot size measurement, the expression in equation (14) boils down to equation (8) where the inverse relationship between plot-size and productivity would be inflated because of the usually negative

[14] In a slightly simplified similar (classical) setting, Bouis and Haddad (1992) and Subramanian and Deaton (1996) show that common measurement problems that affect the dependent and independent variables may create upward bias that can outweigh the attenuation bias associated with classical measurement error in the independent variable.

correlation between measurement errors in production and true plot size. Following the theory of the second best, this bias in the inverse relationship can be more consequential than ignoring both types of measurement errors, so correcting the one measurement error may aggravate the inferential problem, not solve it. This can be expected for cases where the correlations between measurement errors (the last term in equation) is positive and strong enough to dampen part of the overestimation in the inverse relationship caused the third term in equation (14).

Finally, we can judge the relative power of the different types of measurement errors and hence the parameters in equation (14). For example, assuming that there is no statistically significant relationship between farm size and productivity $(\beta = 0)$, then the first term in equation (14) disappears. Then with similar correlations between measurement errors and true area $(\alpha = \lambda)$, NCME in self-reported production can generate a spuriously negative SPR estimate (through the third term in equation (14)), while the positive correlation between measurement errors may generate a spuriously positive one. This suggests that in the presence of correlation between measurement errors in farm size and productivity, the strength of this correlation is a key parameter that may define the direction and size of the relationship between farm-size and productivity.

Overall, the generic analytical expression in equation (14) refines and qualifies recent studies that argue that measurement error in size or productivity spuriously generates the standard inverse SPR (Carletto et al., 2015; Holden and Fisher, 2013; Carletto et al., 2015; Gourlay et al., 2017; Desiere and Jolliffe, 2018). The intuitive expressions in equation (12) and (14) highlight the intricacies through which measurement errors in cultivated area and crop output affect this oft-explored hypothesis. Our analytical framework shows that predicting the direction of bias associated with self-reported size and production is more complex than the existing literature makes it seem. As in any other settings, adding more covariates to equation (1) also complicates the prediction of the direction of the biases, particularly if these covariates are correlated with cultivated area or the measurement errors, as will commonly be true for agricultural inputs such as labor, fertilizer and machinery use.

We note that some of the above analytical regularities are generic while some of them are specific to the SPR, where area cultivated appears on both sides of the regression. The general expressions in equation (12) and (14) apply to any regressions involving non-classical mismeasurement in outcome and explanatory variables as well as for cases where these measurement errors may be correlated. In Table 1 we summarize the key analytical results of this section.

4. Data and Descriptive Statistics

We employ an original experimental dataset collected from rural wheat framers in Ethiopia. The dataset covers 36 *kebeles* spanning in 18 *woredas* (or districts) in the Oromia, Amhara, and Tigray regions of the country.[15] The sampling design followed a four-stage approach.[16] In the first stage, 18 *woredas* were randomly selected from a sampling frame that was developed based on the *woredas'* wheat production potential. In the second stage, 2 *kebeles* per *woreda* were randomly selected. In the third stage, 14 farmers were randomly selected from each *kebele* using household lists maintained by the local *kebele* administrations. Lastly, one reference wheat plot was randomly selected for the experiment for sample households who cultivated wheat on multiple plots during the 2013/14 *meher* season.[17] We only have one reference plot, so the terms 'area' and 'size' all refer exclusively to the plot under study. Focusing on wheat farmers and wheat production helps alleviate other sources of mismeasurement that may arise from mixed cropping patterns.

This paper first relies on data collected from crop-cut measurement conducted in November and December 2013 by experts from the Central Statistical Agency (CSA). These involved both the measurement of the whole plot area using the compass-and-rope method, harvesting of one random subplot (4 meters × 4 meters) in each experimental plot, and weighing of the corresponding grain harvest. Of the 504 sampled farmers, crop-cut wheat production was successfully measured on 382 plots.[18] Second, a wheat growers' household survey was conducted in February and March 2014, after harvest was completed. The survey covered the same farmers, and gathered information related to input use, wheat production and commercialization behaviors.

Table 2 presents the summary statistics of main household and plot level characteristics. The first four rows provide alternative measures of plot size and production while the remaining

---

[15] *Kebele* is the smallest administrative unit in rural Ethiopia.
[16] See Abate et al. (2015) for detailed discussion on the sampling design.
[17] *Meher* is the long (main) rainy and production season in Ethiopia.
[18] Crop cuts could not be measured for the remaining 122 plots for three reasons. First, seven of the farmers had no wheat plot during the 2013 *meher* season. Second, five farmers could not be identified by anyone in their respective *kebeles* at the time of the household survey. Third, the remaining 110 farmers harvested their wheat plots early before the crop-cut survey. There were no refusals. In Table A1, we show that these nonresponses are not systematic and hence uncorrelated with the household and plot characteristics.

rows report household and plot characteristics. In particular, we consider detailed plot level characteristics that might confound accurate measurement of plot size and production. As shown in the top of the table, there are significant discrepancies between self-reported and objective measures of land area and production.

Table 2 shows that sampled plots have, on average, about nine corners, indicating that precise measurement of such plots using scientific methods can also be difficult. Nonetheless, the closure error is one percent, on average.[19] About 40 and 60 percent of the sample households used standard units for reporting their plot size and production, respectively. One can argue that the use of standard units (e.g., kg or ha) may introduce considerable errors since these measurements might not be commonly used in some rural areas. On the other hand, local measurement units are likely to vary between regions, villages and even farmers. For this reason, we control for these measurement units in our empirical characterizations of measurement errors.

5. Descriptive Results and Characterization of Measurement Errors

5.1 Measurement Error in Self-Reported Plot Area

Farmers' self-reported estimates and traversing (also known as compass-and-rope) are the two conventional methods of measuring the surface of land area. With the advent of new technologies, there are now alternative ways of measuring land area — i.e., GPS and remote sensing (see Carletto et al., 2015 and Carletto et al., 2016 for detailed discussion on these methods). While farmers' self-reported area measurement through household surveys are the most cost-effective method of estimating plot area, they can be subject to considerable measurement errors. Self-reported land area measurements are commonly based on traditional units whose conversion factor varies across regions and hence introduces some errors. For example, farmers in Ethiopia commonly measure and report land areas in *oxen days.* But that measure will necessarily vary with weather conditions, slope, drainage and texture of soils, animal breed and condition, etc.

On the other hand, compass-and-rope is the most reliable method to accurately measure land area (Keita and Carfagna, 2009; Fermont and Benson, 2011; Carletto et al., 2015; Carletto et al., 2016). Compared to GPS-based area measurement, the compass-and-rope method is expensive,

[19] Closure error is the shortest line of unknown length and direction connecting the initial and final station of the polygon or traverse. When the closing error is larger than 3 percent of the perimeter of the polygon, repeating the measurement procedure is highly recommended (Casley and Kumar, 1988)

though some argue that the level of accuracy is worth the extra time and cost (Diskin, 1997), in particular because GPS-based area measurement may be imprecise for smaller plots (Schoning et al., 2005; Keita and Carfagna, 2009; Fermont and Benson, 2011; Carletto et al., 2015).

Table 3 reports the average sizes of plots based on self-reported measures and compass-and-rope method. The discrepancy in measurement between the two methods is sizeable and statistically significant—i.e. the bias in self-reported land area measure amounts 0.05 hectares. This is a substantial difference given the sizes of agricultural plots in Ethiopia, amounting to 14 percent of the mean size of plots. Disaggregating the analysis by plot size categories (constructed based on the actual plot sizes) shows that the bias in self-reported area measurement declines with farm size, ranging from an overestimation of 150% in the category of smallest plots to an underestimation of 28% in the category of largest plots. These differences are statistically significant at the bottom and top of the distribution, and non-significant towards the middle where differences are negligible.

Next, we explore potential sources of mismeasurement in self-reported plot sizes. Figure 1 provides joint distributions of the compass-and-rope and self-reported plot sizes. As one can see, the distributions are remarkably distinct, the distribution of the compass-and-rope plot size measure being much smoother than the distribution of self-reported plot size. Figure 1 further hints that rounding can be a key component of the biases in self-reported plot size, as evidenced by clear heaping on values that correspond to the conversion factor between the common local unit and hectare (e.g. ½ *oxen day*=0.125 ha; 1 *oxen day*=0.25 ha) in the distribution of self-reported plot sizes.

Rounding-based biases can be more clearly seen in Figure 2 below, where panel (a) shows horizontal bunching of self-reported land areas on some of the values corresponding to Figure 1. With panel (b) we further investigate whether rounding is the only explanation, by exploring the distribution of the compass-and-rope plot size measures at the self-reported plot size rounding intervals. Unless farmers systematically round up or down, one should expect a symmetrical distribution of true land size around the self-reported value, for those households who reported the most typical rounding numbers for their plots. However, we find strikingly asymmetric distributions across the self-reported rounding levels, showing heavier right tails for small plot sizes and heavier left tails for large plot sizes. Moreover, the interval distributions of actual land size strongly increase as the self-reported rounded off plot size increases.

Overall, panel (b) of figure 2 suggests that other factors may also cause mismeasurements in self-reported plot size estimates. In the Ethiopian context, mismeasurement of plot sizes could emanate in part from the traditional measurement units of land itself. Oxen days (*timad*) is the most common unit of area measurement and can be subject to a wide range of errors, including biases from differences in length of working hours and traction capacity of oxen and in weather conditions, as well as plot characteristics (e.g., slope, soil texture and drainage, etc.). Moreover, some of those same plot characteristics (along with shape, fertility, and ownership of the plot) and household characteristics can affect farmers' estimation of plot size. We assess the effect of these factors in our characterization of measurement errors in the subsequent section.

5.2 Measurement Error in Self-Reported Production

Crop-cuts and farmer self-reported estimates are the two methods used to measure production in developing countries. The crop-cut method is based on harvesting one or multiple random subplots in each plot. The method involves randomly locating a sub-plot(s) prior to the harvest and the subplot(s) will be harvested by survey enumerators at the time of maximum crop maturity. Then, the harvest is processed (e.g. dried) and weighed. Total plot level production is then estimated by extrapolating the sampled crop production. Crop-cuts are commonly regarded as the most reliable and unbiased method for estimating crop production (Fermont and Benson, 2011). However, obtaining production estimates through crop-cuts can be costly—i.e. it is both time and labor-intensive undertaking. Instead, estimating crop production based on farmers' self-report is more common in agricultural surveys, as it can be easily incorporated in standard household surveys. Recently, high-resolution satellite imagery-based remote sensing techniques are also being used to estimate crop yield, with some promising results (e.g., Lobell et al., 2015; Gourlay et al., 2017).

Table 4 compares the mean values of production obtained both from crop-cuts and farmers' self-reports, by plot size groups. As described in Section 3, the crop-cut production is an extrapolated value based on the production obtained from 16 m^2 random sub-plot to the whole plot as measured by compass-and-rope. We note that while crop-cut is considered to be most reliable and objective production measurement, the extrapolation may introduce errors due mainly to variations in the productivity of plot parts (e.g. interior vs. periphery or edge).[20] However, these

[20] For example, previous agronomic studies indicate that the periphery of a plot is often more productive than its interior (Little and Hills, 1978; Barchia and Cooper, 1996; Ward et al., 2016). More recently, Bevis and Barrett (2017) argue that this could be one explanation for the inverse size-productivity relationship. We explore that hypothesis

types of errors are expected to be uncorrelated with true measures of production and relevant explanatory variables, implying that they should have little consequence in estimating the farm-size and productivity relationships. We also account for crop-cut distance to edges to minimize such problems, if any. Self-reported production estimates were collected as part of a household survey shortly after harvest.

Looking at the overall mean values at the bottom of Table 4, the discrepancy between the two production measures is substantial and statistically significant. Moreover, disaggregating the analysis by farm size categories indicates systematic differences between the two production measures—i.e. the bias in self-reported production measurement decreases with plot size.

In Figure 3 we report the joint distribution of self-reported and crop-cut production estimates. From the bar graphs, one finds important clustering on the lower end of the distribution on both graphs, albeit the crop-cut production graph is more clustered. This is consistent with our expectation since wheat productivity in the country is low. In the self-reported productivity measure, however, there is clear heaping on (close to) the five scale levels and the distribution is heavier from the middle to high ends of the scale, than in crop-cut production bar charts. This support the over-reporting trend presented in Table 4.

Figure 4 further explores potential sources of measurement error in self-reported production. Panel (a) corroborates the important discrepancy between the two production estimates. Panel (b) further reports the distribution of the crop-cut production estimates at the self-reported production rounding levels (at each five scale level points) to assess whether rounding is the main source of the over-reporting patterns. As in Figure 2, the distributions are largely asymmetric across the self-reported rounding levels, showing heavier right tails for distribution on the low end of the scale. Moreover, the distributions of the crop-cut production at the self-reported production intervals encompass a large range of crop-cut production values, particularly for values on the high end of the scale. Overall, over- and under-reporting across the intervals indicates that rounding is not the only factor that leads to systematic over-reporting of self-reported production.

below.

5.3 Characterization of Measurement Errors in Production and Land Area

In this section, we parametrically characterize measurement errors (defined as self-reported minus true measures) in production and land area. We mainly conduct the following three empirical exercises: (i) investigate whether measurement errors in self-reported production and area measurements are correlated with true production (crop-cuts) and true area measurements (compass-and-rope); (ii) estimate correlations in measurement errors associated with self-reported production and plot size measurements; (iii) estimate potential correlations between measurement errors and explanatory variables commonly used in household surveys. We consider household and plot level characteristics that may affect farmers' perception of plot size and production. By doing so, we can understand whether measurement errors associated with self-reported production and land area measurements are generally classical or not. These exercises can inform the consequences of ignoring measurement errors in one or more of these variables.

We estimate unconditional and conditional correlations between measurement errors in self-reported production (land area) and crop-cut production (compass-and-rope) measurements. In all our characterizations, we control for village (*kebele*) fixed effects to capture village-level misreporting arising from local measurement units and enumerator bias, since local units vary across villages and different villages were surveyed by different enumerators.

Table 5 reports regression results characterizing measurement error in production, expressed as differences in logarithmic values of self-reported production and crop-cut production (ln(self-reported)-ln(crop-cut)). Column 1 of Table 5 characterizes this measurement error as a function of crop-cut production and *kebele* fixed effects, while the remaining columns extend this specification by adding other household and plot level characteristics.

Table 5 clearly shows that measurement error in self-reported production is strongly and negatively correlated with crop-cut production, suggesting the type of mean-reverting measurement error documented in earnings (Bound and Krueger, 1991) and consumption (Gibson et al., 2015). This negative correlation remains robust to the inclusion of many household and plot level characteristics. Measurement error in self-reported production is also correlated with soil quality. These pieces of evidence broadly suggest that measurement errors associated with self-reported production are non-classical and hence can be expected to be more consequential in estimating the conventional farm size-productivity relationship. Following our analytical framework in Section 3, this may lead to underestimation of the inverse relationship.

Table 6 provides a slightly different characterization of measurement error in self-reported production as a function of true farm size. These results show that measurement error in self-reported production is also correlated with true measure of farm size (compass-and-rope). These correlations are much higher than those reported in Gourlay et al. (2017) and Desiere and Jolliffe (2018). As shown in our analytical framework, this correlation between measurement error in the outcome variable and true explanatory variable induces overestimation of the inverse relationship. However, comparing the sizes of correlations in Table 5 and 6, it seems that both types of correlations are mainly driven by the pattern in Table 6 (the correlation between measurement error in production and true plot size). To probe this, we also had some robustness exercises which show that one of the correlations disappears when we control for both objectively measured production and plot size.[21] This suggests that the overestimation of the inverse relationship caused by the correlations in Table 6 (corresponding to Case 2 of our analytical section) may dominate the underestimation caused by the correlations in Table 5 (corresponding to Case 1 of our analytical section). Comparing the other correlations (between measurement error in production and other explanatory variables) in Table 5 and 6, reveals that once we control for plot size all significant correlations in Table 5 are statistically insignificant (Table 6). This again reinforces our argument that the correlation between measurement error in production and plot size drives most of the significant correlations in Table 5.

Table 7 provides similar characterizations of measurement errors in self-reported farm size. To facilitate comparison of estimates, we are restricting our sample to those plots for which crop-cuts are available and (almost identical) full sample results are given in the Appendix (Table A5). These regressions show that measurement error in self-reported land area is negatively correlated with true farm size (measured by compass-and-rope method). This correlation remains unaffected by the inclusion of many household and plot level characteristics. The size of this correlation between measurement error and true farm size is larger than those reported by Carletto et al. (2013) and Carletto et al. (2015). One potential explanation for these differences could be related to the land area measure we are using in this paper. Carletto et al. (2013) and Carletto et al. (2015) as well as other previous studies investigating measurement error in plot size use GPS-based land area measurement, which might be susceptible to some systematic biases (especially when it comes to smaller plots), while we are using a method commonly considered as the most

[21] These regression results are not reported to conserve space.

accurate method to estimate land area. As shown in equation (12) the size of this correlation along with the other parameters in equation (12) crucially defines the implication of mismeasurement in land area.

Finally, in Table 8 we estimate correlations between both types of measurement errors. We regress measurement error in self-reported production as a function of measurement error in farm size and other covariates, particularly those which appear to be correlated with either type of biases in Tables 5-7. We can observe that both types of measurement errors are strongly and positively correlated. Although including true land area seems to slightly affect the size of the relationship, the correlations remain strong. These correlations along with our analytical framework in Section 3 suggest that correcting for either type of measurement errors may not sufficiently alleviate potential biases in estimating the relationship between productivity and plot size. However, as shown in equation (14) of our analytical framework, the implication of this correlation on estimating the size-productivity relationship depends on many parameters including the strength of this correlation in measurement errors.

Due to the skewed distribution of some of our variables (e.g., plot size), we also re-estimate the above regressions that characterize measurement errors using the inverse hyperbolic sine transformation of our main variables of interest and found similar results (see Tables A2 – A4 in the Appendix).[22]

To sum up, the empirical characterization of measurement errors in self-reported production and farm size uncover four cases that may lead to biased estimates of the SPR: (i) non-classical measurement error in self-reported production caused by negative correlation between measurement error and crop-cut production; (ii) correlation between measurement errors in self-reported production and true plot size; (iii) non-classical measurement error in self-reported plot size, caused by negative correlation between the bias in plot size and its true value; and (iv) positive correlation between measurement errors in self-reported production and plot size. As our analytical framework shows, each of these generates predictable bias in the estimated SPR. However, with several opposing biases introduced simultaneously via measurement errors, the net effect of these measurement errors is ambiguous.

[22] The inverse hyperbolic sine transformation better handles extreme values than the commonly used log transformation (Burbidge et al., 1988). In our case, it also overcomes potential expansion of the heterogeneity of the distribution of biases for values between 0 and 1 due to the log transformation.

6. Measurement Errors and the Estimated Size—Productivity Relationship

In this section, we estimate the SPR considering the four measurements and scenarios. We particularly estimate the following production function:

$$\ln(production/area)_h = \beta_0 + \beta_1 \ln(area)_h + \beta_2 hhold_h + \beta_3 plot_h + village + \varepsilon_h \quad (15)$$

Where *production* represents self-reported or crop-cut output measured in quintals. Similarly, *area* stands for either self-reported or compass-and-rope measures of cultivated area. Thus, the relationship between plot size and productivity can be captured by estimating the scalar parameter β_1, while the vectors β_2 and β_3 quantify the implications of other household and plot level characteristics, respectively, and ε_h captures mean zero error.

We run both unconditional and conditional regressions. In all estimations, we cluster standard errors at village (*kebele*) level to capture correlation of errors terms at village level, for instance, due to our reliance on locally defined measurement units. Table 9 provides benchmark estimates of the true relationship between plot size and productivity, using crop-cut production and plot size measurement based on compass-and-rope. The unconditional correlations between plot size and productivity appear to be negative and statistically significant, albeit modest in magnitude—i.e., a 1 percent increase in plot size is associated with only a 0.08 percent reduction in productivity. However, this correlation becomes statistically insignificant when we control for plot and soil characteristics. Our proxy for the edge effect that Bevis and Barrett (2017) hypothesize could explain the inverse SPR, distance of crop-cut from the edge, is statistically insignificant.

Next, we estimate similar relationships, but without accounting or correcting for measurement errors in one or more of the key variables of interest (plot size and production). Table 10 reports the estimated relationships using self-reported production and plot size measured by compass-and-rope.[23] Recall our analytical representation of this case of measurement error only in the outcome variables, in equations (5-8). In our descriptive results, we showed that measurement error in self-reported crop production is correlated with crop-cut production as well as true plot size. The former correlation is expected to underestimate the inverse relationship while the latter overestimates the inverse relationship. As shown in Tables 6 and 7, it appears that the

[23] Full sample results, including those plots for which crop-cut production is missing, are given in Table A6 (in the Appendix). The empirical relationships are almost identical to those in Table 10.

latter effect dominates. Thus, we can expect OLS to overestimate the inverse relationship in this scenario. The results in Table 10 show that using self-reported production leads to substantial overestimation of the inverse relationship. This corroborates recent findings by Gourlay et al. (2017) and Desiere and Jolliffe (2018), which similarly show that self-reported production measures can generate inverse relationship relationships even in the absence of it.

Then in Table 11 we estimate similar relationships between plot size and productivity, but now without correcting for measurement error in plot size, only correcting the output measure. In our analytical framework, we showed that measurement errors in plot size can have ambiguous consequence in estimating the relationship between plot size and productivity. As shown in equation (12) the direction of bias associated with measurement in plot size depends on the relationship between the variance of self-reported and true area measurements as well as on the size (and sign) of the correlation between the measurement error and true area of land. Our descriptive statistics (Table 2) indicate that variance of the self-reported plot size is smaller than that of the true area, implying a negative correlation between measurement error in plot size and true land area measure. Thus, we may expect OLS estimation using self-reported plot size to overestimate the inverse relationship. The results in Table 11 confirm this analytical prediction. It is also consistent with the pattern reported by Carletto et al. (2015) but in contrast to the results in Carletto et al. (2013). This evidence implies that the implication of measurement error in plot size may vary across contexts and sources of measurement errors, again highlighting the cost of inaccurate land measurements.

Finally, in Table 12 we estimate the relationship between plot size and productivity, without correcting measurement errors in area and production.[24] These relationships are estimated using self-reported plot size and production. The estimated relationships in Table 12 suggest a significant inverse relationship between plot size and productivity. However, the size of this inverse relationship is significantly smaller than those in Tables 10 and 11, implying that some features of measurement errors drag the bias in opposite directions. This is consistent with our analytical expression in equation (14), which shows that positive correlation of measurement errors in the dependent and independent variables may cancel out part of the bias due to measurement error in the dependent or independent variable(s). Apparently, because of the opposing direction

[24] Full sample results, including those plots for which crop-cut production is missing, are given in Table A7 (in the Appendix). The relationships between plot size and productivity are similar to those in Table 12.

of bias, ignoring both types of measurement errors appear to bias the parameter of interest less (i.e., less overestimate of the inverse relationship) than does controlling for either source of measurement error alone. This underscores the hazard in partial correction of multiple, correlated, NCME. Table 13 below summarizes the key empirical relationships considering the alternative empirical scenarios.

Comparing the other estimates associated with the other explanatory variables in our regressions, we also observe some importance differences among Tables 9-12 in other parameter estimates of interest. For example, our measures of soil quality (soil fertility and soil color) are significantly associated with productivity when we use correct measures of plot size and productivity, while this is not the case when using self-reported measures. This is theoretically anticipated as some of these soil quality indicators are correlated with the measurement errors in production and plot size (see Tables 5 and 6). This is consistent with previous studies arguing that omitted attributes, including unobservable soil quality, may contribute to the disputed inverse size-productivity relationship (Benjamin, 1995; Assuncao and Braido, 2007). Similarly, some plot characteristics (number of corners and crop-cut distance to the edge) appear to be significant only when we use crop-cut production along with self-reported plot size (Table 11). These spurious correlations between productivity and plot characteristics are potentially driven by farmers' misperception of plot size and associated endogenous investments. These types of behavioral mechanisms are discussed by Bevis and Barrett (2017). More generally, these pieces of evidence suggest that the implication of measurement errors in size and production may go beyond the inverse relationship and hence affect other relationships and inferences.

7. Concluding Remarks

We analytically investigate correlated NCME in both dependent and independent variables within a standard regression framework. We set up a generic analytical framework where both dependent and explanatory variables can suffer from NCME and these errors can be correlated. We show that the signs and magnitude of resulting biases are ambiguous and depend on several parameters characterizing measurement errors in these variables as well as the relationship under investigation. We also show that accounting for measurement error in only one of the variables may worsen the bias in estimated parameters.

We use this framework to shed further light on the longstanding policy debate about the relationship between plot size and agricultural productivity. This relationship has considerable implications for agricultural development policy and previous, widespread findings of an inverse relationship were widely invoked to support land reform programs. Most previous empirical studies rely, however, on farmer self-reports of output and area cultivated, with considerable room for NCME. And while recent studies have attempted to correct biases on either one of the variables (e.g., through GPS devices for area cultivated, or crop-cuts for production), none to our knowledge has investigated the relationship more generally, by addressing measurement issues on both sides of the equation, nor explored the implications for incomplete correction for correlated NCME.

We rely on a unique dataset combining self-reported and gold standard measurements of both agricultural output and area cultivated in Ethiopia. These data enable us to empirically validate our analytical results, showing that the inverse size-productivity relationship that we find in the self-reported data vanishes with more accurate measures. We also find that fixing measurement error in just one of the variables does not solve the problem and may effectively worsen bias in the parameter estimate of interest. These findings carry strong implications, not only for work that relies on conventional survey data, but also for studies that correct incompletely for measurement issues on both sides of the equation, which may prove inferior to a "second best" approach that uses multiple variables measured with error. These findings are relevant to many economic applications and estimation problems involving multiple error-ridden variables. It may also be relevant to aggregate metrics constructed from multiple variables suffering from competing sources and patterns of bias.[25]

[25] For example, Arthi et al. (2018) show that aggregating households' labor use involves competing biases, over-reporting at the extensive margin of labor use and under-reporting at the intensive margin, with these errors ultimately cancelling each other out to minimize aggregate bias.

References

Abate, G. T., de Brauw, A., Minot, N., Bernard, T., 2015. The impact of the use of new technologies on farmers' wheat yield in Ethiopia: Evidence from a randomized controlled trial. IFPRI Discussion Paper 1462. Washington, D.C.

Ali, N., Kahn, K., Patel, P., Gorelick, J., 2009. Moving the Learning Forward: From Incremental to Transformational Impact on Empowering Smallholder Farmers and Women. Impact Assessment of the Agricultural Marketing Initiative, West Nile, Uganda. Capstone Report. George Washington University, Washington, DC.

Arthi, V., Beegle, K., De Weerdt, J., Palacios-López, A., 2017. Not your average job: measuring farm labor in Tanzania. *Journal of Development Economics* (forthcoming).

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

Baird, S., Özler, B., 2012. Examining the reliability of self-reported data on school participation. *Journal of Development Economics*, *98*(1), 89-93.

Barchia, I., Cooper, J.L., 1996. Edge row effects on yields in an oat trial. *Animal Production Science* 36(4), 459–461.

Barrett C.B., Bellemare, M., Hou, J., 2010. Reconsidering Conventional Explanations of the Inverse Productivity–Size Relationship. *World Development* 38 (1), 88–97.

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

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

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

Bevis, L., Barrett, C. B., 2017. Close to the edge: do behavioral explanations account for the inverse productivity relationship? Paper Presented at the 2017 Centre for the Study of African Economies (CSAE) Conference.

Borjas, G. J., 1980. The relationship between wages and weekly hours of work: The role of division bias. *Journal of Human Resources* 15(3), 409-423.

Bouis, H.E., Haddad, L.J., 1992. Are estimates of calorie-income elasticities too high? A recalibration of the plausible range. *Journal of Development Economics* 39 (2), 333–364.

Bound, J., Brown, C., Duncan, G.J., Rodgers, W. L., 1994. Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics* 12 (3), 345–368.

Bound, J., Brown, C., Mathiowetz, N., 2001. Measurement error in survey data, in Heckman J. and Leamer E. (eds), *Handbook of Econometrics*, Vol. 5, Amsterdam: Elsevier, pp. 3705–3843.

Bound, J., Krueger, A. B., 1991. The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right? *Journal of Labor Economics* 9 (1), 1–24.

Burbidge J.B., Magee L., Robb L., 1988. Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association* 83, 123-127.

Carletto, C., Gourlay, S., Winters, P., 2015. From guesstimates to GPStimates: land area measurement and implications for agricultural analysis. *Journal of African Economies* 24(5), 593–628.

Carletto, C., Savastano, S., Zezza A., 2013. Fact or Artefact: The Impact of Measurement Errors on the Farm Size–Productivity Relationship. Journal of Development Economics, 103: 254–61.

Carletto, G., Gourlay, S., Murray, S., Zezza, A., 2016. Cheaper, faster, and more than good enough: is GPS the new gold standard in land area measurement? World Bank Policy Research Working Paper 7759.

Casley, D. J., Kumar, K., 1988. *The Collection, Analysis and Use of Monitoring and Evaluation Data*. Baltimore, MD: Johns Hopkins University Press.

Chao, L.W., Szrek, H., Peltzer, K., Ramlagan, S., Fleming, P., Leite, R., Magerman, J., Ngwenya, G.B., Pereira, N.S., Behrman, J., 2012. A comparison of EPI sampling, probability sampling, and compact segment sampling methods for micro and small enterprises. *Journal of Development Economics*, *98*(1), 94-107.

Das, J., Hammer, J., Sánchez-Paramo, C., 2012. The impact of recall periods on reported morbidity and health seeking behavior. *Journal of Development Economics*, *98*(1), 76-88.

David, I. P., 1978. Non-sampling Errors in Agricultural Surveys. Review, Current Findings, and Suggestions for Future Research. Paper presented at the Philippine Statistical Association Annual Conference, Manila, Philippines.

De Groote H., Traoré O., 2005. The Cost of Accuracy in Crop Area Estimation. *Agricultural Systems*, 84 (1), 21–38.

De Janvry, A., Sadoulet, E., 2000. Rural poverty in Latin America: dterminants and exit paths. *Food Policy* 25 (2000) 389–409.

Desiere, S., Jolliffe, D., 2018. Land productivity and plot size: Is measurement error driving the inverse relationship? *Journal of Development Economics* (forthcoming).

Diewert, W.E., Fox, K.J., 1999. Can measurement error explain the productivity paradox? *Canadian Journal of Economics*, *32*(2), 251-280.

Diskin, P., 1997. Agricultural Productivity Indicators Measurement Guide. Food and Nutrition Technical Assistance Project. Washington: US Agency for International Development.

Feder, G., 1985. The Relation Between Farm Size and Farm Productivity: The Role of Family Labor, Supervision and Credit Constraints. *Journal of Development Economics* 18 (2–3), 297–313.

Fermont A., Benson, T., 2011. Estimating Yield of Food Crops Grown by Smallholder Farmers, A Review in the Uganda Context. IFPRI Discussion Paper 01097, Washington, D.C.

French, E., 2004. The labor supply response to (mismeasured but) predictable wage changes. *Review of Economics and Statistics*, *86*(2), 602-613.

Gibson, J., Kim, B., 2010. Non-classical measurement error in long-term retrospective surveys. *Oxford Bulletin of Economics and Statistics*, Vol. 72, pp. 687–695.

Gibson J., Beegle K., De Weerdt J., Friedman J., 2015. What Does Variation in Survey Design Reveal about the Nature of Measurement Errors in Household Consumption? *Oxford Bulletin of Economics and Statistics* 77 (3): 466–474.

Gourlay, S., Talip, K., David, L., 2017. Could the Debate Be Over? Errors in Farmer-Reported Production and Their Implications for Inverse Scale - Productivity Relationship in Uganda. Paper Presented at the 2017 World Bank Land and Poverty Conference, March 23, 2017, Washington, DC.

Gottschalk, P., Huynh, M., 2010. Are Earnings Inequality and Mobility Overstated? The Impact of Non-Classical Measurement Error. *Review of Economics and Statistics* 2(2): 302–315.

Holden, S. T., Fisher, M., 2013. Can Area Measurement Error Explain the Inverse Farm Size Productivity Relationship? CLTS Working Paper No. 12, Norwegian University of Life Sciences, Ås, Norway.

Howard, J., Said, A., Molla, D., Diskin, P., Bogale, S., 1995. Towards Increased Domestic Cereals Production in Ethiopia: Using a Commodity Systems Approach to Evaluate Strategic Constraints and Opportunities. Michigan State University.

Hyslop, D.R., Imbens, G.W., 2001. Bias from classical and other forms of measurement error. *Journal of Business & Economic Statistics*, *19*(4), 475-481.

Keita, N., Carfagna, E., 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.*

Kim, B., Solon, G., 2005. Implications of mean-reverting measurement error for longitudinal studies of wages and employment. *Review of Economics and Statistics* 87, 193–196.

Larsen, A.F., Headey, D., Masters, W.A. 2017. Misreporting Month of Birth: Implications for Nutrition Research. IFPRI Discussion Paper 1617.

Lipsey, R. G., Lancaster, K. 1956. The General Theory of Second Best. *Review of Economic Studies* 24 (1), 11–32.

Little, T.M., Hills, F.J., 1978. *Agricultural experimentations– Design and analysis*. John Wiley and Sons. New York, NY, USA.

Lobell, D. B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. *Remote Sensing of Environment* 164, 324–333.

McKenzie, D., Rosenzweig, M., 2012. Preface for symposium on measurement and survey design. *Journal of Development Economics* 98(1):1-2.

Naschold, F., Barrett, C.B., 2011. Do short-term observed income changes overstate structural economic mobility? *Oxford Bulletin of Economics and Statistics*, *73*(5), 705-717.

Özler, B., 2013. Economists have experiments figured out. What's next? (Hint: It's Measurement). World Bank Development Impact blog. https://blogs.worldbank.org/impactevaluations/economists-have-experiments-figured-out-what-s-next-hint-it-s-measurement.

Schoning P., Apuuli, E., 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. Oslo: Statistics Norway.

Sen A. K., 1966. Peasants and Dualism with or Without Surplus Labor. *Journal of Political Economy* 74 (5), 425–50.

Subramanian, S., Deaton, A., 1996. The Demand for Food and Calories. *Journal of Political Economy* 104 (1), 133–62.

Ward, B., Roe, B. E., Batte, M.T., 2016. Potential Profitability of Strip Intercropping with Corn and Soybeans. *Journal of the ASFMRA.*

Tables and Figures

Figures

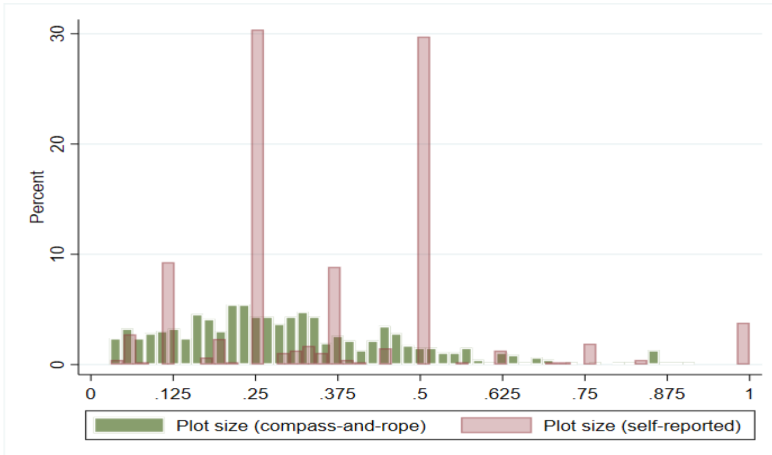

Figure 1: Distributions of Compass and Rope (CR) and Self-Reported (SR) plot areas

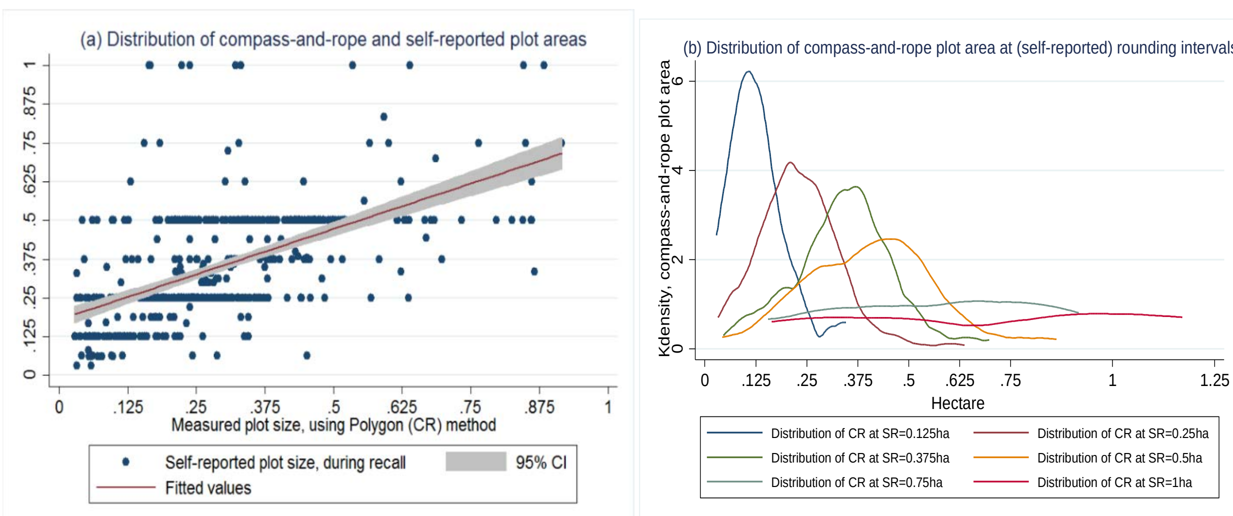

Figure 2: Distributions of Compass and Rope (CR) Areas at Self-Reported (SR) Rounding Intervals

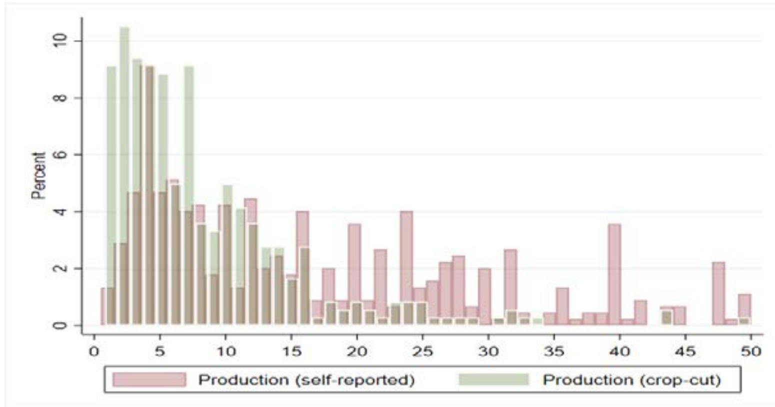

Figure 3: Distributions of Crop-Cut (CC) estimated and Self-Reported (SR) production

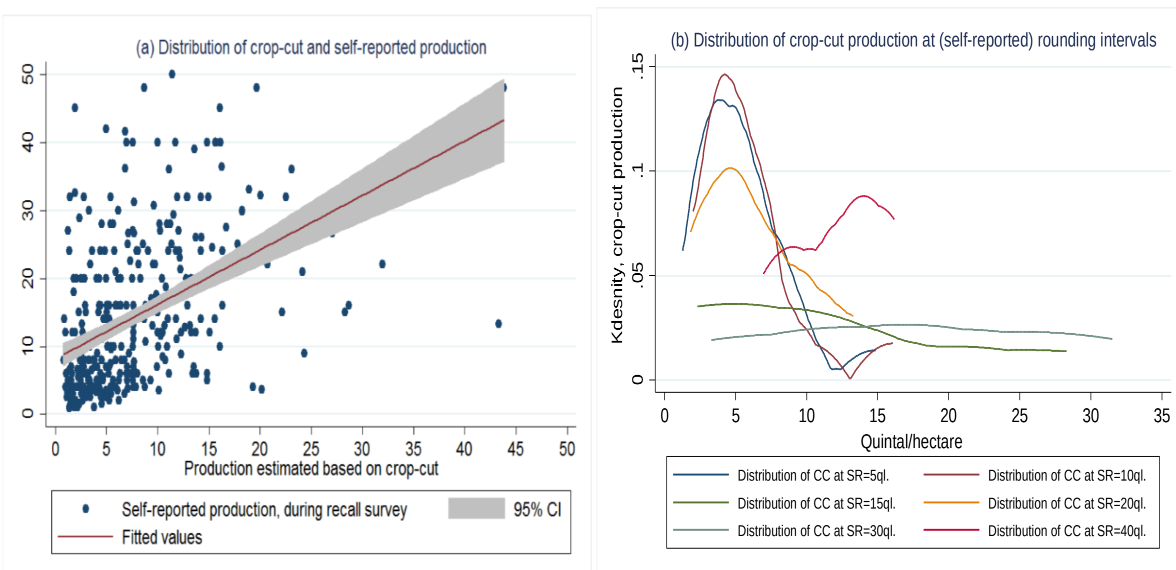

Figure 4: Distributions of Crop-Cut (CC) Production at Self-Reported (SR) Rounding Intervals

Tables

Table 1: Summary of Analytical Results

Source of non-classical measurement error	Key Parameters δ	λ	α	π	Estimated SPR	Direction of bias on the SPR
No error	0	0	0	0	β	No bias
Error in production	<0	0	0	0	$(1+\delta)\beta$	Underestimation of ISPR
Error in production	*	<0	0	0	$\beta+\lambda$	Overestimation of ISPR
Error in plot size	*	0	<0	0	$\beta(1+\alpha)\Phi-\alpha(1+\alpha)\Phi$	Ambiguous
Error in both	*	<0	<0	0	$\beta(1+\alpha)\Phi-\alpha(1+\alpha)\Phi-\lambda\Phi$	Ambiguous
Error in both	*	<0	<0	>0	$\beta(1+\alpha)\Phi-\alpha(1+\alpha)\Phi-\lambda\Phi+\pi\Phi$	Ambiguous

Notes: we have used our empirical data and analysis to get an insight of the sign of the key parameters of interest.

$$\Phi = \operatorname{var}(X^*)/\operatorname{var}(X) = \frac{\rho x_*^2}{(1+\alpha)^2 \rho x_*^2 + \rho \ell^2}.$$

* refers that the value of these parameter can be zero or negative. SPR stands for the size-productivity relationship while the acronym ISPR represents the inverse size-productivity relationship.

Table 2: Summary Statistics

Variable	Description	Mean	Std. Dev.	Min	Max	Obs.
Area SR	Self-reported area size (ha)	0.42	0.36	0.03	4.00	488
Area measured	Measured area size during crop-cut (ha)	0.37	0.39	0.03	3.80	483
Production SR	Self-reported production for reference plot (qt.)	21.05	19.18	0.50	120.00	488
Production measured	Estimated production based on crop-cut (qt.)	8.98	9.91	0.81	101.5	365
Yield SR	Self-reported (production/area), (qt./ha)	30.69	18.18	1.00	96.00	488
Yield measured	Measured (production/area), (qt./ha)	28.23	15.05	2.78	95.38	366
Age of HH head	Age of the household head in completed years	45.67	10.84	20.00	77.00	488
Gender of HH head	Gender of the household head	0.86	0.34	0.00	1.00	488
HH size	Number of household members	6.79	2.39	1.00	16.00	488
Literacy of HH head	=1 if the household head is literate	0.64	0.48	0.00	1.00	488
No. of corners	Number of corners of the reference plot	8.74	4.88	4.00	23.00	484
Closure error	Closure error in plot area measurement	1.09	0.89	0.02	4.50	483
Area unit†	=1 if farmers used ha for SR area measurement	0.39	0.49	0.00	1.00	488
Total owned area†	Total farm land owned by sample farmers	2.31	2.14	0.00	20.00	488
Crop-cut to edge	Distance between the crop-cut and shortest or closest plot edge (meters)	25.83	18.57	1.40	148.00	374
Production unit	=1 if farmers used kg for SR production measurement	0.59	0.49	0.00	1.00	488
Total wheat produced†	Total wheat production during 2013/14 *meher*	46.64	75.26	0.95	755.00	488
Soil fertility†						
High	=1 if the fertility of the reference plot is high	0.44	0.49	0.00	1.00	488
Medium	=1 if the fertility of the reference plot is medium	0.49	0.50	0.00	1.00	488
Poor	=1 if the fertility of the reference plot is poor	0.07	0.26	0.00	1.00	488
Soil color†						
Red	=1 if the color of the reference plot is red	0.26	0.44	0.00	1.00	488
Black	=1 if the color of the reference plot is black	0.54	0.49	0.00	1.00	488
Grey/sand	=1 if the color of the reference plot is grey or sandy	0.20	0.40	0.00	1.00	488
Distance to plot†	Walking distance between the dwelling and the plot (in minutes)	30.98	9.94	0.00	120.00	488
Plot ownership	=1 if the reference plot owned by the HH	0.82	0.38	0.00	1.00	488

Notes: † denotes that the values are self-reported by farmers during the household survey. HH refers to household.

Table 3: Discrepancy between measured (CR) and self-reported (SR) plot size

Plot size group (CR)	Number of obs.	Self-Reported (SR) (1)	Compass-and-Rope (CR) (2)	Bias (SR) – (CR)		Difference in mean (p-value) (5)
				Bias=(1)-(2) (3)	%Bias=(3)/(2) (4)	
≤0.125 ha	70	0.20	0.08	0.12	150%	0.000
0.125–0.25 ha	132	0.31	0.19	0.12	63%	0.000
0.25–0.375 ha	125	0.38	0.30	0.08	27%	0.000
0.375–0.5 ha	74	0.46	0.44	0.02	5%	0.350
0.5–0.75 ha	46	0.60	0.58	0.02	3%	0.783
0.75–1 ha	12	0.64	0.85	–0.21	–25%	0.005
>1.0 ha	24	1.22	1.70	–0.48	–28%	0.019
Total	483	0.42	0.37	0.05	14%	0.002

Note: CR refers compass-and-rope, while SR stands for self-reported farm size.

Table 4: Discrepancy between Crop Cut (CC) and self-reported (SR) production

Plot size group (CR)	Number of obs.	Self-reported (SR) (1)	Crop-cut (CC) (2)	Bias (SR) – (CC)		Difference in mean (p-value) (5)
				Bias=(1)-(2) (3)	%Bias=(3)/(2) (4)	
≤0.125 ha	59	9.1	2.6	6.5	250%	0.000
0.125–0.25 ha	108	13.9	5.6	8.3	148%	0.000
0.25–0.375 ha	87	16.3	7.7	8.6	111%	0.000
0.375–0.5 ha	50	19.1	11.7	7.4	63%	0.000
0.5–0.75 ha	33	26.1	13.6	12.5	91%	0.000
0.75–1 ha	9	24.2	21.8	2.3	10%	0.800
>1.0 ha	19	46.5	32.2	14.3	44%	0.064
Total	365	17.5	8.9	8.5	95%	0.000

Note: CC refers crop-cut and SR stands for self-report, while CR stands for compass-and-rope measurement of farm-size.

Table 5: Characterizing measurement errors in production

Explanatory variables	Dependent variable: ln (self-reported production/crop-cut production)		
	(1)	(2)	(3)
ln (crop-cut production)	-0.656***	-0.658***	-0.670***
	(0.055)	(0.043)	(0.055)
Age of HH head		0.028	0.016
		(0.021)	(0.024)
Age square		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		0.072	0.032
		(0.092)	(0.098)
Size of HH		-0.016	-0.007
		(0.017)	(0.014)
Education of HH head		0.013	-0.001
		(0.076)	(0.104)
Total landholding size		0.038*	0.041
		(0.020)	(0.024)
Soil fertility			
Medium			-0.150**
			(0.072)
Poor			-0.289***
			(0.082)
Soil color			
Black			-0.035
			(0.102)
Grey or sandy			0.234
			(0.146)
Distance from home			0.003
			(0.004)
Distance to the edge			0.000
			(0.003)
Number of corners			-0.002
			(0.010)
Own plot (1=yes)			-0.101
			(0.081)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	0.987***	0.324	0.739
	(0.093)	(0.502)	(0.552)
Observations	365	365	360
R-squared	0.609	0.617	0.635

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table 6: Characterizing measurement errors in production

Explanatory variables	Dependent variable: ln (self-reported production/crop-cut production)		
	(1)	(2)	(3)
ln (compass-and-rope plot size)	-0.596***	-0.590***	-0.558***
	(0.073)	(0.058)	(0.080)
Age of HH head		0.020	0.009
		(0.024)	(0.022)
Age square		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		0.004	-0.020
		(0.105)	(0.101)
Size of HH		-0.018	-0.013
		(0.019)	(0.013)
Education of HH head		0.018	0.022
		(0.087)	(0.115)
Total landholding size		0.024	0.034
		(0.022)	(0.027)
Soil fertility			
Medium			-0.059
			(0.089)
Poor			-0.051
			(0.103)
Soil color			
Black			0.121
			(0.128)
Grey or sandy			0.260*
			(0.151)
Distance from home			0.003
			(0.004)
Distance to the edge			0.000
			(0.003)
Number of corners			-0.014
			(0.010)
Own plot (1=yes)			-0.130
			(0.090)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	-1.327***	-1.793***	-1.379**
	(0.147)	(0.574)	(0.608)
Observations	365	365	360
R-squared	0.495	0.501	0.516

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. High and red are reference categories for soil fertility and color, respectively.

Table 7: Characterizing measurement errors in land area

Explanatory variables	Dependent variable: ln (self-reported area/compass-and-rope plot size)		
	(1)	(2)	(3)
ln (compass-and-rope plot size)	-0.550***	-0.540***	-0.532***
	(0.045)	(0.044)	(0.042)
Age of HH head		0.014	0.010
		(0.015)	(0.015)
Age squared		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		0.005	-0.047
		(0.088)	(0.090)
Size of HH		-0.020*	-0.013
		(0.011)	(0.011)
Education of HH head		0.017	0.007
		(0.066)	(0.068)
Total landholding size		0.061***	0.059***
		(0.016)	(0.017)
Soil fertility			
Medium			-0.098*
			(0.055)
Poor			-0.262***
			(0.084)
Soil color			
Black			-0.137
			(0.087)
Grey or sandy			-0.003
			(0.082)
Distance from home			0.004
			(0.003)
Distance to the edge			0.001
			(0.002)
Number of corners			-0.002
			(0.008)
Own plot (1=yes)			-0.007
			(0.075)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	-0.889***	-1.162***	-0.981**
	(0.090)	(0.341)	(0.415)
Observations	365	365	360
R-squared	0.463	0.494	0.518

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table 8: Correlation between both types of measurement errors

Explanatory variables	Dependent variable: ln (self-reported production/crop-cut production)		
	(1)	(2)	(3)
ln (land area bias)	0.623***	0.542***	0.492***
	(0.091)	(0.105)	(0.114)
Ln (CR Plot size)		-0.310**	-0.246*
		(0.125)	(0.126)
Age of HH head			0.010
			(0.021)
Age square			-0.000
			(0.000)
Gender of HH head			-0.026
			(0.094)
Size of HH			-0.013
			(0.015)
Education of HH head			0.030
			(0.114)
Total landholding size			0.007
			(0.029)
Soil fertility			
Medium			-0.022
			(0.081)
Poor			0.068
			(0.127)
Soil color			
Black			0.181
			(0.122)
Grey or sandy			0.275*
			(0.152)
Distance from home			0.001
			(0.004)
Distance to the edge			-0.004
			(0.003)
Number of corners			-0.029**
			(0.012)
Own plot (1=yes)			-0.125
			(0.097)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	-0.263***	-0.194***	-0.076
	(0.020)	(0.039)	(0.499)
Observations	365	365	360
R-squared	0.481	0.494	0.521

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table 9: Benchmark results: plot size-productivity relationship *(correcting for both area and production measurement errors)*

Explanatory variables	Dependent variable: ln (**crop-cut production/compass-and-rope plot size)**		
	(1)	(2)	(3)
ln (compass-and-rope plot size)	-0.083**	-0.086*	-0.104
	(0.040)	(0.042)	(0.063)
Age of HH head		0.012	0.012
		(0.012)	(0.012)
Age square		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		0.098	0.069
		(0.082)	(0.085)
Size of HH		0.002	0.007
		(0.009)	(0.010)
Education of HH head		-0.005	-0.028
		(0.068)	(0.073)
Total landholding size		0.021	0.011
		(0.024)	(0.020)
Soil fertility			
Medium			-0.137***
			(0.049)
Poor			-0.363***
			(0.118)
Soil color			
Black			-0.237***
			(0.069)
Grey or sandy			-0.043
			(0.074)
Distance from home			-0.001
			(0.002)
Distance to the edge			-0.001
			(0.002)
Number of corners			0.012
			(0.012)
Own plot (1=yes)			0.045
			(0.065)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	3.542***	3.247***	3.351***
	(0.082)	(0.263)	(0.426)
Observations	365	365	360
R-squared	0.518	0.525	0.562

Notes: Standard errors are clustered at the *kebele* level and given in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. High and red are reference categories for soil fertility and color, respectively.

Table 10: Plot size-productivity relationship *(correcting for area measurement only)*

Explanatory variables	Dependent variable: ln (self-reported production/**compass-and-rope plot size**)		
	(1)	(2)	(3)
ln (compass-and-rope plot size)	-0.679***	-0.675***	-0.662***
	(0.079)	(0.083)	(0.074)
Age of HH head		0.033	0.021
		(0.025)	(0.025)
Age square		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		0.102	0.049
		(0.113)	(0.109)
Size of HH		-0.015	-0.006
		(0.014)	(0.014)
Education of HH head		0.013	-0.006
		(0.103)	(0.105)
Total landholding size		0.045*	0.045
		(0.027)	(0.027)
Soil fertility			
Medium			-0.196***
			(0.067)
Poor			-0.414***
			(0.089)
Soil color			
Black			-0.116
			(0.092)
Grey or sandy			0.217
			(0.151)
Distance from home			0.002
			(0.004)
Distance to the edge			-0.001
			(0.003)
Number of corners			-0.002
			(0.010)
Own plot (1=yes)			-0.085
			(0.085)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	2.215***	1.454**	1.972***
	(0.160)	(0.572)	(0.608)
Observations	365	365	360
R-squared	0.576	0.587	0.607

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. High and red are reference categories for soil fertility and color, respectively.

Table 11: Plot size-productivity relationship *(correcting for production measurement only)*

Explanatory variables	Dependent variable: ln (**crop-cut production**/self-reported plot size)		
	(1)	(2)	(3)
ln (self-reported plot size)	-0.410***	-0.404***	-0.578***
	(0.067)	(0.070)	(0.077)
Age of HH head		-0.019	-0.013
		(0.017)	(0.013)
Age square		0.000	0.000
		(0.000)	(0.000)
Gender of HH head		0.252**	0.187*
		(0.115)	(0.109)
Size of HH		0.029*	0.037**
		(0.017)	(0.018)
Education of HH head		-0.072	-0.095
		(0.091)	(0.078)
Total landholding size		-0.029	-0.027
		(0.031)	(0.026)
Soil fertility			
Medium			-0.091
			(0.085)
Poor			-0.162
			(0.137)
Soil color			
Black			-0.142
			(0.092)
Grey or sandy			0.009
			(0.126)
Distance from home			-0.002
			(0.004)
Distance to the edge			0.011***
			(0.003)
Number of corners			0.072***
			(0.012)
Own plot (1=yes)			0.028
			(0.094)
Village level dummies	Yes	Yes	Yes
Constant	2.752***	2.921***	1.682***
	(0.121)	(0.389)	(0.443)
Observations	365	365	360
R-squared	0.403	0.424	0.535

Notes: Standard errors are clustered at *kebele* level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table 12: Plot size—productivity relationship ***(with no correction of measurement errors)***

Explanatory variables	Dependent variable: ln (self-reported production/self-reported plot size)		
	(1)	(2)	(3)
ln (self-reported plot size)	-0.154**	-0.155**	-0.204***
	(0.062)	(0.061)	(0.073)
Age of HH head		0.027	0.022
		(0.022)	(0.022)
Age square		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		0.108	0.067
		(0.086)	(0.084)
Size of HH		-0.009	-0.002
		(0.017)	(0.018)
Education of HH head		-0.024	-0.033
		(0.082)	(0.084)
Total landholding size		0.002	0.005
		(0.030)	(0.029)
Soil fertility			
Medium			-0.063
			(0.063)
Poor			-0.184
			(0.111)
Soil color			
Black			-0.062
			(0.089)
Grey or sandy			0.127
			(0.121)
Distance from home			0.001
			(0.003)
Distance to the edge			0.002
			(0.002)
Number of corners			0.014
			(0.010)
Own plot (1=yes)			-0.065
			(0.071)
Village level dummies	Yes	Yes	Yes
Constant	7.693***	7.128***	7.060***
	(0.111)	(0.476)	(0.566)
Observations	365	365	360
R-squared	0.459	0.465	0.476

Notes: Standard errors are clustered at *kebele* level and given in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table 13: Summary of Empirical Relationships

Source of non-classical measurement error	Key empirically estimated parameters				Estimated SPR	Relative implication on the SPR
	δ	λ	α	π		
No error	NA	NA	NA	NA	-0.104 (0.063)	Insignificant ISPR estimated
Error in production	-0.670*** (0.055)	-0.558*** (0.080)	NA	NA	-0.659*** (0.074)	Strongest ISPR estimated
Error in plot size	NA	NA	-0.532*** (0.042)	NA	-0.578*** (0.077)	Strong ISPR estimated
Error in both	-0.670*** (0.055)	-0.558*** (0.080)	-0.532*** (0.042)	0.492*** (0.114)	-0.204*** (0.073)	Weaker ISPR estimated

Notes: we extracted the above estimates and standard errors (given in parenthesis) from our conditional regressions. NA refers that these parameters are either not relevant or not empirically estimated. SPR stands for the size-productivity relationship while ISPR represents the invers size-productivity relationship.

APPENDIX

Table A1: Characterizing non-responses in crop-cut production

Explanatory variables	Dependent variable: Crop-cuts (1=yes)
	(1)
Age of HH head	0.001
	(0.008)
Age square	0.000
	(0.000)
Gender of HH head	-0.065
	(0.039)
Size of HH	-0.002
	(0.008)
Education of HH head	-0.023
	(0.039)
Total landholding size	0.006
	(0.006)
Soil fertility	
Medium	0.036
	(0.022)
Poor	-0.006
	(0.046)
Soil color	
Black	-0.003
	(0.039)
Grey or sandy	-0.020
	(0.043)
Distance from home	0.000
	(0.001)
Own plot (1=yes)	0.018
	(0.035)
Constant	0.822^{***}
	(0.167)
Observations	488
R-squared	0.692

Note: Standard errors are clustered at *kebele* level and given in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table A2: Characterizing measurement errors in production

Explanatory variables	Dependent variable: Inverse Hyperbolic Sine (self-reported production/crop-cut production)		
	(1)	(2)	(3)
IHS (crop-cut production)	-0.576***	-0.578***	-0.595***
	(0.057)	(0.038)	(0.057)
Age of HH head		0.031*	0.021
		(0.018)	(0.021)
Age square		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		0.084	0.051
		(0.078)	(0.088)
Size of HH		-0.017	-0.009
		(0.014)	(0.010)
Education of HH head		0.017	-0.001
		(0.064)	(0.088)
Total landholding size		0.034**	0.037
		(0.017)	(0.023)
Soil fertility			
Medium			-0.152**
			(0.063)
Poor			-0.236***
			(0.071)
Soil color			
Black			-0.041
			(0.083)
Grey or sandy			0.192
			(0.121)
Distance from home			0.003
			(0.003)
Distance to the edge			0.000
			(0.002)
Number of corners			0.000
			(0.007)
Own plot (1=yes)			-0.090
			(0.063)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	2.238***	1.496***	1.851***
	(0.138)	(0.431)	(0.506)
Observations	365	365	360
R-squared	0.606	0.617	0.637

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. IHS refers to Inverse Hyperbolic Sine transformation High and red are reference categories for soil fertility and color, respectively.

Table A3: Characterizing measurement errors in land area

Explanatory variables	Dependent variable: Inverse Hyperbolic Sine (ratio of self-reported and compass-and-rope plot size)		
	(1)	(2)	(3)
IHS (compass-and-rope plot size)	-0.766***	-0.730***	-0.543***
	(0.132)	(0.136)	(0.121)
Age of HH head		0.020	0.016
		(0.014)	(0.013)
Age square		-0.000	-0.000
		(0.000)	(0.000)
Gender of HH head		-0.041	-0.069
		(0.086)	(0.082)
Size of HH		-0.019**	-0.018*
		(0.009)	(0.010)
Education of HH head		0.018	0.012
		(0.063)	(0.063)
Total landholding size		0.048***	0.051***
		(0.016)	(0.017)
Soil fertility			
Medium			-0.096*
			(0.052)
Poor			-0.181**
			(0.067)
Soil color			
Black			-0.113
			(0.077)
Grey or sandy			0.026
			(0.083)
Distance from home			0.003
			(0.004)
Distance to the edge			-0.003*
			(0.002)
Number of corners			-0.022**
			(0.008)
Own plot (1=yes)			-0.007
			(0.066)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	1.190***	0.816**	1.255***
	(0.021)	(0.321)	(0.346)
Observations	365	365	360
R-squared	0.295	0.330	0.382

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. IHS refers to Inverse Hyperbolic Sine transformation. High and red are reference categories for soil fertility and color, respectively.

Table A4: Correlation between both types of measurement errors

Explanatory variables	Dependent variable: Inverse Hyperbolic Sine (ratio of self-reported and crop-cut production)		
	(1)	(2)	(3)
IHS (area bias)	0.694***	0.621***	0.567***
	(0.101)	(0.106)	(0.114)
compass-and-rope Plot size		-0.269***	-0.225**
		(0.095)	(0.097)
Age of HH head			0.015
			(0.018)
Age square			-0.000
			(0.000)
Gender of HH head			0.001
			(0.086)
Size of HH			-0.014
			(0.012)
Education of HH head			0.027
			(0.097)
Total landholding size			0.006
			(0.026)
Soil fertility			
Medium			-0.035
			(0.067)
Poor			0.050
			(0.098)
Soil color			
Black			0.143
			(0.097)
Grey or sandy			0.217*
			(0.121)
Distance from home			0.001
			(0.003)
Distance to the edge			-0.003
			(0.002)
Number of corners			-0.022**
			(0.010)
Own plot (1=yes)			-0.109
			(0.082)
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	0.117	0.238*	0.201
	(0.108)	(0.120)	(0.445)
Observations	365	365	360
R-squared	0.494	0.509	0.534

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. IHS refers to Inverse Hyperbolic Sine transformation. High and red are reference categories for soil fertility and color, respectively.

Table A5: Characterizing measurement errors in land area using the full sample data

Explanatory variables	Dependent variable: ln (self-reported area/compass-and-rope plot size)		
	(1)	(2)	(3)
ln (compass-and-rope plot size)	-0.553***	-0.544***	-0.536***
	(0.040)	(0.039)	(0.042)
Household characteristics	No	Yes	Yes
Plot characteristics	No	No	Yes
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	-0.844***	-1.018***	-0.994**
	(0.082)	(0.300)	(0.392)
Observations	483	483	373
R-squared	0.461	0.483	0.524

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table A6: Plot size-productivity relationship *(correcting for area measurement only)* using the full sample data

Explanatory variables	Dependent variable: ln (self-reported production/**compass-and-rope plot size**)		
	(1)	(2)	(3)
ln (compass-and-rope plot size)	-0.671***	-0.664***	-0.659***
	(0.072)	(0.073)	(0.074)
Household characteristics	No	Yes	Yes
Plot characteristics	No	No	Yes
Village (*kebele*) dummies	Yes	Yes	Yes
Constant	2.229***	1.766***	2.049***
	(0.148)	(0.489)	(0.581)
Observations	483	483	373
R-squared	0.603	0.608	0.617

Notes: Standard errors are clustered at *kebele* level and given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.

Table A7: Plot size—productivity relationship *(with no correction of measurement errors)* using the full sample data

Explanatory variables	Dependent variable: ln (self-reported production/self-reported plot size)		
	(1)	(2)	(3)
ln (self-reported plot size)	-0.174***	-0.172***	-0.223***
	(0.062)	(0.060)	(0.073)
Household characteristics	No	Yes	Yes
Plot characteristics	No	No	Yes
Village (*kebele*) dummies	Yes	Yes	Yes
Observations	488	488	374
R-squared	0.466	0.469	0.480

Notes: Standard errors are clustered at *kebele* level and given in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. High and red are reference categories for soil fertility and color, respectively.