ANNUAL REVIEWS

Annual Review of Resource Economics

Inferential and Behavioral Implications of Measurement Error in Agricultural Data

Kibrom A. Abay,[1] Tesfamicheal Wossen,[2] Gashaw T. Abate,[3] James R. Stevenson,[4] Hope Michelson,[5] and Christopher B. Barrett[6]

[1]International Food Policy Research Institute (IFPRI), Cairo, Egypt; email: K.Abay@cgiar.org

[2]International Institute of Tropical Agriculture (IITA), Nairobi, Kenya

[3]International Food Policy Research Institute (IFPRI), Washington, DC, USA

[4]CGIAR Standing Panel on Impact Assessment, Rome, Italy

[5]Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[6]Charles H. Dyson School of Applied Economics and Management, Cornell University, Ithaca, New York, USA

Annu. Rev. Resour. Econ. 2023. 15:63–83

First published as a Review in Advance on June 7, 2023

The *Annual Review of Resource Economics* is online at resource.annualreviews.org

https://doi.org/10.1146/annurev-resource-101422-090049

JEL codes: C12, C81, C83, D13, O12, Q12

Keywords

bias, misperception, misreporting, statistical inference

Abstract

An evolving literature evaluates the inferential and behavioral implications of measurement error (ME) in agricultural data. We synthesize findings on the nature and sources of ME and potential remedies. We provide practical guidance for choosing among alternative approaches for detecting, obviating, or correcting for alternative sources of ME, as these have different behavioral and inferential implications. Some ME biases statistical inference and thus may require econometric correction. Other types of ME may affect (and shed light on) farmers' decision-making processes even if farmers' responses are objectively incorrect. Where feasible, collecting both self-reported and objectively measured data for the same variable may enrich understanding of policy-relevant agricultural and behavioral phenomena.

1. INTRODUCTION

Agricultural data collected using a range of methods are important to the design and evaluation of policies in low- and middle-income countries (LMICs), where agriculture remains a major employer and source of income. Unfortunately, these data are often characterized by measurement error (ME) arising from a range of sources. ME can distort descriptive, inferential, predictive, and prescriptive analysis, with significant consequences. For example, the often-observed inverse relationship between farm size and productivity has long motivated land reform and other smallholder-oriented policies, but several recent studies suggest that such findings might arise purely due to ME (e.g., Desiere & Jolliffe 2018, Abay et al. 2019, Gourlay et al. 2019). Similarly, Gollin & Udry (2021) show that ME accounts for a substantial share of the observed productivity dispersion in African agriculture, calling into question the wisdom of policies intended to reduce seemingly high allocative inefficiency.

The last decade has seen important innovations to improve the measurement, collection, and maintenance of agricultural data, as recently reviewed by Carletto et al. (2021) and De Weerdt et al. (2020). Many such advances rely on new technologies, ranging from computer-assisted personal interviewing (CAPI) and global positioning system (GPS) devices to the more specialized genotyping sequencers and soil spectrometers. These technologies not only advance data collection accuracy with known levels of precision, but they also provide an objective benchmark against which self-reported survey data can be fruitfully compared.

Data collection innovations and objective benchmarks have sparked renewed research on, and a deeper understanding of, ME in agricultural data. Recent studies find widespread ME in key agricultural metrics, including land size (e.g., Carletto et al. 2013, 2015; Abay et al. 2019, 2021; Dillon et al. 2019; Bevis & Barrett 2020; Burke et al. 2020); crop variety type (e.g., Kosmowski et al. 2019; Wossen et al. 2019a,b; Wineman et al. 2020; Maredia & Bartle 2023); soil quality (e.g., Berazneva et al. 2018, Gourlay et al. 2019); labor inputs (e.g., Dillon et al. 2012, Arthi et al. 2018, Gaddis et al. 2021); fertilizer, seed, pesticide, and herbicide quality (Bold et al. 2017, Ashour et al. 2019, Michelson et al. 2021), and output (Desiere & Jolliffe 2018, Abay et al. 2019, Gourlay et al. 2019, Lobell et al. 2019, Abay 2020; Kosmowski et al. 2021). Several of these studies advance methods to improve data quality and statistical inferences based on ME-ridden survey data.

Yet objective measures may themselves suffer from ME. For example, measuring fertilizer and seed quality in the laboratory can introduce new errors (Michelson et al. 2021); the multidimensional criteria needed to define quality for these inputs can also make their evaluation contingent and uncertain. Similarly, a sensor degradation can systematically bias satellite-based measures (Jain 2020). Moreover, not all mismeasurements are amenable for correction, and incomplete correction of errors can prove counterproductive, even increasing bias relative to not correcting (Abay et al. 2019). And despite the significant progress in improved data collection methods (e.g., laboratory-based soil testing through spectral analysis, DNA fingerprinting of crop germplasm), such innovations add (sometimes quite substantial) costs to surveys that necessarily limit actual deployment of objective measures at scale (Carletto et al. 2021).

In fact, ME can itself provide important information. While much of the econometrics literature treats ME as misreporting by survey respondents or technical mismeasurement as merely a statistical challenge to be overcome, several recent studies underscore that some ME may reveal survey respondents' mistaken beliefs—often termed misperceptions—with real consequence for statistical inference as well as respondents' actual decision-making processes and understanding these decisions. For example, Abay et al. (2021) show that a large share of ME in plot sizes reported by farmers in four African countries (Ethiopia, Malawi, Tanzania, and Uganda) appears to reflect farmers reporting mistaken beliefs rather than misreporting of accurate beliefs. Similarly, Wossen et al. (2022) demonstrate that much of the misclassification in crop variety identification arises from farmer misperceptions. Other studies provide similar evidence of misperception in soil quality (Berazneva et al. 2018, Abay et al. 2022a), chemical fertilizer quality (Michelson et al. 2021), and crop variety (Wineman et al. 2020, Maredia & Bartle 2023). ME in agricultural data might therefore matter to farmers' decision-making processes. Indeed, some studies argue that mistaken beliefs regarding the quality, type, and nature of agricultural inputs limits farmers' purchase and use (Bold et al. 2017, Ashour et al. 2019, Michelson et al. 2021, Abay et al. 2022b). Knowing when, whether, and how to correct for ME is thus an increasingly important skill for applied economists. This review synthesizes recent findings on ME in agricultural data and offers practical guidance to improve the production and use of agricultural and household data. We present a systematic approach to identify alternative sources of ME and to assess whether or not to correct for rather than directly study the ME. Specifically, we provide important insights regarding (*a*) what sources generate ME in agricultural data, (*b*) how to identify alternative sources of ME, (*c*) what the implications of ME are for data collection and econometric inference, (*d*) when and how to use (accurate) objective versus (erroneous) self-reported measures, and (*e*) whether farmer misperceptions can be easily corrected. Two recent reviews provide comprehensive discussions of improvements in survey designs and agricultural data measurement (De Weerdt et al. 2020, Carletto et al. 2021), so we focus on the inferential and behavioral implications of ME as well as potential remedies to improve statistical inferences and address behavioral anomalies driving potential mismeasurement in survey data. Although these issues apply globally, we focus on lessons learned from studies of African agriculture, where most of these studies took place. Our synthesis is relevant to practitioners and researchers working to improve the measurement and use of agricultural data in LMICs. We summarize our key insights in a decision tree for applied researchers to navigate these issues (**Figure 1**).

2. TYPOLOGY AND PROPERTIES OF MEASUREMENT ERROR IN AGRICULTURAL DATA

We begin with a review of ME typologies and properties. The inferential and behavioral implications of ME in agricultural data depend on (*a*) whether ME is classical or nonclassical (i.e., randomly or systematically distributed), (*b*) whether the (mis)measured variable of interest is continuous or discrete, (*c*) the role of the (mis)measured variable in inference and behavioral analysis, in particular whether it is an outcome or explanatory variable, (*d*) ME origins (misperception, misreporting, or both), and (*e*) correlation structure among ME across potential variables of interest (correlated or uncorrelated mismeasurements).

We start with the following measurement system in the context of a linear model:

$$Y^* = \alpha + \beta X^* + \epsilon \qquad 1.$$

$$X = X^* + \mu \qquad 2.$$

$$Y = Y^* + \omega, \qquad 3.$$

where X^* is the true but unobserved explanatory variable of interest (e.g., true land size or variety type), Y^* is the true unobserved outcome of interest (e.g., production or yield), and ϵ is a mean-zero error term. In the above measurement system, we only observe X and Y, noisy measures of X^* and Y^*, respectively, which are contaminated by MEs (μ and ω). We assume that X^* is strictly exogenous [i.e., $E(\epsilon X^*) = 0$]. The specification in Equation 1 assumes that ME in survey data enters additively, although the model can accommodate multiplicative ME through the use of a log linear specification.

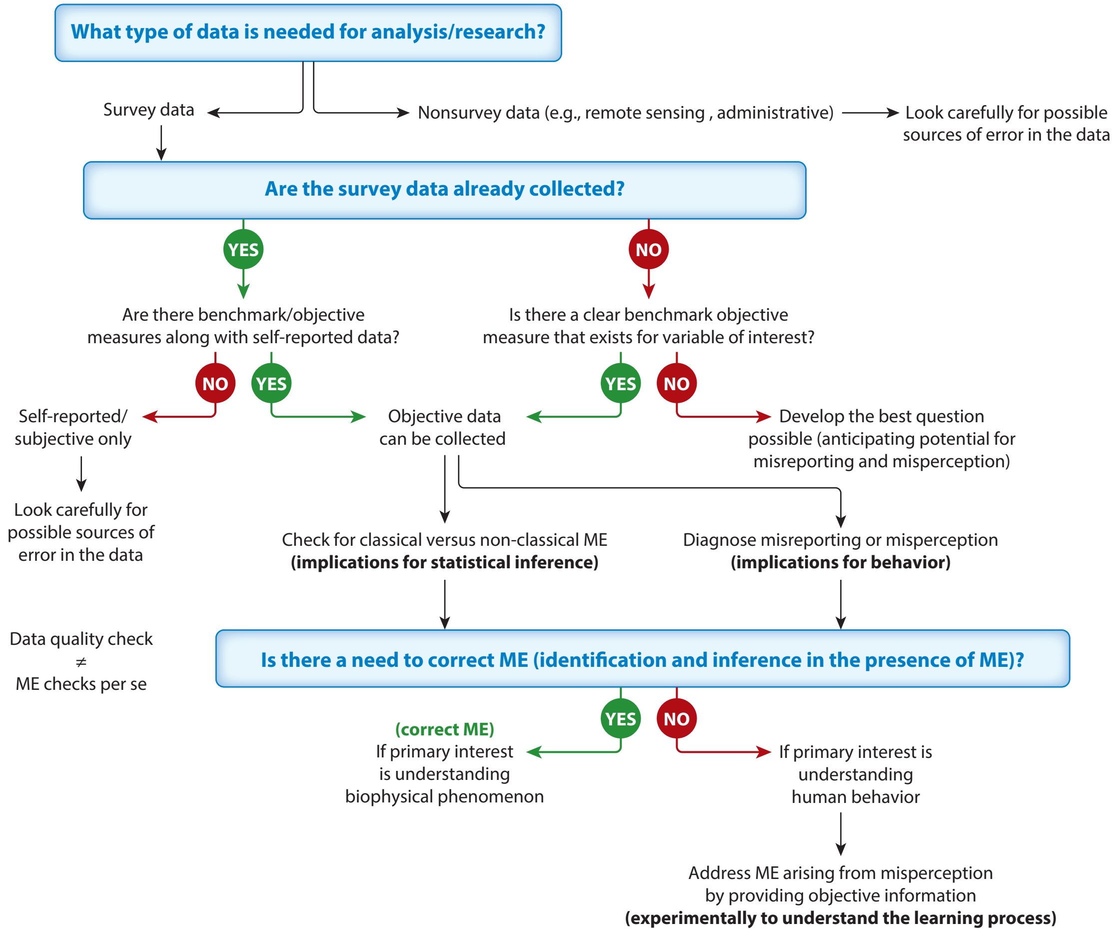

Figure 1

Decision tree: identifying, testing, and correcting measurement error (ME).

2.1. Classical Versus Nonclassical Measurement Error

In Equations 1–3, classical measurement error (CME) arises when μ and ω are uncorrelated with both X^* and Y^* as well as between themselves. For example, in land size or production surveys, CME arises when farmers misreport their true land size or harvest in a nonsystematic way. CME in dependent variables is innocuous and only attenuates estimated standard errors; but CME in explanatory variables biases the estimated slope coefficient (β) toward zero, known as attenuation bias. Nonclassical measurement error (NCME) arises when μ or ω is correlated with X^* or Y^*. For example, in land size and production surveys, NCME in self-reported land size and harvest arise when farmers with larger (smaller) plots (harvests) systematically under (over) report their land area (harvest). NCME patterns appear commonplace in self-reported plot size and harvest data (Carletto et al. 2013; Desiere & Jolliffe 2018; Abay et al. 2019, 2021; Gourlay et al. 2019; Bevis & Barrett 2020). When the dependent or independent variables in Equations 1–3 are contaminated by NCME, the estimated slope coefficient can be biased upward or downward.

2.2. Measurement Error in Continuous Versus Discrete Variables

When X^* is continuous (i.e., has no support restrictions), ME can be classical or nonclassical. When X^* is discrete, however, ME (commonly referred to as misclassification) is inherently nonclassical, as the latent true value and its measurement share the same discrete support (Mahajan 2006, Lewbel 2007, Hu 2008, Meyer & Mittag 2017). For example, in the context of agricultural technology adoption surveys, any error in binary adoption status indicators must always be nonclassical (Wossen et al. 2019a, 2022).[1]

2.3. Correlated Versus Uncorrelated Measurement Errors

Multiple agricultural metrics may suffer from various forms of ME, and these errors can be correlated or uncorrelated among each other. Correlated MEs occur when the ME in the dependent variable in Equation 3 (ω) is correlated with the ME in the explanatory variable from Equation 2 (μ). Such correlation may arise due to common behavioral phenomenon, such as focal point rounding or heaping (discussed below), that independently contaminate multiple series, creating spurious correlation in ME. The ME in other variables may have structural origins, as in the case of Berkson-type ME (discussed below). Multiple examples exist of agricultural metrics that exhibit correlated ME. For example, MEs in land size and production appear strongly correlated (Abay et al. 2019). Correlated ME can complicate identification of the inferential biases associated with mismeasurement in agricultural data and impede corrections based on incompletely improved data sources. Abay et al. (2019) show that when MEs in the dependent and independent variables are positively correlated, correcting only one erroneous variable with an accurate, objective measure may aggravate the bias in estimated regression coefficients relative to estimation using two erroneous variables. Given how pervasive ME is, one should assume correlated MEs unless one can credibly explain why MEs among key variables in the model would be uncorrelated.

2.4. Misreporting Versus Misperception

Traditionally, econometricians treat ME as a reporting error by survey respondents who know and act upon the true value of the variable(s) of interest, and simply misreport it, whether intentionally or inadvertently. In such cases, it becomes imperative to correct, as best as one can, for ME. But what is commonly interpreted as misreporting—that is, the differences between self-reported and objectively measured values—may instead reflect, at least in part, mistaken beliefs or misperceptions honestly held by the survey respondent, especially if the true value of the variables of interest is beyond the respondent's control or capacity for knowledge (Lewbel 2007; Abay et al. 2021, 2022b; Wossen et al. 2022). For example, a farmer may have been allocated a plot of land by a customary authority or a family member without any formal, cadastral measure of the area transferred and thus not know its true size. Frequent subdivision of parcels into plots may also lead to misperception of plots sizes. In some contexts, traditional measurement units may cause misperceptions. For example, land area in parts of Ethiopia is traditionally measured by oxen days needed to plow or till it. But this measure can create mistaken beliefs about plot size if farmers fail to correct for factors that vary across plots, including travel distance to the plot, soil texture, plot slope, and livestock power/health. Or a farmer may have been told a seed type by an input dealer or extension agent but be unable to verify the variety herself by direct inspection. The informality of

[1]For example, considering adoption of agricultural technologies, when the true adoption status is one (i.e., adopter status), the ME can take a value of only negative one (false negative). Similarly, when the true adoption status is zero (i.e., nonadopter status), ME can take a value of only one (false positive).

seed markets in most developing countries can be a major source of misperception about the planting material farmers use (Wossen et al. 2022). In such cases, self-reported information might not reflect misreporting per se so much as it reflects agents' underlying beliefs and decision-making processes (Drerup et al. 2017, Abay et al. 2021, Wossen et al. 2022). Under this scenario, the analyst needs to decide what truth she wishes to study. Self-reported information may have more explanatory power for understanding agents' decision-making processes and behavioral responses such as input allocation decisions, while objective measures may be more relevant for predicting biophysical relationships such as yield response to an input (Abay et al. 2021).

2.5. Optimal Prediction Error (Berkson-Type Error)

Berkson-type error is a special case of ME based on misperception that may arise when respondents have imperfect control over a given variable and form their best prediction of the true value given available information, almost surely yielding a prediction that deviates from the true value (Berkson 1950, Hyslop & Imbens 2001, Hoderlein & Winter 2010, Schennach 2020). Farm households may predict rather than know agricultural metrics for several reasons. For example, keeping track of frequent, variable agricultural input applications (e.g., labor, fertilizers, pesticides) can be difficult; farmers may predict these outcomes using other characteristics (e.g., land size multiplied by a recommended application rate or the weight of a purchased container but ignoring amounts spilled or given to others). Rounding can also generate Berkson-type error (Schennach 2020). As shown by Hyslop & Imbens (2001), the inferential implications of these types of errors depend on the information set agents use for predicting a variable and whether a predicted variable appears as a dependent or an explanatory variable in predictive models. Unlike the usual CME, Berkson-type error is innocuous when the predicted variable appears as an explanatory variable and agents employ only the information set associated with the dependent variable (Hyslop & Imbens 2001). On the other hand, Hyslop & Imbens (2001) show that Berkson-type ME in the dependent variable leads to downward bias.

2.6. Differential and Nondifferential Measurement Error

Nondifferential ME occurs when the error is independent of the outcome (dependent variable) of interest, while differential error involves ME that is correlated with the outcome of interest (Imai & Yamamoto 2010). However, nondifferential ME is not necessarily equivalent to CME: Unlike CME, nondifferential error may still depend on the true value of the mismeasured variable. For instance, in Equations 1–2 differential ME may arise due to (*a*) the correlation between the ME and other observables in the model, including the dependent variable, and (*b*) the correlation between the ME and the unobservable [i.e., $\mathrm{Cov}(\mu, \epsilon) \neq 0$]. Conceptually, the distinction between differential and nondifferential ME resembles that between misreporting and misperceptions.

3. SOURCES OF MEASUREMENT ERROR IN AGRICULTURAL DATA

This section reviews sources and potential causes of ME in agricultural data (see Carletto et al. 2021 for a more comprehensive discussion of sources of ME in agricultural data). MEs may arise from errors generated by the choices researchers make regarding the data-generation process and analysis (i.e., survey design, data collection, and processing), from participant (e.g., farmers) misreporting (e.g., erroneous response), and misperceptions (i.e., mistaken beliefs). ME can also arise from third parties (e.g., enumerators, lab tests, and technologies) involved in the production and processing of agricultural data.

3.1. Measurement Error Arising from Researchers' Design and Choice

The decisions researchers make about the data-generation process (e.g., questionnaire design, field implementation modes, and protocols) often involve trade-offs related to budget constraints and are among the main sources of ME (Biemer 2010, Dillon et al. 2020). Poorly designed survey instruments are a major source of ME. Questionnaire length and complexity, sequencing and skip patterns, durations and differences in reference periods, (incomplete) response categories, and (im)precise definitions, among others, can have important implications (Schwarz et al. 1991, Beaman & Dillon 2012, Kilic & Sohnesen 2019, De Weerdt et al. 2020, Ambler et al. 2021b). For example, Wollburg et al. (2021) find that key agricultural input, output, and productivity variables are systematically related to the length of the recall period, with longer recall periods systematically generating higher reported quantities of harvest, labor, and fertilizer inputs. A recent survey experiment that randomly assigned seven different household consumption and expenditure survey designs in Tanzania finds that hunger rates can vary by up to 50 percentage points and that differences in hunger rates vary systematically by household characteristics depending on the design (De Weerdt et al. 2015, 2020). Similarly, a survey experiment that varied the level of detail in a labor survey module (e.g., with and without screening questions) shows that questionnaire design impacts employment statistics, with effects varying by gender (Bardasi et al. 2012).[2] Survey specification and unit errors may also introduce substantial ME. Survey specification errors occur when the concept implied by the survey question differs from the concept that should have been measured (e.g., lack of clarity in defining plots relative to parcels), and unit errors occur due to missing conversion factors (i.e., failure to collect adequate detail on the conversion of local units of measure to more universal ones) or from forcing respondents to report in standardized units that they do not typically use (Biemer 2010, De Weerdt et al. 2015, Carletto et al. 2021). These types of MEs can lead to incorrect constructs, biased parameter estimates, and invalid inferences.

Survey modes (e.g., face-to-face versus by phone/web, on-paper versus electronic format, self-administered versus interviewer-led) and associated choices made by researchers can also affect data quality. Recent experimental studies show differences between information collected by interviewing face-to-face versus telephone. Abate et al. (2023) find a 23% difference in reported consumption and show that this lower reported consumption via phone interviewing results in an estimated poverty headcount twice as high in the phone survey relative to the in-person survey. Conversely, evidence of overreporting in phone survey mode is found in dietary diversity measures (Lamanna et al. 2019) and contraceptive use (Greenleaf et al. 2020).

The environment and setting in which a survey is conducted can also introduce ME if the context contributes to intentional misreporting by respondents, especially for sensitive topics subject to economic and social desirability bias (e.g., food insecurity, land size and use, technology adoption, employment, contraceptive use). Kilic et al. (2020) show that nonprivate interviewing results in significant underreporting of employment compared to measurement through private interviews, with a stronger effect for women. Researchers' decisions related to both who within the household to interview and what to use as an enumerator (gender, age group, income level) may also affect data quality through context-specific social norms and customs (e.g., Caroli & Weber-Baghdiguian 2016). For instance, social norms governing interactions between men and women may shape what can be reported to or discussed with an interviewer of the opposite gender. Other fieldwork design choices likely to contribute to measurement error include decisions

[2] Survey length and associated respondent and enumerator fatigue may also introduce substantial ME (e.g., Laajaj & Macours 2021, Ambler et al. 2021b, Abay et al. 2022c).

about survey visit timing (time of day and time of year). Phung et al. (2015) show that time of day and season significantly affect data quality, while LoPalo (2022) find that the temperature at the time of interview affects the quantity and quality of interviewer output.

Besides errors that arise because of poor research design, researchers' decisions in the construction of variables from raw data can also generate substantial ME, including aggregation bias (Sharma & Gibson 2019, Huntington-Klein et al. 2021). For example, researchers' desire to measure imprecisely defined concepts (e.g., ability), choice of imputation (Huntington-Klein et al. 2021), and aggregation process (Sharma & Gibson 2019) can introduce errors that may significantly affect statistical inferences.

3.2. Measurement Error Arising from Survey Respondents

ME also originates from study participants, mainly due to misreporting and misperception. Misreporting occurs when respondents provide incorrect information to a response to a question, intentionally or unintentionally (Biemer 2010). Recall errors (because of memory decays) are perhaps the most common such ME. Questions with longer recall periods or that elicit information about less frequent and less salient events are often subject to recall errors mainly because of the cognitive demand on the respondents (Fermont & Benson 2011, Beegle et al. 2012, Gourlay et al. 2019). De Nicola & Giné (2014) found that beyond a certain recall period, respondents switch from enumeration to estimation, which translates into different error patterns. Similarly, Wollburg et al. (2021) and Arthi et al. (2018) document the presence of nonrandom MEs systematically related to the length of the recall period in multiple agricultural surveys.

Closely related to recall error that may lead to (unintentional) misreporting is telescoping, wherein respondents misdate events by either recalling more distant events as occurring more recently (forward telescoping) or pushing recent events further back in time (backward telescoping) (Sudman & Bradburn 1973, Gaskell et al. 2000). Recent experiments confirm that responses from recalls with no memorable recall marker (unbounded recall) significantly differ from bounded recalls that provide a salient recall marker (reference points) in respondents' memory to better contextualize responses (De Weerdt et al. 2020, Abate et al. 2022). For instance, the typical seven-day food consumption recall without a salient recall marker results in a 16% higher consumption (equivalent to an entire extra day worth of consumption) compared to consumption levels based on bounded recall (Abate et al. 2022).

Measurement error is also a common problem when the variables of interest are unobservable (latent), either for privacy reasons (e.g., crime, sexual preference, gender roles) or because they are difficult to observe or measure directly (e.g., ability, risk aversion, empowerment) (Schennach 2020, Calvi et al. 2022). This is often the case in gender empowerment studies, particularly when both hidden information (adverse selection) and hidden actions (moral hazard) are present (Ambler et al. 2021a, Kilic et al. 2021). Social norms and customs associated with specific socioeconomic groups (e.g., gender, age group and income levels) may influence rates of intentional misreporting (Caroli & Weber-Baghdiguian 2016). For instance, a recent study in Ghana found systematic differences in amounts of agricultural labor reported by men versus women and by their perceived quality of marriage (Dervisevic & Goldstein 2023).

Misreporting may also arise from the use of proxy respondents (e.g., typically the male head of the households in agricultural surveys) who inaccurately represent activities or events related to household members that occur outside the respondent's purview (Biemer 2010). Kilic et al. (2021) show that interviewing one knowledgeable household member results in higher measured rates of exclusive ownership of land assets among men and lower rates of joint ownership among women, compared to those reported by self-respondents. Bardasi et al. (2012) show that response by proxy rather than self-report produces substantially lower male employment rates, mostly due to underreporting of agricultural activities with effects attenuated when the proxy respondents are spouses and individuals with some schooling. Respondents recall facts about their own lives more reliably than others' characteristics (Baird et al. 2008, Hogset & Barrett 2010). Reliance on proxy informants to report on household-level agricultural variables may also lead to misreporting due to aggregation errors (Dillon & Mensah 2021).

Reporting error due to respondent rounding (focal point bunching) occurs often in farmers' self-reported land area and production, with regression-to-mean patterns as farmers owning relatively larger (smaller) landholding under(over)-reporting landholdings (e.g., Carletto et al. 2013, 2015; Desiere & Jolliffe 2018; Abay et al. 2019, 2021; Abay 2020; Gourlay et al. 2019). Carletto et al. (2015) and Abay et al. (2019) confirm that the presence of rounding contributes to increasing ME mainly due to asymmetric heaping (i.e., bunching by farmer with small landholdings); Abay et al. (2022b) show that rounding explains about 18% of the variation in ME in land area in Malawi.

Respondents may also intentionally misreport (Bound et al. 2001, Norwood & Lusk 2011, Krumpal 2013). In the case of landholdings, respondents may deliberately overreport due to status considerations or underreport because of taxation or redistribution concerns. A number of agricultural variables including modern input use, hired labor and child labor, and adoption of sustainability standards and climate-smart practices can similarly be subject to social desirability bias. Jouvin (2021) shows that child labor use in Ivorian cocoa farms is twice as high when measured using a list experiment approach that reduces social desirability relative to a direct survey question related to child labor. Another study on environmental protection behavior found that more than 60% of the self-reports were inaccurate, because of both under- and overreporting (Moore & Rutherfurd 2020).

Misperception is another mechanism through which respondents introduce ME in survey data. Although it is challenging to distinguish misperception from misreporting in survey data, few recent studies use empirical techniques to apportion MEs between misperception and misreporting (e.g., Abay et al. 2021, Wossen et al. 2022). One distinction between misreporting and misperception relates to the impact of ME on respondents' actual decision-making processes. In the case of misreporting, respondents know and base their decision on their latent, accurate perceptions instead of its mismeasured values (Lewbel 2007, Chetty 2012). Misperception, however, reflects inaccurate beliefs that affect respondents' underlying decision-making processes (Chetty 2012, Drerup et al. 2017, Hu & Wansbeek 2017, Wilhelm 2019, Abay et al. 2021). Abay et al. (2021) show that MEs in plot sizes represent a mixture of farmers' misperception and misreporting, whereas Wossen et al. (2022) show that crop variety misclassification is mostly driven by misperception (i.e., farmers acquiring planting material that is not what they think).

Besides misreporting and misperception, respondents may also introduce MEs through optimal prediction error. Some farmers may base their assessment of soil quality and soil type on crop yields (Berazneva et al. 2018) or their plot size based on the seed quantity used (David 1978). Prediction error may also be introduced by the length of the recall period because respondents may switch from enumeration to estimation beyond a certain recall length (De Nicola & Giné 2014).

3.3. Measurement Error Arising from Third-Party Actors

ME in agricultural data may also be introduced by third-party actors such as enumerators or technologies deployed for data collection. Survey enumerators' actions and characteristics may systematically affect respondents' responses. Enumerator-related ME arises when differences across enumerators (e.g., in gender, education, ability) and their behavior (e.g., pace of interview, compliance with survey procedure, fatigue) affect the interview process (Biemer 2010, Berazneva 2014, Carletto et al. 2021). Enumerators' characteristics and their differences from respondent characteristics can generate systemic biases (Baird et al. 2008, Lavrakas 2008, De Weerdt et al.

2015, Phung et al. 2015, Himelein 2016, Di Maio & Fiala 2019). Baird et al. (2008) find systematic errors in survey data due to enumerator fatigue (i.e., error rates increase with the number of interviews completed per day). The nature and magnitude of enumerator effects are also influenced by question salience and sensitivity (Himelein 2016, Laajaj & Macours 2021).

Researchers are increasingly relying on alternative data sources such as sensors as low-cost and scalable solutions to collect agricultural data (Burke & Lobell 2017, Lobell et al. 2019). While these relatively new data sources offer opportunities for collecting agricultural data with rich spatial and temporal dimensions, these modes of data collection are not immune from ME (Japec et al. 2015, Amaya et al. 2020). For instance, several studies show that poorly equipped technologies including the lack of calibration in satellite-based models can generate MEs in satellite-based estimates of yield (Lobell et al. 2007, 2019; Jain 2020), agriculture/irrigation water use (Foster et al. 2020), economic activities (Gibson et al. 2020, 2021), economic well-being (Yeh et al. 2020), and pollution concentrations (Fowlie et al. 2019), among others. Gibson et al. (2020, 2021) show that nightlight data extracted from the Defense Meteorological Satellite Program (DMSP) satellite sensors suffer from NCME because of blurring, top-coding, and lack of calibration. Michler et al. (2022) show that the sign and magnitude of the relationship between agricultural production and weather varies across remote sensing weather products, while Li & Ortiz-Bobea (2022) demonstrate how researchers' decisions related to the aggregation of remotely sensed weather data can generate NCME. Amaya et al. (2020), Jain (2020), Fowlie et al. (2019), and Gibson et al. (2021) provide extensive discussions on various sources of ME in satellite data and potential avenues to address them.

4. TESTING AND CORRECTING FOR MEASUREMENT ERROR IN HOUSEHOLD SURVEY DATA

The appropriate responses to ME depend on the nature and type of the errors as well as the analytical objective. **Figure 1** summarizes the pertinent questions and appropriate responses in the form of a decision tree. The first question relates to the nature of data needed: survey or nonsurvey (e.g., remote sensing, administrative) data, or both. The choice of data should consider potential vulnerability to mismeasurement as well as potential remedies to address these inaccuracies. As discussed in Section 3, the sources of ME in survey data and nonsurvey data differ, requiring different diagnostics.

In the case of survey data still to be collected, some ME can be addressed by improving research design and measurement protocols. As described by De Weerdt et al. (2020) and Carletto et al. (2021), definitional clarity, length, complexity, and modality of delivering survey questionnaires all affect the quality and relevance of agricultural data to inform policy. Although some of these improvements can be costly, reducing recall length and increasing visit frequency can reduce both cost and ME in survey data (Beegle et al. 2012, Wollburg et al. 2021). Bounding recall periods and references significantly eases the cognitive burden on respondents and hence minimizes errors (Abate et al. 2022). Carletto et al. (2021) argue that proper enumerator recruitment, training, and monitoring can minimize enumerator-driven ME. Given many prospective sources of survey error, however, addressing all types and causes in a single research design can be impractical; some ME will likely remain even in those data collected with best-practice survey designs.

Although ex post analysis and adjustment may generate some insights to circumvent these data quality deficiencies, addressing ME ex ante through appropriate survey design and technologies is typically the best approach. Once data are collected, researchers have fewer corrective options. Increasing effort by researchers to understand (*a*) data-generating processes, (*b*) units of observation, and (*c*) the local context and actors involved in the design, collection and processing of agricultural data can minimize those MEs.

If the data are already collected and researchers have limited opportunities to address potential mismeasurement, analysts should focus on (*a*) assessing the presence and type of ME in the data, (*b*) evaluating the inferential implications and assessing whether an error should be corrected, and (*c*) if an error should be corrected, understanding how to make those corrections. Several alternative approaches exist to test for the presence and properties of ME. Although some approaches can be employed as diagnostic tests without additional information, most require validation or external information.

Among those diagnostic tests to evaluate the presence of potential mismeasurement in a specific variable of interest, several studies employ Benford's law to test the properties and associated distribution of a specific variable of interest. Judge & Schechter (2009) test the distribution of the first significant digit of various data using Benford's law and identify potential mismeasurement in agricultural surveys in Paraguay. Similarly, Schündeln (2018) employs Benford's law and shows that a longer recall period increases ME in consumption data. However, these types of diagnostic tests are only applicable for sufficiently continuous measures.

Other methods for testing the presence of ME require either proxies (e.g., repeated measures) or external information that could serve as an instrumental variable (IV). For example, some methods estimate linear regressions using ordinary least squares and IV estimators and attribute the differences between the two estimates to ME (e.g., Hausman 1978, Hausman et al. 1998, Mahajan 2006, Hu 2008, Wilhelm 2019, Lee & Wilhelm 2020). Mahajan (2006) proposes a test for the presence of ME for binary explanatory variables. This test requires an additional random variable that is correlated with the unobserved true underlying variable but unrelated to the ME. Wilhelm (2019) extends this approach, relying on testing whether ME distorts key outcomes of interest instead of directly testing the presence or absence of ME. Lee & Wilhelm (2020) provide a Stata package to implement this simple nonparametric test of the no ME hypothesis, which is inferred based on the relevance of ME (i.e., whether measurement distorts expected relationships) by comparing distribution functions for a given variable of interest.

Testing for the presence and properties of ME is considerably easier in the presence of benchmark measurement or validation data. When a researcher has the option to design data collection, building in validation checks on measures of key variables is typically a wise investment of scarce survey time and resources. For example, one can evaluate whether ME behaves classically or nonclassically by regressing ME on the true value of the variable of interest (i.e., either μ onto X^* or ω onto Y^*, considering the expression in Equation 1). A statistically significant correlation between ME (μ and ω) and true values implies NCME. Several studies use GPS-based plot size as a benchmark for this type of empirical test to confirm NCME in self-reported land size and production measures (Desiere & Jolliffe 2018; Abay et al. 2019, 2021, 2022b).

Identifying the presence of ME is not sufficient for understanding the inferential implications, nor is identifying the presence sufficient to decide whether an error should be corrected. As discussed, the distinction between misreporting and misperceptions especially matters for both statistical inference and understanding respondents' actual decision-making processes (e.g., farm management). Considering the expression in Equations 1–3, misperception implies that ME in an explanatory variable (μ) is likely to matter for farm management decisions and hence production outcomes. Abay et al. (2021) employ an empirical test to evaluate the predictive role of measurement error on farm management decisions. In the absence of misperceptions conditional on X^*, the additional information contained in X must be uninformative about farmers' behaviors, a hypothesis that entails testing whether ME in agricultural inputs (μ) predicts input use decisions by directly introducing ME as a regressor in the spirit of a control function approach as follows:

$$Y = \alpha + \beta X^* + \lambda\mu + \epsilon. \qquad 4.$$

If $\lambda = 0$, then there is no evidence of misperceptions. The hypothesis test in Equation 4 requires validation data and is vulnerable to misspecification of the functional form of the relationship between Y and X^* (Wilhelm 2019). To overcome those limitations, Wilhelm (2019) proposed a simple nonparametric test for ME (including NCME).

Wossen et al. (2022) combine self-reported crop variety adoption with DNA fingerprinting results and, assuming these variables as two proxies of the latent variety perception, conduct a simple nonparametric test of the no ME null hypothesis by testing whether ME distorts the conditional mean of farmers' complementary input use functions (particularly fertilizer and herbicide inputs). Wossen et al.'s test explicitly evaluates equality of the input use expectation, conditional on the alternative proxies employed. This is a nonparametric evaluation of how the conditional mean of Y^* depends on the various proxies assumed (Mahajan 2006, Wilhelm 2019, Lee & Wilhelm 2020). These tests do not necessarily imply the absence of ME in the two adoption proxies, only that, if any ME exists, it does not affect the conditional mean of the input use function. When researchers have access to repeated measurements, they may be able to infer the origins of ME by testing the joint conditional expectations for various variables of interest (e.g., land use, consumption, production choices). However, these tests are only as credible as the assumptions used. Thus, it is prudent to scrutinize whether such assumptions are plausible in view of the data-generating process and context.

Whether researchers need (and how) to correct MEs depends mainly on the type/nature of the error and on the purpose of the analysis. Errors that arise due to misperceptions affect both statistical inference as well as respondents' decision-making processes. Thus, in the case of misperception, researchers interested in estimating behavioral parameters associated with respondents' decision-making processes and behavioral responses can benefit from using misperceived values. In contrast, researchers interested in the prediction of actual biophysical phenomena (e.g., yield response to an input) should typically employ objective measurements of these variables. Conceptually, identifying misperception and providing updated information to respondents who display misperceptions may allow them to update mistaken beliefs and thereby improve investment and other management choices and associated outcomes. Moreover, mistaken beliefs about one object may spill over to affect beliefs about other objects, as Abay et al. (2022b) find in the case of plot size misperceptions' effects on self-reports of fertilizer and labor use in Malawi.

Some misperceptions and mistaken beliefs might be driven, however, by respondents' behavioral anomalies such as inattention, self-esteem, or confirmation bias or by local agricultural input market imperfections (e.g., the informality of input markets). Such factors might impede respondents from updating their beliefs in response to information-based interventions (e.g., agricultural extension programming). Abay et al. (2022b) evaluate the impact of an information-based intervention meant to address farmers' misperceptions about plot size and show that learning is limited, heterogeneous, and asymmetric.

To summarize, while some types and sources of ME may be addressed by improving the survey design or measurement protocol or using improved measurement technologies, many other types of ME are less amenable to correction during data collection, and an analyst might not even want to study corrected data. Characterizing and examining the nature of ME are crucial to justify investment in correction or identify appropriate statistical remedies to minimize the consequences of these errors on policy relevant inferences. Similarly, evaluating whether these sources of ME affect just one or multiple variables of interest can inform whether one should invest in correcting ME in agricultural data. Researchers should routinely test for the presence of ME in agricultural data and characterize important features of these errors just as they do other specification tests in data analysis.

Almost all methods aiming to conduct statistical inference in the presence of ME require some additional information, whether validation data or some assumptions on error distributions, repeated measurements, or instruments (Schennach & Hu 2013). Auxiliary or validation data associated with a specific variable of interest can facilitate IV estimation to overcome bias due to ME (Mahajan 2006; Lewbel 2007; Hu 2008; Kreider et al. 2012; Schennach 2016, 2020; Di Traglia & García-Jimeno 2019). For example, the availability of multiple imperfect measurements of the same unobserved underlying true variable can provide a practical way to correct inferential biases associated with ME, particularly when the error in one measure is plausibly classical and hence can serve as an instrument (Schennach 2020). Along these lines, Battistin et al. (2014) use two repeated measures of education qualification to estimate returns to education. When multiple imperfect measurements and IVs are available, Calvi et al. (2022) propose an ME robust LATE (MR-LATE) estimation approach. It may be reasonable to assume that ME (if any) in crop variety identification through DNA fingerprinting or land size measured via GPS is random (i.e., it is CME), allowing researchers to use these measurements as instruments to overcome bias due to ME in self-reported information. Cohen (2019) uses this logic to address potential bias due to ME in land area. However, IV-based correctives require strong assumptions about the nature of ME or the relationship between the instrument and ME. Those assumptions often cannot be validated empirically, especially in the presence of NCME, such as the core IV assumption that the instrument is strongly correlated with the unobserved true underlying variable but unrelated to the ME.

5. BEHAVIORAL IMPLICATIONS OF MEASUREMENT ERROR IN AGRICULTURAL DATA

No data will be error-free. But as discussed above, when ME at least partly represents respondents' mistaken beliefs, it also adds new, potentially valuable information to an analysis. Researchers are beginning to take farmer misperceptions seriously as objects of analysis in their own right.

As discussed by Abay et al. (2022b), the existence and persistence of farmer misperceptions may reflect important behavioral phenomena, including rational inattention (Gabaix et al. 2006, Gabaix 2017), selective inattention (Hanna et al. 2014, Schwartzstein 2014) and confirmation bias (Rabin & Schrag 1999), and motivated reasoning (Nyhan 2020). In the presence of rational or selective inattention, farmers fail to notice and thus fail to learn from information available to them. Confirmation bias and motivated reasoning prompt farmers to pay attention mainly (or only) to new information that confirms their prior beliefs or generates some anticipated payoff.

These behavioral explanations may help explain the existence or persistence of misperceptions, yet they have little to say about their initial formation. The causal mechanisms behind farmer misperceptions—their formation and evolution—are important, nascent research areas, and analytically distinguishing among potential behavioral explanations for ME is an area of research with considerable promise.

Misperceptions about the quality, quantity, and type of agricultural inputs could result from, and be reinforced by, inaccurate information about input quality and purchasing decisions. For example, frequent measurement of land area is not common in many African countries, so farmers form beliefs about the size of their plots through varied, often inaccurate land allocation processes. Similarly, in the absence of lab-based testing facilities and in contexts where agricultural inputs and markets are informal and prone to adulteration, farmers' assessments of fertilizer and pesticide (or herbicide) quality likely suffers from misperceptions (Bold et al. 2017, Ashour et al. 2019, Michelson et al. 2021, Hoel et al. 2022, Maredia & Bartle 2023).

Regardless of the causes of misperception, mistaken beliefs about the quantity and quality of agricultural inputs and related technologies can affect farm management and agricultural investment decisions. A large literature has demonstrated the relevance of farmer uncertainty about inputs and technologies to adoption decisions. Farmer uncertainty about seed identity and traits and uncertainty regarding implementation of the bundled management choices that come with switching from a known variety to a new variety have long been concerns of development economics research, dating back at least to the framework laid out by Feder (1980). Uncertainty about agricultural input quality is associated with lower expected returns, reducing incentives for adoption and use (Bold et al. 2017, Michelson et al. 2021). Bulte et al. (2023) use a randomized controlled trial to show that farmers who are uncertain about whether their maize seed is a hybrid or local variety reallocate labor away from maize production, lowering yields by 30% relative to the complete information scenario.

A slight but important distinction exists between uncertainty and misperception, however. A misperception is an incorrect understanding rather than an understanding characterized by complete uncertainty. Building on the study of the importance of uncertainty to examine farmer misperceptions may provide additional insights into farm management and investment decisions. Recent work suggests a strong relationship between farmer beliefs about the type and quality of different agricultural inputs and farming outcomes. For example, recent studies show significant relationships between farmer beliefs about soil quality, input application, and crop yields (Marenya et al. 2008, Berazneva et al. 2018, Abay et al. 2022a). Mullally (2012) argues that farmers hold mistaken beliefs about the mean of the cotton yield distribution in their region; farmers believe mean yields are higher than the researchers can establish with extant data. As a result, farmers are unwilling to purchase the index insurance offered at an actuarily fair price based on the researcher-established mean yield.

Besides their impact on technology adoption and investment decisions, farmer misperceptions can ultimately reduce agricultural productivity and allocative efficiency. Wossen et al. (2022) show that misperceptions induce crowding-in(out) of complementary agricultural inputs. More importantly, Wossen et al. find that these misperception-driven input allocations are associated with lower yields, implying that such mistaken beliefs may induce allocative or technical inefficiency. Similarly, Mallia (2022) estimates the impact of varietal misperception on farmer decisions, productivity, and profits using data from farmers who either grow an improved maize variety and know it (true positive) or grow an improved variety and think that they do not (false negative). Farmers that correctly perceive their status as improved maize adopters invest in more variable inputs (fertilizers and labor) and have higher yields.

Although the literature on the implication of farmer misperceptions remains nascent, the limited studies to date highlight the potential prevalence, magnitude, and costs of misperceptions and associated inaccurate information in agricultural systems. This implies that identifying and addressing farmer misperceptions and associated agricultural market imperfections may improve farmers' investment choices and productivity outcomes. The degree to which farmers' misperceptions can be corrected likely depends on the kind of misperception and degree to which information sufficiently addresses its underlying behavioral phenomenon, however. For example, mistaken beliefs that arise from agricultural market imperfections and frictions may be addressed by providing access to quality information, verification, and regulatory mechanisms. Information-based campaigns and interventions may also ameliorate farmers' inherent mistaken beliefs, although the empirical validity of this presumption remains largely untested, and the scant existing evidence is not encouraging (e.g., Abay et al. 2022b). Some behavioral phenomena that may generate misperceptions (e.g., inattention, confirmation, and self-esteem bias) might also obstruct learning and updating of mistaken beliefs. These behavioral anomalies and potential systematic variation in these attributes can affect the effectiveness of information interventions as well as heterogeneities across observable and unobservable characteristics of farmers. For example, some individuals are skeptical about, and thus less accepting of, information that is inconsistent with their preferences, a tendency known as directionally motivated reasoning. If individuals scrutinize any information inconsistent with their preferences, information campaigns may need to be tailored asymmetrically based on whether they largely reinforce or counter target recipients' preferences (Ditto & Lopez 1992).

Efforts to correct misperceptions among farmers have achieved mixed success. Maertens et al. (2022) study patterns of farmer beliefs and beliefs updating about urea fertilizer quality in response to a randomized information campaign in Tanzania. They find that farmers exposed to the information campaigns update their beliefs significantly, revising upward their belief that urea is of good quality. They also find that urea purchases increase, especially among farmers who were not previously using the input and who believed quality was bad at baseline. Tamim et al. (2022) find that farmers update their beliefs about the soil fertility of their plots after receiving input vouchers, but that farmers who receive both vouchers and plot testing do not change their beliefs, nor do farmers who receive the information alone. Abay et al. (2022b) use a randomized controlled trial to show that even after farmers are provided GPS measures of their plot size, only a small number correct their self-reports: 37% do not correct at all, 28% partly correct (meaning that they move closer to the GPS-measured value), 13% fully correct, and 13% move from one focal point value to another. Those that are likely to correct are those who underestimated at baseline relative to the GPS measure.

6. IMPLICATIONS FOR FUTURE RESEARCH AND SMALLHOLDER AGRICULTURAL DEVELOPMENT

MEs are pervasive in data of all sorts, and analysts must be alert to those errors. Failure to account for ME can lead to mistaken inferences, including about important policy-relevant phenomena, such as the inverse-size productivity relationship (Desiere & Jolliffe 2018, Abay et al. 2019, Gourlay et al. 2019) or the magnitude of allocative inefficiency in African agriculture (Gollin & Udry 2021). The fact that stylized facts that routinely inform policy design can at least partially be explained by measurement should serve as a caution to empirical analysts about proceeding without careful consideration of the prospective causes and consequences of ME in their data.

The pervasiveness and implications of ME justify further investments to improve measurement of agricultural data, about which two very nice surveys offering excellent guidance have been written recently (De Weerdt et al. 2020, Carletto et al. 2021). But while more careful data collection is essential, one can never fully correct for ME. Evolving scholarship on the nature and implications of ME raises several ripe opportunities for future research.

First, more careful thought needs to be devoted to whether, when, and how to correct for ME. For example, if one is fundamentally interested in farmer behaviors that might be informed by misperceptions, or if NCME is correlated among multiple variables, some of which cannot be corrected, then erroneous data may be—counterintuitively—more appropriate objects for analysis than partly corrected data series. Researchers do not yet have good rules of thumb regarding when to correct and when not to correct.

Second, growing evidence on the pervasiveness of misperceptions within ME suggests the importance of evaluating the quality and veracity of information provided to, rather than received from or about, farmers through both new and old dissemination platforms as a part of agricultural extension or policy evaluation efforts. Inaccurate information can itself be a source of misperceptions or can serve to reinforce existing misperceptions. If misperceptions are associated, especially causally, with underinvestment or underperformance, greater attention needs to be paid to the origins and evolution of farmer misperceptions. Moreover, providing research subjects with inaccurate information can inadvertently cause harm, so researchers need to exercise caution in information experiments (Barrett & Carter 2014). Important differences exist between providing farmers with validated information on their own circumstances (e.g., the GPS-measured size of their plot, the DNA-identified variety of their seed, or the lab chemistry analysis of their soils or purchased fertilizer or pesticides) versus giving them estimates based on large spatial or intertemporal sample averages or statistical models that almost surely differ from the subject's individual-specific parameter value(s).

Third, the introduction and scaling of technologies that dramatically reduce the costs of both collecting information on or from farmers and providing information to farmers raise the stakes and the salience of the issues we discuss in this review. For example, digital agriculture, which uses information communications technology–based means of providing information to farmers, is the focus of a significant and growing industry (reviewed by Birner et al. 2021, Spielman et al. 2021, Abate et al. 2023), and text messages are now a primary digital channel used to reach farmers with information in Africa at scale, at least in part because the costs of providing information texts to smallholder farmers are extremely low.[3] As a consequence, an intervention with only modest effects on knowledge and uptake can potentially deliver positive social returns if the process of generating actionable information for farmers is sufficiently cheap. Yet African agriculture is characterized by considerable heterogeneity in growing conditions across relatively small distances that suggests a need for caution (Tittonell et al. 2005, 2013; Vanlauwe et al. 2019) because any management recommendation is only as accurate as the spatial scale of the most coarsely measured variable used to derive it. Information interventions that are careful about contextualized measurement and external validity and that vet the information appropriately for accuracy may be difficult and expensive to implement, undercutting one of the main arguments for information interventions: their seemingly low unit cost.

The recent literature on ME has focused heavily on survey data. But as greater attention turns to ME in administrative and remotely sensed data, as well as to the evaluation of digital agricultural interventions, appropriate attention should be paid to the errors that such methods may introduce into researchers' analyses but also to the potential negative costs borne by farmers from receiving wrong information at scale.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Abate GT, de Brauw A, Gibson J, Hirvonen K, Wolle A. 2022. Telescoping error in recalled food consumption: evidence from a survey experiment in Ethiopia. *World Bank Econ. Rev.* 36(4):889–908

Abate GT, de Brauw A, Hirvonen K, Wolle A. 2023. Measuring consumption over the phone: evidence from a survey experiment in urban Ethiopia. *J. Dev. Econ.* 161:103026

Abay KA. 2020. Measurement errors in agricultural data and their implications on marginal returns to modern agricultural inputs. *Agric. Econ.* 51(3):323–41

Abay KA, Abate GT, Barrett CB, Bernard T. 2019. Correlated non-classical measurement errors, "second best" policy inference, and the inverse size-productivity relationship in agriculture. *J. Dev. Econ.* 139:171–84

[3]Across Africa, network coverage on cropland remains low: 33% for 3G and 9% for 4G services (Mehrabi et al. 2021), and the adoption of smart phones lags far behind 2G text and voice cell phone adoption.

Abay KA, Abay MH, Amare M, Berhane G, Aynekulu E. 2022a. Mismatch between soil nutrient deficiencies and fertilizer applications: implications for yield responses in Ethiopia. *Agric. Econ.* 53(2):215–30

Abay KA, Barrett CB, Kilic T, Moylan H, Ilukor J, Vundru WD. 2022b. *Nonclassical measurement error and farmers' response to information reveal behavioral anomalies*. Policy Res. Work. Pap. 9908, World Bank Group, New York

Abay KA, Berhane G, Hoddinott J, Tafere K. 2022c. Respondent fatigue reduces dietary diversity scores reported from mobile phone surveys in Ethiopia during the COVID-19 pandemic. *J. Nutr.* 152(10):2269–76

Abay KA, Bevis LEM, Barrett CB. 2021. Measurement error mechanisms matter: agricultural intensification with farmer misperceptions and misreporting. *Am. J. Agric. Econ.* 103(2):498–22

Amaya A, Biemer PP, Kinyon D. 2020. Total error in a big data world: adapting the TSE framework to big data. *J. Surv. Stat. Methodol.* 8(1):89–119

Ambler K, Doss C, Kieran C, Passarelli S. 2021a. He says, she says: spousal disagreement in survey measures of bargaining power. *Econ. Dev. Cult. Change* 69(2):765–88

Ambler K, Herskowitz S, Maredia MK. 2021b. Are we done yet? Response fatigue and rural livelihoods. *J. Dev. Econ.* 153:102736

Arthi V, Beegle K, De Weerdt J, Palacios-López A. 2018. Not your average job: measuring farm labor in Tanzania. *J. Dev. Econ.* 130:160–72

Ashour M, Gilligan DO, Hoel JB, Karachiwalla NI. 2019. Do beliefs about herbicide quality correspond with actual quality in local markets? Evidence from Uganda. *J. Dev. Stud.* 55(6):1285–1306

Baird S, Hamory J, Miguel E. 2008. *Tracking, attrition and data quality in the Kenyan Life Panel Survey Round 1 (KLPS-1)*. Work. Pap., Cent. Int. Dev. Econ. Res., Univ. Calif., Berkeley

Bardasi E, Beegle K, Serneels PM, Dillon A. 2012. Do labor statistics depend on how and to whom the questions are asked? Results from a survey experiment in Tanzania. *World Bank Econ. Rev.* 25(3):418–47

Barrett CB, Carter MR. 2014. Retreat from radical skepticism: rebalancing theory, observational data and randomization in development economics. In *Field Experiments and Their Critics*, ed. DL Teele, pp. 58–77. New Haven, CT: Yale Univ. Press

Battistin E, De Nadai M, Sianesi B. 2014. Misreported schooling, multiple measures and returns to educational qualifications. *J. Econom.* 181(2):136–50

Beaman L, Dillon A. 2012. Do household definitions matter in survey design? Results from a randomized survey experiment in Mali. *J. Dev. Econ.* 98(1):124–35

Beegle K, De Weerdt J, Friedman J, Gibson J. 2012. Methods of household consumption measurement through surveys: experimental results from Tanzania. *J. Dev. Econ.* 98(1):3–18

Berazneva J. 2014. Audio recording of household interviews to ensure data quality. *J. Int. Dev.* 26(2):290–96

Berazneva J, McBride L, Sheahan M, Güereña D. 2018. Empirical assessment of subjective and objective soil fertility metrics in east Africa: implications for researchers and policy makers. *World Dev.* 105:367–82

Berkson J. 1950. Are there two regressions? *J. Am. Stat. Assoc.* 45:164–80

Bevis L, Barrett CB. 2020. Close to the edge: high productivity at plot peripheries and the inverse size-productivity relationship. *J. Dev. Econ.* 143:102377

Biemer PP. 2010. Total survey error: design, implementation, and evaluation. *Public Opin. Q.* 74(5):817–48

Birner R, Daum T, Pray C. 2021. Who drives the digital revolution in agriculture? A review of supply-side trends, players and challenges. *Appl. Econ. Perspect. Policy* 43:1260–85

Bold T, Kaizzi KC, Svensson J, Yanagizawa-Drott D. 2017. Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda. *Q. J. Econ.* 132:1055–1100

Bound J, Brown C, Mathiowetz N. 2001. Measurement error in survey data. In *Handbook of Econometrics*, Vol. 5, ed. JJ Heckman, E Leamer, pp. 3705–43. Amsterdam: Elsevier

Bulte E, Di Falco S, Kassie M, Vollenweider X. 2023. Low-quality seeds, labor supply and economic returns: experimental evidence from Tanzania. *Rev. Econ. Stat.* **https://doi.org/10.1162/rest_a_01285**

Burke M, Lobell DB. 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *PNAS* 114(9):2189–94

Burke WJ, Morgan S, Namonje T, Muyanga M, Mason NM. 2020. *Beyond the "inverse relationship": area mismeasurement affects actual productivity, not just how we understand it*. Res. Pap. 159, Feed Fut. Innov. Lab Food Secur. Policy, Mich. State. Univ., Lansing

Calvi R, Lewbel A, Tommasi D. 2022. LATE with missing or mismeasured treatment. *J. Bus. Econ. Stat.* 40:1701–17

Carletto C, Dillon A, Zezza A. 2021. Agricultural data collection to minimize measurement error and maximize coverage. In *Handbook of Agricultural Economics*, Vol. 5, ed. CB Barrett, DR Just, pp. 4407–80. Amsterdam: Elsevier

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

Carletto C, Savastano S, Zezza A. 2013. Fact or artifact: the impact of measurement errors on the farm size-productivity relationship. *J. Dev.* 103:254–61

Caroli E, Weber-Baghdiguian L. 2016. Self-reported health and gender: the role of social norms. *Soc. Sci. Med.* 1:153:220–29

Chetty R. 2012. Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply. *Econometrica* 80(3):969–18

Cohen A. 2019. Estimating farm production parameters with measurement error in land area. *Econ. Dev. Cult. Change* 68(1):305–34

David IP. 1978. *Non-sampling errors in agricultural surveys. Review, current findings, and suggestions for future research*. Paper presented at the Philippine Statistical Association Annual Conference, June 19, Manila. **https://www.pssc.org.ph/wp-content/pssc-archives/The%20Philippine%20Statistician/1978/Num%201-2/06_Non-Sampling%20Errors%20in%20Agricultural%20Surveys.pdf**

De Nicola F, Giné X. 2014. How accurate are recall data? Evidence from coastal India. *J. Dev. Econ.* 106:52–65

Dervisevic E, Goldstein M. 2023. He said, she said: the impact of gender and marriage perceptions on self and proxy reporting of labor. *J. Dev. Econ.* 161:103028

Desiere S, Jolliffe D. 2018. Land productivity and plot size: Is measurement error driving the inverse relationship? *J. Dev. Econ.* 130(1):84–98

De Weerdt J, Beegle K, Friedman J, Gibson J. 2015. What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxf. Bull. Econ. Stat.* 77(3):466–74

De Weerdt J, Gibson J, Beegle K. 2020. What can we learn from experimenting with survey methods? *Annu. Rev. Resour. Econ.* 12:431–47

Dillon A, Bardasi E, Beegle K, Serneels P. 2012. Explaining variation in child labor statistics. *J. Dev. Econ.* 98(1):136–47

Dillon A, Gourlay S, McGee K, Oseni G. 2019. Land measurement bias and its empirical implications: evidence from a validation exercise. *Econ. Dev. Cult. Change* 67(3):595–624

Dillon A, Karlan D, Udry C, Zinman J. 2020. Good identification, meet good data. *World Dev.* 127:104796

Dillon A, Mensah ER. 2021. *Respondent biases in household surveys*. Glob. Poverty Res. Lab. Work. Pap. 21-103, Kellog Sch. Manag., Northwestern Univ., Chicago

Di Maio M, Fiala N. 2019. Be wary of those who ask: a randomized experiment on the size and determinants of the enumerator effect. *World Bank Econ. Rev.* 34(3):654–69

Di Traglia FJ, García-Jimeno C. 2019. Identifying the effect of a mis-classified, binary, endogenous regressor. *J. Econom.* 209(2):376–90

Ditto PH, Lopez DF. 1992. Motivated skepticism: use of differential decision criteria for preferred and nonpreferred conclusions. *J. Pers. Soc. Psychol.* 63(4):568–84

Drerup T, Enke B, von Gaudecker H-M. 2017. The precision of subjective data and the explanatory power of economic models. *J. Econom.* 200(2):378–89

Feder G. 1980. Farm size, risk aversion and the adoption of new technology under uncertainty. *Oxf. Econ. Pap.* 32(2):263–83

Fermont A, Benson T. 2011. *Estimating yield of food crops grown by smallholder farmers*. Discuss. Pap. 01097, Int. Food Policy Res. Inst., Washington, DC

Foster T, Mieno T, Brozović N. 2020. Satellite-based monitoring of irrigation water use: assessing measurement errors and their implications for agricultural water management policy. *Water Resour. Res.* 56(11):e2020WR028378

Fowlie M, Rubin E, Walker R. 2019. Bringing satellite-based air quality estimates down to earth. *AEA Pap. Proc.* 109:283–88

Gabaix X. 2017. *Behavioral inattention*. NBER Work. Pap. 24096
Gabaix X, Laibson D, Moloche G, Weinberg S. 2006. Costly information acquisition: experimental analysis of a boundedly rational model. *Am. Econ. Rev.* 96(4):1043–68
Gaddis I, Oseni G, Palacios-Lopez A, Pieters J. 2021. Measuring farm labor: survey experimental evidence from Ghana. *World Bank Econ. Rev.* 35(3):604–34
Gaskell GD, Wright DB, O'Muircheartaigh CA. 2000. Telescoping of landmark events: implications for survey research. *Public Opin Q*. 64(1):77–89
Gibson J, Olivia S, Boe-Gibson G. 2020. Night lights in economics: sources and uses. *J. Econ. Surv.* 34(5):955–80
Gibson J, Olivia S, Boe-Gibson G, Li C. 2021. Which night lights data should we use in economics, and where? *J. Dev. Econ.* 149:102602
Gollin D, Udry C. 2021. Heterogeneity, measurement error, and misallocation: evidence from African agriculture. *J. Political Econ.* 129(1):1–80
Gourlay S, Kilic T, Lobell DB. 2019. A new spin on an old debate: errors in farmer-reported production and their implications for inverse scale-productivity relationship in Uganda. *J. Dev. Econ.* 141:102376
Greenleaf AR, Gadiaga A, Guiella G, Turke S, Battle N, Ahmed S, Moreau C. 2020. Comparability of modern contraceptive use estimates between a face-to-face survey and a cellphone survey among women in Burkina Faso. *PLOS ONE* 15(5):e0231819
Hanna R, Mullainathan S, Schwartzstein J. 2014. Learning through noticing: theory and evidence from a field experiment. *Q. J. Econ.* 129(3):1311–53
Hausman JA. 1978. Specification tests in econometrics. *Econometrica* 46:1251–71
Hausman JA, Abrevaya J, Scott-Morton FM. 1998. Misclassification of the dependent variable in a discrete-response setting. *J. Econom.* 87(2):239–69
Himelein K. 2016. Interviewer effects in subjective survey questions: evidence from Timor-Leste. *Int. J. Public Opin. Res.* 28(4):511–33
Hoderlein S, Winter J. 2010. Structural measurement errors in nonseparable models. *J. Econom.* 157:432–40
Hoel J, Michelson H, Norton B. 2022. *Misattribution and uncertainty about beliefs prevent learning*. Work. Pap., Colo. Coll., Colo. Springs. **https://www.hopemichelson.org/_files/ugd/bb24e9_86fa26ec96024e03a165cb01134ea96c.pdf**
Hogset H, Barrett CB. 2010. Social learning, social influence, and projection bias: a caution on inferences based on proxy reporting of peer behavior. *Econ. Dev. Cult. Change* 58(3):563–89
Hu Y. 2008. Identification and estimation of nonlinear models with misclassification error using instrumental variables: a general solution. *J. Econom.* 144(1):27–61
Hu Y, Wansbeek T. 2017. Measurement error models: editors' introduction. *J. Econom.* 200(2):151–53
Huntington-Klein N, Arenas A, Beam E, Bertoni M, Bloem J, et al. 2021. The influence of hidden researcher decisions in applied microeconomics. *Econ. Inq.* 59:944–60
Hyslop DR, Imbens GW. 2001. Bias from classical and other forms of measurement error. *J. Bus. Econom. Stat.* 19(4):475–81
Imai K, Yamamoto T. 2010. Causal inference with differential measurement error: nonparametric identification and sensitivity analysis. *Am. J. Political Sci.* 54(2):543–60
Jain M. 2020. The benefits and pitfalls of using satellite data for causal inference. *Rev. Environ. Econ. Policy* 14(1):157–69
Japec L, Kreuter F, Berg M, Biemer P, Decker P, et al. 2015. Big data in survey research. *Public Opin. Q.* 79(4):839–80
Jouvin M. 2021. *Addressing social desirability bias in child labor measurement: an application to cocoa farms in Côte d'Ivoire*. Econ. Work. Pap. 2021-08, Bordeaux Sch. Econ., Fr.
Judge G, Schechter L. 2009. Detecting problems in survey data using Benford's law. *J. Hum. Resour.* 44(1):1–24
Kilic T, Koolwal GB, Moylan HG. 2020. *Are you being asked? Impacts of respondent selection on measuring employment*. Policy Res. Work. Pap. 9152, World Bank Dev. Data Group, New York
Kilic T, Moylan H, Koolwal G. 2021. Getting the (gender-disaggregated) lay of the land: impact of survey respondent selection on measuring land ownership and rights. *World Dev.* 146:105545
Kilic T, Sohnesen TP. 2019. Same question but different answer: experimental evidence on questionnaire design's impact on poverty measured by proxies. *Rev. Income Wealth* 65(1):144–65

Kosmowski F, Aragaw A, Kilian A, Ambel A, Ilukor J, Yigezu B, et al. 2019. Varietal identification in household surveys: results from three household-based methods against the benchmark of DNA fingerprinting in southern Ethiopia. *Exp. Agric.* 55(3):371–85

Kosmowski F, Chamberlin J, Ayalew H, Sida T, Abay K, Craufurd P. 2021. How accurate are yield estimates from crop cuts? Evidence from smallholder maize farms in Ethiopia. *Food Policy* 102:102122

Kreider B, Pepper JV, Gundersen C, Jolliffe D. 2012. Identifying the effects of SNAP (food stamps) on child health outcomes when participation is endogenous and misreported. *J. Am. Stat. Assoc.* 107(499):958–75

Krumpal I. 2013. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual. Quant.* 47(4):2025–47

Laajaj R, Macours K. 2021. Measuring skills in developing countries. *J. Hum. Resour.* 56(4):1254–95

Lamanna C, Hachhethu K, Chesterman S, Singhal G, Mwongela B, et al. 2019. Strengths and limitations of computer assisted telephone interviews (CATI) for nutrition data collection in rural Kenya. *PLOS ONE* 14(1):e0210050

LoPalo M. 2022. Temperature, worker productivity, and adaptation: evidence from survey data production. *Am. Econ. J. Appl. Econ.* 13(3):101–24

Lavrakas PJ. 2008. *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: Sage

Lee YJ, Wilhelm D. 2020. Testing for the presence of measurement error in Stata. *Stata J.* 20(2):382–404

Lewbel A. 2007. Estimation of average treatment effects with misclassification. *Econometrica* 75(2):537–551

Li Z, Ortiz-Bobea A. 2022. On the timing of relevant weather conditions in agriculture. *J. Agric. Appl. Econ. Assoc.* 1(2):180–95

Lobell DB, Azzari G, Burke M, Gourlay S, Jin Z, Kilic T, Murray S. 2019. Eyes in the sky, boots on the ground: assessing satellite- and ground-based approaches to crop yield measurement and analysis. *Am. J. Agric. Econ.* 102(1):202–19

Lobell DB, Ortiz-Monasterio JI, Falcon WP. 2007. Yield uncertainty at the field scale evaluated with multi-year satellite data. *Agric. Syst.* 92(1–3):76–90

Maertens A, Magomba C, Michelson H. 2022. *Restoring trust: evidence from the fertilizer market in Tanzania*. Work. Pap., Univ. Sussex, UK. **https://www.hopemichelson.org/_files/ugd/bb24e9_1394339f607d45219ffd678c208c33c1.pdf**

Mahajan A. 2006. Identification and estimation of regression models with misclassification. *Econometrica* 74(3):631–65

Mallia P. 2022. *You reap what (you think) you sow? Evidence on farmers' behavioral adjustments in the case of correct crop varietal identification*. Work. Pap., Paris Sch. Econ.

Maredia MK, Bartle B. 2023. Excess demand amid quality misperceptions: the case for low-cost seed quality signalling strategies. *Eur. Rev. Agric. Econ.* 50:360–94

Marenya P, Barrett CB, Gulick T. 2008. *Farmers' perceptions of soil fertility and fertilizer yield response in Kenya*. Work. Pap., Cornell Univ., Ithaca, NY

Mehrabi Z, McDowell MJ, Ricciardi V, Levers C, Diego C, et al. 2021. The global divide in data-driven farming. *Nat. Sustain.* 4:54–160

Meyer BD, Mittag N. 2017. Misclassification in binary choice models. *J. Econom.* 200(2):295–311

Michelson H, Fairbairn A, Ellison B, Maertens A, Manyong V. 2021. Misperceived quality: fertilizer in Tanzania. *J. Dev. Econ.* 148:102579

Michler JD, Josephson A, Kilic T, Murray S. 2022. Privacy protection, measurement error, and the integration of remote sensing and socioeconomic survey data. *J. Dev. Econ.* 158:102927

Moore HE, Rutherfurd ID. 2020. Researching agricultural environmental behaviour: improving the reliability of self-reporting. *J. Rural Stud.* 76:296–304

Mullally C. 2012. *Perceptions and participation: research design with low program enrollment and heterogeneous impacts in development*. Work. Pap., Dep. Agric. Resour. Econ., Univ. Calif., Davis

Norwood FB, Lusk JL. 2011. Social desirability bias in real, hypothetical, and inferred valuation experiments. *Am. J. Agric. Econ.* 93(2):528–34

Nyhan B. 2020. Facts and myths about misperceptions. *J. Econ. Perspect.* 34(3):220–36

Phung TD, Hardeweg B, Praneetvatakul S, Waibel H. 2015. Non-sampling error and data quality: What can we learn from surveys to collect data for vulnerability measurements? *World Dev.* 71:25–35

Rabin M, Schrag JL. 1999. First impressions matter: a model of confirmatory bias. *Q. J. Econ.* 114(1):37–82
Schennach SM. 2016. Recent advances in the measurement error literature. *Annu. Rev. Econ.* 8:341–77
Schennach SM. 2020. Mismeasured and unobserved variables. In *Handbook of Econometrics*, Vol. 7A, ed. S Durlauf, L Hansen, J Heckman, R Matzkin, pp. 487–565. Amsterdam: Elsevier
Schennach SM, Hu Y. 2013. Nonparametric identification and semiparametric estimation of classical measurement error models without side information. *J. Am. Stat. Assoc.* 108:177–86
Schwarz N, Knäuper B, Hippler HJ, Noelle-Neumann E, Clark L. 1991. Rating scales: numeric values may change the meaning of scale labels. *Public Opin. Q.* 55(4):570–82
Schwartzstein J. 2014. Selective attention and learning. *J. Eur. Econ. Assoc.* 12(6):1423–52
Schündeln M. 2018. Multiple visits and data quality in household surveys. *Oxf. Bull. Econ. Stat.* 80(2):380–405
Sharma H, Gibson J. 2019. *Civil war and international migration from Nepal: evidence from a spatial durbin model*. Work. Pap. Econ. 6/19, Univ. Waikato, Hamilton, NZ
Spielman D, Lecoutere E, Makhija S, Van Campenhout B. 2021. Information and communications technology (ICT) and agricultural extension in developing countries. *Annu. Rev. Resour. Econ.* 13:177–201
Sudman S, Bradburn NM. 1973. Effects of time and memory factors on response in surveys. *J. Am. Stat. Assoc.* 68(344):805–15
Tamim A, Harou A, Burke M, Lobell D, Madajewicz M, et al. 2022. *Relaxing credit and information constraints: 5-year experimental evidence from Tanzanian agriculture*. Work. Pap., Dep. Econ., Univ. Calif., Berkeley
Tittonell P, Muriuki A, Klapwijk CJ, Shepherd KD, Coe R, Vanlauwe B. 2013. Soil heterogeneity and soil fertility gradients in smallholder farms of the East African highlands. *Soil Sci. Soc. Am. J.* 77(2):524–38
Tittonell P, Vanlauwe B, Leffelaar PA, Shepherd KD, Giller KE. 2005. Exploring diversity in soil fertility management of smallholder farms in western Kenya: II. Within-farm variability in resource allocation, nutrient flows and soil fertility status. *Agric. Ecosyst. Environ.* 110(3–4):166–84
Vanlauwe B, Coe R, Giller KE. 2019. Beyond averages: new approaches to understand heterogeneity and risk of technology success or failure in smallholder farming. *Exp. Agric.* 55(S1):84–106
Wilhelm D. 2019. *Testing for the presence of measurement error*. Cemmap Work. Pap. CWP45/18, Inst. Fiscal Stud., Dep. Econ., Univ. Coll. London
Wineman A, Njagi T, Anderson CL, Reynolds TW, Alia DY, et al. 2020. A case of mistaken identity? Measuring rates of improved seed adoption in Tanzania using DNA fingerprinting. *J. Agric. Econ.* 71(3):719–41
Wollburg P, Tiberti M, Zezza A. 2021. Recall length and measurement error in agricultural surveys. *Food Policy* 100:102003
Wossen T, Abay KA, Abdoulaye T. 2022. Misperceiving and misreporting input quality: Implications for input use and productivity. *J. Dev. Econ.* 157:102869
Wossen T, Abdoulaye T, Alene A, Nguimkeu P, Feleke S, et al. 2019a. Estimating the productivity impacts of technology adoption in the presence of misclassification. *Am. J. Agric. Econ.* 101(1):1–16
Wossen T, Alene A, Abdoulaye T, Feleke S, Rabbi IY, Manyong V. 2019b. Poverty reduction effects of agricultural technology adoption: the case of improved cassava varieties in Nigeria. *J. Agric. Econ.* 70(2):392–407
Yeh C, Perez A, Driscoll A, Azzari G, Tang Z, et al. 2020. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nat. Commun.* 11(1):2583

Annual Review of Resource Economics

Volume 15, 2023

Contents

Development Economics

Resource Economics

Health and Nutrition Economics

Research and Development

Errata

An online log of corrections to *Annual Review of Resource Economics* articles may be found at http://www.annualreviews.org/errata/resource