Abstract

This report compiles measures commonly used to track agricultural productivity and discusses the ways in which they are subject to error, bias, and other data limitations. Though each measure has limitations, choosing the measure(s) most appropriate to the goals of an analysis and understanding the sources of variation allows for more effective and closely targeted investments and policy and program recommendations, particularly when measures suggest different drivers of productivity growth and links to poverty reduction.

Introduction

Agricultural productivity growth has been empirically linked to poverty reduction across a range of measures for both staple and export crops (Timmer, 1995; Datt & Ravallion, 1998; Mellor, 1999; Fan, Hazell, & Thorat, 1999; Irz, Lin, Thirtle et al., 2001; Thirtle, Irz, Lin et al., 2001; Minten & Barrett, 2008; Byerlee, Dao, & Jackson, 2009; Muyanga, Jayne, & Burke, 2010; Pingali, 2012). Many public and private organizations have thus made it a priority to increase farm productivity, and have invested billions toward this end (Coelli & Rao, 2005; Ludena, Hertel, Preckel et al., 2007; Fuglie, 2008; Fuglie & Schimmelpfennig, 2010; Nin-Pratt & Yu, 2010; O’Sullivan, Banerjee, Gulati et al., 2014; FAO, 2015; USAID, 2015; BMGF, 2015).

This report examines measures of productivity rather than measures of output. Although the literature does not always clearly distinguish between the two, output (e.g., kilograms of crop produced) and productivity (e.g., total value of crop produced per unit of input used) may have a direct or inverse relationship depending on the circumstances, with different consequences for poverty. For example, expanding cultivated area may raise output and profits without raising yields. Alternatively, a yield-enhancing technology may increase output, but if input costs also increase, productivity and profit may remain constant or decline. Use of agricultural best practices may reduce input use while output remains constant, raising productivity. And an input price reduction will lower production costs and increase profit, but may not affect output or productivity. Thus, growth in productivity does not necessarily imply a reduction in poverty (Thirtle et al., 2001).

According to Carletto et al. (2015b), “improvement in the measurement of land productivity has been identified as the highest priority in new research by the Global Strategy, a recent multi-agency initiative endorsed by the United Nations Statistical Commission in February 2010.” Yet measuring land productivity is not simple (Carletto, Gourlay, & Winters, 2015a; Fermont & Benson, 2011), and measuring productivity becomes even more complicated when inputs besides land are taken into account. A range of measures of partial and total factor productivity exist, with differing theoretical and practical justifications. We discuss...
these measures, along with their limitations, in detail below. Because the use of different measures can lead to different understandings of the drivers of productivity growth, considering the comparative merit of each productivity measure and selecting the most appropriate one for the goals of the analysis allows for development of intervention strategies to most closely cater to targeted populations and development objectives.

Defining Productivity Measures

The literature defines and measures smallholder agricultural productivity in several ways, all of which are subject to error, bias, and measurement challenges. This report does not provide a comprehensive survey of productivity measures, but rather discusses the strengths and weaknesses of those used commonly in the published literature. Partial factor productivity measures, including common crop yield, account for only some inputs and outputs, most frequently land area and crop harvest quantity. Table 1 compares partial and total factor productivity measures.

Table 1. Agricultural Productivity Measures

<table>
<thead>
<tr>
<th>PARTIAL FACTOR PRODUCTIVITY MEASURES</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological yield (Gross yield)</td>
<td>$\sum \text{Quantity produced (before harvest or postharvest loss)} \over \sum \text{Area harvested}$</td>
</tr>
<tr>
<td>Harvested yield (Common crop yield)</td>
<td>$\sum \text{Quantity harvested (before postharvest loss)} \over \sum \text{Area harvested}$</td>
</tr>
<tr>
<td>Economic yield</td>
<td>$\sum \text{Quantity available for use (after harvest and postharvest loss)} \over \sum \text{Area harvested}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TOTAL FACTOR PRODUCTIVITY MEASURES</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Factor Productivity</td>
<td>$\ln(y_{it}) = f(x_{it}, t, \beta) + u_{it} - v_{it}$</td>
</tr>
</tbody>
</table>

Yield measurement

Yield includes several distinct measures of land productivity that consider land the sole input and crop production the sole output. These measures differ in when during the crop production cycle the numerator (crop production) and denominator (land area) are measured.

In terms of crop production weight, **biological** or **gross yield** is measured before harvest-related and postharvest loss. **Harvested yield** is measured after harvest-related losses, but before postharvest loss. Biological yield and harvested yield are often not clearly distinguished - it is sometimes unclear from reports whether and how harvest-related losses are accounted for in yield calculations. **Economic yield** takes both
harvest-related and postharvest losses into account, but is rarely used (Fermont & Benson, 2011), though a growing literature on postharvest losses (e.g., rodents), spoilage (e.g., postharvest physiological deterioration in cassava), and contamination (e.g., mycotoxins) underscores the roles of postharvest processing and storage in increasing food available for consumption and securing higher prices (Reynolds et al., 2015).

Land area in yield calculations may be measured at planting (yield by area planted, including all land area planted to a given crop) or at harvest (yield by area harvested, including only plots or sections of plots where harvest occurred). Harvested yield, which designates crop production harvested by area harvested, is by far the most common yield measure reported in the literature (Fermont & Benson, 2011).

However, measuring crop area at the time of harvest ignores potential sources of preharvest crop damage or failure that can reduce harvested area substantially. Smallholder farmers may experience partial or complete area loss between planting and harvesting due to poor germination or damage from pests, flooding, or drought. Additionally, farmers may forego harvesting land because of a lack of labor availability or market opportunities (Kaminski & Christiaensen, 2014). In cases where area harvested is less than area planted, yield by area harvested measures are likely to overestimate farmer productivity by excluding from analysis the portions of plots with null yield (Anderson, Reynolds, & Slakie, 2015). We discuss measurement error and other data limitations in the next section.

All yield measures listed in Table 1 are measures of actual yield. Agronomists and others interested in increasing land productivity also estimate potential yields, or maximum yields attainable under given growing conditions. Agronomists use three primary methods to calculate potential yields: simulation using crop models, field experiments, and measurement of maximum farmer-achieved yields. The gap between potential yield and actual yield is known as the yield gap, and is used by agronomists and policymakers to predict the potential for land productivity increases and to target yield-enhancing interventions (Lobell, Cassman, & Field, 2009; Fischer, Byerlee, & Edmeades, 2009). Because this report focuses on measuring realized productivity, we will not further discuss potential yield or yield gaps.

Alternatives to yield measurement

Use of crop yield measures to proxy for farmer productivity has many limitations and its validity has been questioned (Cassidy et al., 2014; West et al., 2014). First, yield is calculated based on one input (land) and one output (crop production), and does not allow for consideration of other benefits or costs like labor, purchased inputs, or environmental damage (Tittonell & Giller, 2013). Further, most yield measures do not account for intercropping, a common practice among smallholder farmers, and among women in particular, where multiple crops are grown in one plot (Khan et al., 2014). Though intercropping has potential benefits in terms of soil fertility, dietary diversity, and risk management, yield measures typically underestimate the productivity of intercropped plots, because farmers plant each individual crop more sparsely on intercropped plots (Anderson et al., 2015).

Technical efficiency accounts for some costs to the farmer (inputs) besides land, but still is based on crop production as the sole output. It measures farmers' productivity compared to a maximum potential crop output achievable with fixed quantities of basic inputs included in the formula, such as land, seed, fertilizer, and water. The random variable \( u \) can account for sociodemographic factors that affect farmers' technical efficiency (Diagne, 2002). A technically inefficient farmer fails to operate on the production frontier - an equivalent production quantity could be grown by a more efficient farmer using proportionally less of each input. Thus a gain in technical efficiency can be achieved either by increasing production quantity while
Total factor productivity (TFP) captures total crop and livestock output, including intercrops and byproducts for fodder and fuel, and all inputs, including land, labor, seed and agrochemicals, and production technology (Rezek, Campbell, & Rogers, 2011). Livestock activities, which partial factor productivity measures like crop yield do not capture, are important to the welfare of many rural and agricultural populations, so their measurement is important in accurately assessing smallholder wellbeing (Carletto, Jolliffe, & Banerjee, 2015b). A change in total factor productivity may be driven by technical change (introducing new input or output factors) or efficiency change (increasing output without changing factors of production) (Dias Avila & Evenson, 2010). Total factor productivity measures are considered superior to partial factor measures in the literature (Fuglie, 2008; Fuglie & Schimmelpfennig, 2010; Alston, Beddow, & Pardey, 2010; Fermont & Benson, 2011; Alston & Pardey, 2014), but partial factor productivity measures are used far more often because they are much simpler to measure and calculate.

Net farm income takes an approach based in finance, calculating net income by subtracting total fixed and variable costs from gross income (Simonyan & Omolehin, 2012; Birthal, Kumar, Negi et al., 2015; Rada, Wang, & Qin, 2015). Proponents of this method emphasize that the highest-yielding strategy for smallholder farmers (maximizing kg/ha) may not align with the best income-generating strategy per crop (maximizing net kg per dollar spent) or per farm (choice of crops to maximize net income earned per dollar spent). Net farm income is still subject to limitations, as it can be difficult to quantify inputs like labor or land in terms of monetary cost, and even crop value is difficult to estimate when farmers consume most of their production rather than selling it and when market prices vary seasonally (Carletto et al., 2015b).

None of these measures effectively considers long-term costs and benefits, externalities, or risk management in rural economies. For example, farming decisions may have social and environmental consequences (pollution, deforestation, or use of water resources) that affect the broader community. The literature on ecosystem services valuation quantifies some of these costs and benefits, and offers measures that can be used alongside or integrated with a chosen productivity measure. Near-term choices, for example maximizing yields in the short run by drawing down soil nutrients without replenishing them, may limit sustained productivity increases over the long term. Additionally, in order to minimize or mitigate risk, farmers may make choices that maximize neither yield nor farm income. Diversification into non-farm activities to increase income, reduce risk, and smooth consumption across agricultural seasons is a common practice among smallholders, indicating that channels besides productivity growth need to be considered to improve smallholder wellbeing (Reardon, 1997; Bryceson, 2002; Davis, Winters, Carletto, et al., 2010; as cited in Carletto et al., 2015b).

Error and Bias in Productivity Measurement and Analysis

Aggregate-level measures

Productivity measures are subject to error and bias at the national aggregate level as well as at the plot level. The FAOSTAT database of the Food and Agriculture Organization (FAO) of the United Nations, which compiles national-level production and yield estimates, is the most widely cited source of yield data (Sandefur & Glassman, 2015). However, the FAO methodology estimates harvested quantity and area based on reports from national ministries of agriculture, and these estimates may be subject to imprecision and inaccuracy as well as political incentives to over- or under-report yield numbers (Jerven, 2014; Sandefur & Glassman, 2015). The FAO considers only two of the 44 countries in Sub-Saharan Africa to have high standards for data collection. Many countries, especially the poorest, lack both financial and human resources to collect accurate agricultural
And in some cases, FAOSTAT yield estimates differ from those published by ministries of agriculture or national statistical agencies themselves (Carletto et al., 2015b).

Furthermore, yields achieved by smallholder farmers may vary substantially from national average estimates when using common crop yield measurements (Waddington, Li, Dixon, et al., 2010), in part due to the sensitivity of averages to outlier data from very large or very small farms. Anderson et al. (2015) demonstrate this with rice yield data from the 2008-2009 Tanzania National Panel Survey.

**Farm- and plot-level measures**

Measuring inputs and outputs on the farm is also subject to inconsistency and bias. Debate exists in the literature over whether farmer-reported estimates of output quantity and land area are more accurate than weights and distances measured by crop researchers or survey enumerators. Area and quantity harvested can prove quite difficult for farmers to estimate, especially for crops such as cassava that are harvested as needed over several months (Fermont & Benson, 2011). But harvesting and measurement by enumerators or researchers is time-consuming and costly, requiring multiple visits to the farm between planting and harvest, and is more easily implemented for cereal crops or cash crops than for continuously harvested crops (Carletto et al., 2015b).

In the case of land area, farmers frequently overestimate the area of small plots, underestimate the area of large plots, and round to even units, all of which can distort productivity measures, especially yields, which have a sole-input denominator (De Groote & Traoré, 2005; Carletto, Savastano, & Zezza, 2013; Carletto et al., 2015a, 2015b). Further, in multiple-visit surveys such as the Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA), farmers occasionally report irreconcilable plot measures, such as an area harvested larger than the area planted or larger than the farmer-reported plot size, presenting a challenge to analysts. Measurement by GPS or rope and compass is possible for land area, but as with crop production, is time-consuming, costly, and subject to bias. For example, plots that go unmeasured in a survey are systematically different - typically farther away from households or roads, and often larger - from those plots that are measured. In the case of some LSMS-ISA data, so many plots in a sample are missing useable measurement data that analysis is impossible or constrained by low statistical power. Measurement error in GPS technology estimates can range from .5 to 4 meters, which is substantial for very small plots (Carletto et al., 2015b). And with any estimation method, accurate measurement is complicated by irregularly shaped plots, obstructions like stumps or anthills, and sloped plots, for which surface area is larger than productive area (Fermont & Benson, 2011).

Unit conversion adds another layer of complexity to yield measurement. “Heaping” in the data, or clustering around common, even estimates such as one acre for plot area, that results from rounding is less apparent to the analyst when measures are converted from one area or weight unit to another (e.g., from 1 acre to .405 hectares). Survey data does not necessarily include the appropriate conversion factors for analysis. And in some countries, commonly reported units are not standard across regions or districts, leading to complex conversion factor estimation processes (Fermont & Benson, 2011; Carletto et al., 2015a, 2015b). For example, in the Malawi LSMS-ISA, farmers may report harvested quantities in kilograms, or in small or large pails, No. 10 or No. 12 plates, bunches, pieces, bales, baskets, or ox-carts. Further, these non-standard measures correspond to different amounts in different regions of the country, and volume-based measures (pails, plates, ox-carts, 50-kg bags) correspond to different harvest weights (kg) depending on the denseness of the harvested crop and whether or not it has been shelled.
Survey instrument design introduces further imprecision and bias. Intercrops in particular tend to be measured imprecisely, and farmers often must designate a primary crop on the plot. Farmers may be asked to estimate the fraction or percentage of the plot planted or harvested, rather than the area in acres or hectares. Instruments or enumerators may neglect questions entirely: for example, common crop yield (by area harvested) cannot be calculated from the Malawi LSMS-ISA survey data, because farmers were not asked what portion of their field was harvested. Surveys may not ask the condition of the crop when quantity is estimated (for example, maize on the cob, shelled, or ground into flour) (Carletto et al., 2015b). And even coordinated survey initiatives such as the LSMS-ISA are not closely aligned across countries. Differences in survey questions and implementation strategies by national agricultural and statistical agencies (such as timing and resources that affect recall periods) make cross-country analysis difficult.

Farmer-reported measures are also subject to bias and error associated with reliance on recall. Surveys often ask farmers to recall crop area harvested, quantity harvested, and input and output prices over a period of weeks or months. Recent research suggests that farmers are able to remember costly investments or market transactions quite accurately (Beegle, Carletto, & Himelein, 2011), but more bias may be introduced for extended-harvest root and tuber crops and crops consumed at home rather than sold. Deininger et al. (2012) suggest harvest diaries kept by farmers throughout the harvest season as a less biased alternative to recall methods. Periodic data collection via mobile phone may be another less costly alternative to frequent enumerator visits (Dillon, 2012; cited in Carletto et al., 2015b).

The cognitive psychology issues surrounding recall are also a problem for measuring labor inputs. Surveys typically have a long recall period, and require respondents to calculate average hours worked per day and days worked per week on the spot (for themselves and for other household members). Farm labor in particular is difficult to estimate because it is neither regular nor a salient event in farmers’ lives (Merfeld & Anderson, 2015). Further, farmers may not have modern tools such as watches or mobile phones to accurately track hours worked, or may knowingly misreport hours worked for strategic reasons. In an ongoing study in Tanzania, De Weerdt (2015) compares agricultural labor estimates from a group surveyed at the end of the season to a group surveyed weekly. He reports that estimates of hours worked per day tended to be both regular and accurately estimated at the end of the season, while days worked per week was irregular and thus poorly estimated by averages, given that respondents often reported their modal number of days worked per week rather than the mean. This recall bias may lead to overestimation of labor hours per plot by as much as a factor of three, though the bias was less pronounced at the household than the individual level. Exaggerating labor may underestimate productivity (in measures that account for labor inputs) (De Weerdt, 2015).

**Problems faced by analysts**

In working with subnational-level yield data, analysts must make decisions about how to compensate for these shortcomings. Agriculture sample surveys such as the LSMS-ISA suffer from missing data, outlier observations that skew means while being difficult to verify, respondent error, and other flaws. Decisions about whether and how to impute missing data and how to address outlier data (by removing the top and bottom percentile in each calculation, by deleting outlier observations entirely, or by replacing outlier values with mean- or median-based figures) can greatly affect the results of the analysis. In yield regression analysis, there is a lack of consensus on the best regression techniques to address data limitations such as spatial autocorrelation (Lambert, Lowenberg-Deboer, & Bongiovanni, 2004; Tittonell, Shepherd, & Vanlauwe, 2008; Chu Su, 2011). Even the accepted experimental methods for technical efficiency and total factor productivity calculations are constantly challenged and improved (Arnade & Jones, 2011; Hoang & Coelli, 2011; Zúñiga González, 2013; Block, 2014; O’Donnell, 2014).
Productivity Measures Used in Recent Literature

In peer-reviewed agricultural economics journals, authors use a variety of productivity measures as outcome variables. Yield is the most commonly used, but calculation methods such as whether yield is denominated by area planted or area harvested typically are not specified. Forty-four percent of the 25 articles on agricultural productivity in developing countries published in 2015 used an undefined measure of yield. One article specified the use of yield by area planted, four others converted yield to sale value using production value per hectare, and the rest used technical efficiency, total factor productivity, a variety of measures of farm and non-farm income, and other measures including value added per land area and yield treatment effect less cost treatment effect. Figure 1 shows the productivity measures used in these 25 peer-reviewed articles.

*Figure 1. Productivity Measures Used in 2015 in Peer-Reviewed Agricultural Economics Journals*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Times measure used (n=25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield by area planted</td>
<td>4%</td>
</tr>
<tr>
<td>Yield, unspecified</td>
<td>44%</td>
</tr>
<tr>
<td>Total Factor Productivity</td>
<td>12%</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>8%</td>
</tr>
<tr>
<td>Farm income</td>
<td>28%</td>
</tr>
<tr>
<td>Value of production/ha</td>
<td>16%</td>
</tr>
<tr>
<td>Other</td>
<td>8%</td>
</tr>
</tbody>
</table>

*Note:* “Farm income” includes both gross and net measures of income. Some studies used more than one productivity measure.

*Sources: Agricultural Economics, American Journal of Agricultural Economics, Canadian Journal of Agricultural Economics, Food Policy, International Journal of Agricultural Economics, 2015*

Implications of Productivity Measurement Choices

The choice of a productivity measure can have real implications for results and recommendations. As Anderson et al. (2015) show in their analysis of the 2008-09 Tanzania National Panel Study, conducting analysis based on yield by area harvested, which overestimates the yields of farmers who plant more area than they harvest, can lead to different recommendations for policies or interventions to raise productivity when compared to analyses of yield by area planted. Using regression analysis, variation in rice yield measured by area harvested was explained principally by soil fertility, while variation in yield by area planted was instead explained by rainfall levels and by access to markets and to hired labor. Further, estimates of yield by area harvested compared to yield by area planted were significantly different between respondents with daily consumption

---

1 We surveyed articles with productivity as an outcome variable published in 2015 in Agricultural Economics, American Journal of Agricultural Economics, Canadian Journal of Agricultural Economics, Food Policy, and International Journal of Agricultural Economics. We searched European Journal of Agricultural Economics and found no relevant articles. We excluded articles on productivity measurement of processed agricultural products and on industrialized country agriculture.
greater than $1.25/day per adult equivalent and those with daily consumption less than $1.25/day, suggesting that the choice of yield measure, and the corresponding estimates of the most important drivers of yield variation and hence likely effective investments, varies systematically by consumption level.\(^2\) Interventions intending to target the most vulnerable and least productive farmers - those most likely to lose area between planting and harvesting - might therefore be designed quite differently when based on analysis using the area planted measure rather than the area harvested measure.

Some researchers are choosing to go beyond yield and include additional inputs and outputs as well as social and environmental costs and benefits in their measures of productivity, as well as attempting to estimate productivity gaps rather than yield gaps (Ball, Lovell, Luu, et al., 2004; Byerlee & Murgai, 2005; Di Falco & Chavas, 2006; World Bank, 2010; Hoang & Coelli, 2011; Zúñiga González, 2013). Total factor productivity and net farm income may more appropriately proxy for farm family wellbeing than yield does, as farmers are likely to make choices that maximize income per dollar spent or minimize upfront investment rather than maximizing yield alone. There are no widely-accepted measures that take risk into account, even though controlling or mitigating risk is likely very important in crop mix and agriculture investment decisions for farmers without insurance, savings, or access to capital. Recent research on crop insurance has begun to measure yield risk (Huang, Wang, & Wang, 2015). Considering the strengths and weaknesses of each measure (as shown in Table 2) together with the goals of the analysis allows analysts and policymakers to better understand the drivers of productivity growth, and thus to more effectively plan strategies that both increase productivity and reduce poverty.

Table 2. Characteristics of Agricultural Productivity Measures

<table>
<thead>
<tr>
<th>Yield measure</th>
<th>Counts land input and crop output</th>
<th>Counts all agricultural inputs</th>
<th>Counts all agricultural outputs</th>
<th>Counts monetary costs to the farmer</th>
<th>Counts monetary benefits to the farmer</th>
<th>Counts social, environmental, or risk costs and benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial factor measures</td>
<td>Yield by area harvested</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yield by area planted</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Production value per area</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Technical efficiency</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total factor measures</td>
<td>Total Factor Productivity (Malmquist index)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>Environmental can be included</td>
</tr>
<tr>
<td></td>
<td>Net farm income</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

Please direct comments or questions about this research to Principal Investigators Leigh Anderson and Travis Reynolds at epar.evans.uw@gmail.com.

\(^2\) Significant at 90% confidence level, p=0.083.
References


Additional References: Recently Published Productivity Studies

The following 25 studies are results of a search for literature published in 2015 that measured agricultural productivity as an outcome in six peer-reviewed agricultural economics journals: Agricultural Economics, American Journal of Agricultural Economics, Canadian Journal of Agricultural Economics, Food Policy, International Journal of Agricultural Economics, and European Journal of Agricultural Economics. European Journal of Agricultural Economics published no relevant articles in 2015.


