INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

IFPRI Discussion Paper 01925

May 2020

Improving Data Quality for the CAADP Biennial Review

A Partnership Initiative Piloted in Five Countries

Samuel Benin

Joseph Karugia

Greenwell Matchaya

Mbaye Yade

Africa Regional Office

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), a CGIAR Research Center established in 1975, provides evidence-based policy solutions to sustainably end hunger and malnutrition and reduce poverty. The Institute conducts research, communicates results, optimizes partnerships, and builds capacity to ensure sustainable food production, promote healthy food systems, improve markets and trade, transform agriculture, build resilience, and strengthen institutions and governance. Gender is considered in all of the Institute's work. IFPRI collaborates with partners around the world, including development implementers, public institutions, the private sector, and farmers' organizations, to ensure that local, national, regional, and global food policies are based on evidence. IFPRI is a member of the CGIAR Consortium.

AUTHORS

Samuel Benin (s.benin@cgiar.org) is Deputy Director for the Africa Regional Office of the International Food Policy Research Institute (IFPRI), based in Davis, California.

Joseph Karugia (j.karugia@cgiar.org) is Coordinator for the Regional Strategic Analysis and Knowledge Support System (ReSAKSS) for Eastern and Central Africa at the International Livestock Research Institute (ILRI), based in Nairobi, Kenya

Greenwell Matchaya (g.matchaya@cgiar.org) is a Senior International Economics Researcher at the International Water Management Institute (IWMI) and Coordinator for ReSAKSS Southern Africa at IWMI, based in Pretoria, South Africa

Mbaye Yade (m.yade@cgiar.org) is Coordinator for ReSAKSS West Africa at the International Institute for Tropical Agriculture (IITA), based in Ibadan, Nigeria.

Notices

[1] IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by IFPRI.

[2] The boundaries and names shown, and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

[3] Copyright remains with the authors. The authors are free to proceed, without further IFPRI permission, to publish this paper, or any revised version of it, in outlets such as journals, books, and other publications.

Contents

Tables

Figures

ABSTRACT

This paper presents results of a data partnership framework for strengthening evidence-based planning and implementation that was initiated in 2019 in five selected African countries (Kenya, Malawi, Mozambique, Senegal, and Togo) during the second round of the CAADP biennial review (BR) process. It analyzes the effect of the activities conducted on the data reporting rate and the quality of data reported in the five pilot countries, compared with what was achieved in like-pilot countries. The like-pilot countries are non-pilot countries that have characteristics like the pilot countries at the baseline which affect selection into the pilot or the data reporting and quality outcomes. Different methods (standard deviations, propensity score matching, and two-stage weighted regression) are used to identify the like-pilot countries, and a difference-in-difference method is used to estimate the effect of the pilot activities on the outcomes.

The capacity-strengthening activities focused on working with the country Biennial Review (BR) team to: assess the inaugural or 2018 BR process and identify the data gaps; constitute and train members of data clusters to compile and check the data for the 2020 BR; and then validate and submit the data. The findings show that the activities helped the pilot countries to improve their performance in the data reporting rate and the quality of data reported in the 2020 BR. The largest improvement is observed in Togo and Senegal, followed by Kenya and Malawi, and then Mozambique.

The average increase in the data reporting rate between 2018 and 2020 BRs for the pilot countries is greater than the average progress made in the like-pilot countries by about 6 to 9 % pts. This derives mostly from improvements in the data reporting rate for the indicators under theme 3 on ending hunger. Regarding the quality of data reported (measured as the percent of the data reported that have issues) too, the pilot countries on average performed better than the like-pilot countries, especially with respect to the data reported under themes 2 on investment in agriculture and 3 on ending hunger. But most of the estimated differences have low or no statistical significance. Implications for sustaining the progress made in the pilot countries, as well as for extending the activities to other countries, for the next rounds of the BR are discussed.

ACKNOWLEDGEMENTS

The authors acknowledge various contributions from several individuals. They include: Paul Guthiga and Nancy Munyiva Laibuni (for Kenya); Readwell Musopole, Emmanuel Jofilisi Mwanaleza, and Zephania Bondera Nyirenda (for Malawi); Benedito Cunguara, Wilson Duque, Sophia Manussa, Joao Mutondo, Delfim Villisa, and Helder Zavale (for Mozambique); Sibusiso Nhlenengethwa (for Malawi and Mozambique); Momath Cisse (for Senegal); and Kenao Lao and Maurice Taondeyande (for Togo). Comments by John Ulimwengu on a previous version of the paper are acknowledged. Funding for this work is from the Bill & Melinda Gates Foundation.

ABBREVIATIONS AND ACRONYMS

%pts	percentage points
2SWRM	two-stage weighted regression method
AATS	Agricultural Transformation Scorecard
ASCI	Agricultural Statistics Capacity Index
ASWG	agricultural sector working group
AU	African Union
AUC	African Union Commission
BR	biennial review
BRS	biennial review score
CAADP	Comprehensive Africa Agriculture Development Programme
CSO	civil society organization
DID	difference-in-difference
DRR	data reporting rate
eBR	electronic or online biennial review
FBO	farmer-based organization
ICT	information and communication technology
IFPRI	International Food Policy Research Institute
IITA	International Institute for Tropical Agriculture
ILRI	International Livestock Research Institute
IWMI	International Water Management Institute
JSR	joint sector review
LAN	local analytical network
LP	like-pilot
M&E	monitoring and evaluation
NAG	non-agricultural sector
NEPAD	New Partnership for Africa's Development
NSA	non-state actor
PAPA	Projet d'Appui Aux Politiques Agricoles
PSMM	propensity score matching method
QDR	quality of data reported
REC	regional economic community
ReSAKSS	Regional Strategic Analysis and Knowledge Support System
SAKSS	Strategic Analysis and Knowledge Support System
s.d.	standard deviation
SDM	standard deviation method
TWG	technical working group

1. INTRODUCTION

In the 2014 Malabo Declaration on Accelerated African Agricultural Growth and Transformation for Shared Prosperity and Improved Livelihood, the African leaders called for a biennial review (BR) that involves tracking, monitoring and reporting on progress in seven commitments: (1) recommitting to the principles and values of the Comprehensive Africa Agriculture Development Programme (CAADP); (2) enhancing investment finance in agriculture; (3) ending hunger in Africa by 2025; (4) reducing poverty by half by 2025 through inclusive agricultural growth and transformation; (5) boosting intra-African trade in agricultural commodities and services; (6) enhancing resilience of livelihoods and production systems to climate variability and other related risks; and (7) ensuring mutual accountability to actions and results (AU 2014). As such, the need for credible, timely, and high-quality data and knowledge are essential, not only for reporting on progress on implementing such an ambitious Declaration, but also for guiding sound agricultural policymaking and investment.

The BR process was launched with the development of: i) Technical Guidelines that profile the indicators used to assess progress toward meeting the seven Malabo commitments; ii) a Country Performance Reporting Template used by each country to collect and report on data, and iii) a Technical Note on the scorecard methodology to evaluate progress (whether "on-track" or "not-on-track") at the country, regional, and continental levels for meeting each of the seven Malabo commitments separately as well as together. The inaugural BR report, including the Africa Agricultural Transformation Scorecard (AATS), was launched on January 29, 2018 during the 30th Ordinary Session of the Assembly of Heads of State and Government of the African Union (AU), held in Addis Ababa, Ethiopia. The launch of the inaugural BR report and AATS marked an important milestone in promoting mutual accountability at the highest political level and in strengthening evidence-based agricultural planning and implementation. This is highly commendable as it is first of its kind in the agricultural sector at the AU level. Out of the 55-member states, 47 (85%) submitted country BR reports, out of which 20 were assessed to be on-track to achieving the Malabo commitments by 2025 (AUC 2018).

The inaugural BR process and report highlighted the problems with agricultural data and statistics in many African countries. Data on several of the indicators to be reported on were missing or had measurement issues (Benin et al. 2018). Some of the main reasons cited for these include lack of a centralized agricultural database, lack of awareness of available data sources, inadequate capacities for data collection and transformation, poor data management, and lack of funding, among others (Matchaya et al. 2018). Another factor is limiting data collection to mostly those that are known, generated, or endorsed by the government, which may reflect the politicization of data. Whereas governments and politicians may appreciate the benefits of statistics and data for improving their policy decisions, they may also fear attribution of poor outcomes to their decisions and may intervene in the data-generating process and affect the quality of the data (Jerven 2013, 2014). Related to the inadequate capacity, other available data sources were not used because they were in a different format than was required and the BR country team did not know how to transform the data into the required format. Most of such data were those generated by independent think tanks, universities, researchers, and external organizations for purposes other than the BR. Whereas data and knowledge produced by these entities tend to be of high statistical quality, their policy relevance and timeliness have been questioned. Similarly, uses of international databases such as FAOStat and the World Development Indicators (WDI) in the BR process have been limited, but due perhaps more to the criticism of their poor data-filling procedures (Jerven 2013).

To help improve the quality of data available for policymaking in CAADP implementation to achieve the Malabo commitments, IFPRI-ReSAKSS, with funding from the Gates Foundation and in collaboration with its implementing partners,[1] initiated a partnership framework for strengthening evidence-based planning and implementation in five selected African countries (Kenya, Malawi, Mozambique, Senegal, and Togo) in 2019 during the second round of the BR. This paper presents the results of the initiative by analyzing its effect on the BR data reporting rate and the quality of data reported in the five pilot countries, compared with achievements in like-pilot countries. The like-pilot countries are non-pilot countries that have characteristics like the pilot countries at the baseline which affect selection into the pilot or the data reporting and quality outcomes. Different methods (standard deviations, propensity score matching, and two-stage weighted regression) are used to identify the like-pilot countries, and a difference-in-difference method is used to estimate the effect of the pilot activities on the outcomes.

In the partnership framework that was initiated, capacity-strengthening focused on working with the country team to: i) assess the 2018 BR process and identify the data gaps; ii) constitute and train members of data clusters to compile and check the data; iii) and then validate and submit the data. Overall, we find that the activities helped the pilot countries to achieve higher reporting rates in the 2020 BR, with the largest improvement in Togo and Senegal, followed by Kenya and Malawi, and then Mozambique. The average increase in the reporting rate between 2018 and 2020 for the pilot countries is greater than the average progress made in the like-pilot countries by about 6 to 9 % pts, which derived mostly from improvements in the data for the indicators under theme 3 on ending hunger. On the quality of data reported (measured as the percent of the data reported that have issues), although the pilot countries on average performed better than the like-pilot countries, especially with the data reported under themes 2 on investment in agriculture and 3 on ending hunger, most of the estimated differences have low or no statistical significance.

The conceptual framework for strengthening evidence-based planning and implementation is presented in the next section, followed by a description of the pilot activities and the empirical approach used for the comparative analysis in sections 3 and 4, respectively. The results are presented in section 5 and the conclusions and implications in section 6.

---

[1] These are the Africa CG centers: International Institute for Tropical Agriculture (IITA), International Livestock Research Institute (ILRI), and International Water Management Institute (IWMI).

2. CONCEPTUAL FRAMEWORK

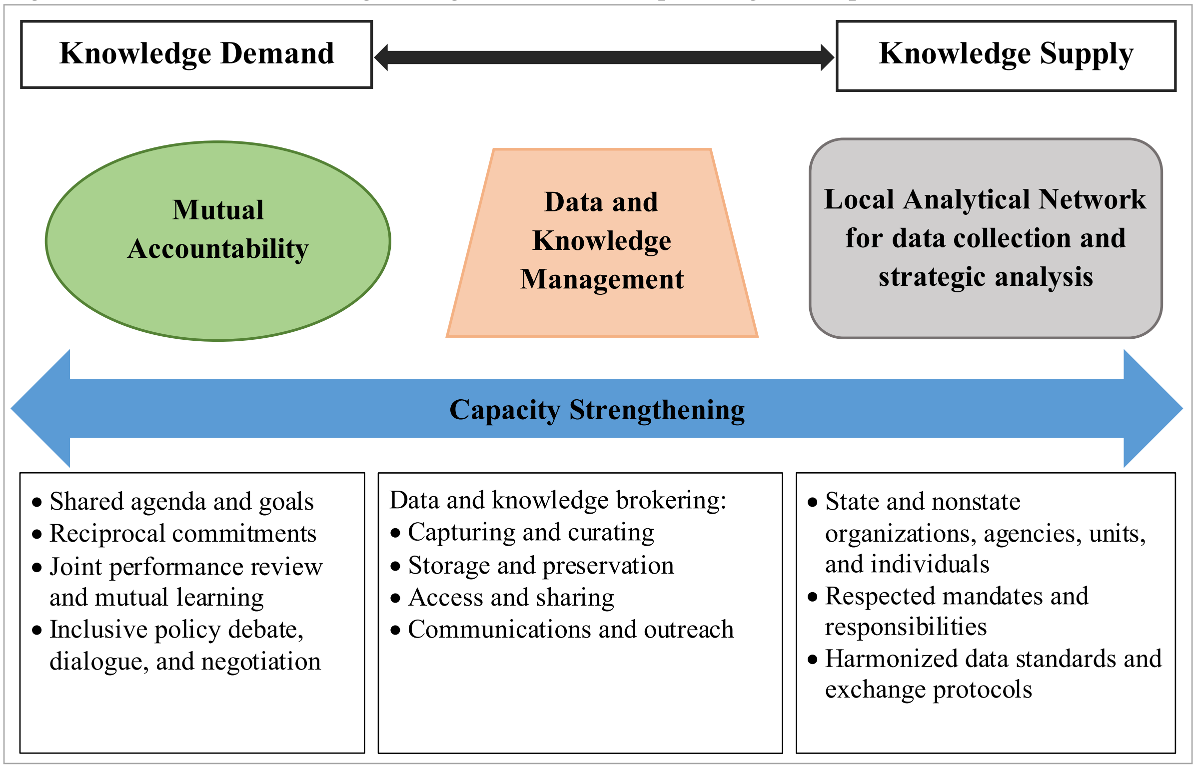
Structure

The conceptual framework for improving data quality to strengthen evidence-based planning and implementation is based on a system for linking knowledge demand and supply, harnessing the existing capacities of different institutions and individuals involved with the generation or use of evidence to inform policy decisions (Figure 1). As no single organization or individual is likely to be successful in having sole responsibility over the diverse data and knowledge relevant for desired agricultural policymaking, the framework is a multi-stakeholder partnership. The knowledge supply part is based on a local analytical network (LAN) that is made up of both state and non-state organizations (e.g., government agencies, research institutes, universities, development organizations, private sector businesses, and farmer-based and civil-society organizations) and individuals (e.g., researchers, policymakers and development practitioners). The LAN is organized into data clusters, with different working groups tasked with collecting, analyzing, and managing specific datasets that are aligned with the mandates and responsibilities of the organizations and individuals within the working group (Figure 2). Thus, harmonization of data standards and exchange protocols across the clusters is important, in addition to having nimble coordination and governance structures to ensure that the data and knowledge generated are of high quality and timely. Essentially, the LAN ensures that a broad array of credible and relevant information is available, and that all potential data providers and analysts are viewed and treated as part of the system's knowledge-supply infrastructure (NASEM 2017).

The knowledge demand part of the system is driven by mutual accountability, which can be defined as a process by which two or multiple parties agree to be held responsible for the commitments that they have made to each other [see e.g., Vance et al. (2013) and Steer and Wathne (2009)]. Mutual accountability is built around 3 key elements: first, the process starts with a shared agenda through clear, specified goals and reciprocal commitments that generate actions; second, these actions need to be jointly monitored and reviewed to assess whether and the extent to which commitments are held; and third, based on evidence, parties engage in debate, dialogue and negotiation around goals and actions. The parties here too include both state and non-state institutions and individuals.

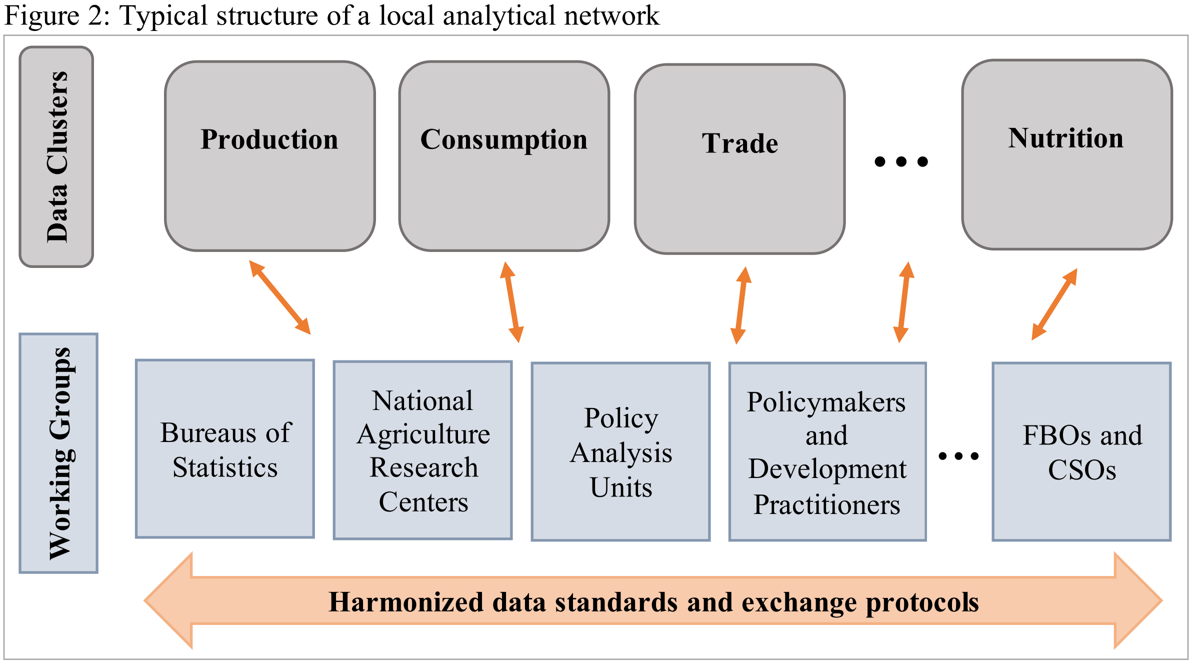
The knowledge demand and supply parts of the system are bridged by the data and knowledge management function. Here, the shared agenda (with the accompanying goals and reciprocal commitments and actions) is made accessible to the LAN to use it to plan their data collection and strategic analysis. Similarly, the data and knowledge generated by the LAN are made accessible as evidence to the parties in the mutual accountability process. Thus, the data and knowledge management function are essentially a brokering function by capturing, curating, storing, and sharing data and knowledge in the policy ecosystem, as well as undertaking broader communications and outreach activities.

Figure 1: Framework for strengthening evidence-based planning and implementation



Source: Authors' illustration.
Notes: FBOs = farmer-based organizations. CSOs = civil society organizations.

Figure 2: Typical structure of a local analytical network



Source: Authors' illustration.
Notes: FBOs = farmer-based organizations. CSOs = civil society organizations.

Capacity strengthening is cross-cutting, with differentiated activities targeted at different parts of the system. With the partnership framework, capacity strengthening may target three different levels: the partners, their relationships, or the system (Hartwich et al. 2007). At the partners level, this involves motivating, providing incentives for, and enhancing the ability of the partners to develop and maintain relationships (linkages, partnerships, agreements, networks, etc.), collaborate, share, and learn from each other. At the relationships level, capacity strengthening focuses on skills and tools for effective communication, negotiation, conflict resolution, and for building trust and social capital (shared agendas, resources, methods, data, knowledge, etc.)—which are critical for successful partnerships (Spielman and von Grebmer 2004). Capacity strengthening at the system level targets the broader political-economy factors (especially policies and institutions—laws, rules, and regulations) that determine or condition the generation or use of data in the agricultural policy ecosystem. Capacity strengthening at the system level may include for example advocacy for rights-to-information laws and open-data access policies for publicly-funded data (Roberts 2010, Yerramareddy and Babu 2018), and data standards and regulation on sampling and sample sizes for national data generated by surveys (NASEM 2017, Citro 2014). The specific capacity strengthening activities must be based on a capacity needs assessment involving all the various partners and stakeholders in the system.

Measures of success

As the motivation for the framework is improving the evidence used to inform the agents in the policy process, rather than to ensure a policy outcome, the metric of success or impact is about the quality of the data or evidence that is utilized in the mutual accountability process. There are many dimensions or attributes of quality that are discussed in the vast literature on data quality (see e.g.: Brodie 1980; Fox et al. 1994; Wang and Strong 1996; Loshin 2011). A survey by Wang and Strong (1996), for example, resulted in 179 attributes—see Figure 3. Here, we focus on defining several of the common attributes including relevance, accuracy, frequency, completeness, consistency, transparency, traceability, validity, reliability, and timeliness of the available data. Each of these attributes embody many of those listed in Figure 3.

Relevance is a measure of the applicability of the data to the analysis or task that the data are required for. For example, to generate evidence on the effect of globalization on gender and youth employment in agriculture will require labor data that are disaggregated by sector, sex, and age for example. Following from the concept of “fit to use” as fundamental to defining data quality (Wang and Strong 1996), then relevance seems a precursor to the other traits. In other words, if data are not relevant, it is useless to worry about their accuracy, reliability, and timeliness, for example.

Figure 3: Data quality attributes

Ability to be joined with	Ability to download	Ability to identify	Ability to upload
Acceptability	Access by competition	Accessibility	**Accuracy**
Adaptability	Adequate detail	Adequate volume	Aestheticism
Age	Aggregability	Alterability	Amount of data
Auditable	Authority	Availability	Believability
Breadth of data	Brevity	Certified data	Clarity
Clarity of origin	Clear data responsibility	Compactness	Compatibility
Competitive edge	**Completeness**	Comprehensiveness	Compressibility
Concise	Conciseness	Confidentiality	Conformity
Consistency	Content	Context	Continuity
Convenience	Correctness	Corruption	Cost
Cost of accuracy	Cost of collection	Creativity	Critical
Current	Customizability	Data hierarchy	Data improves efficiency
Data overload	Definability	Dependability	Depth of data
Detail	Detailed source	Dispersed	Distinguishable updated files
Dynamic	Ease of access	Ease of comparison	Ease of correlation
Ease of data exchange	Ease of maintenance	Ease of retrieval	Ease of understanding
Ease of update	Ease of use	Easy to change	Easy to question
Efficiency	Endurance	Enlightening	Ergonomic
Error-free	Expandability	Expense	Extendibility
Extensibility	Extent	Finalization	Flawlessness
Flexibility	Form of presentation	Format	
Friendliness	Generality	Habit	Historical compatibility
Importance	Inconsistencies	Integration	Integrity
Interactive	Interesting	Level of abstraction	Level of standardization
Localized	Logically connected	Manageability	Manipulatable
Measurable	Medium	Meets requirements	Minimality
Modularity	Narrowly defined	No lost information	Normality
Novelty	Objectivity	Optimality	Orderliness
Origin	Parsimony	Partitionability	Past experience
Pedigree	Personalized	Pertinent	Portability
Preciseness	Precision	Proprietary nature	Purpose
Quantity	Rationality	Redundancy	Regularity of format
Relevance	**Reliability**	Repetitive	Reproducibility
Reputation	Resolution of graphics	Responsibility	Retrievability
Revealing	Reviewability	Rigidity	Robustness
Scope of information	Secrecy	Security	Self-correcting
Semantic interpretation	Semantics	Size	Source
Specificity	Speed	Stability	Storage
Synchronization	Time-independence	**Timeliness**	**Traceable**
Translatable	Transportability	Unambiguity	Unbiassed
Understandable	Uniqueness	Unorganized	Up-to-date
Usable	Usefulness	User friendly	**Valid**
Value	Variability	Variety	Verifiable
Volatility	Well-documented	Well-presented	

Source: Wang and Strong (1996).

Notes: The traits in boldface are explicitly defined in the text and they are composites or variants of several of the other traits.

Once data are available, their accuracy seems to be the most important quality trait in the sense that each data point must be of the correct value, reflect the actual underlying information, and free of error. Common causes of inaccurate data derive from initial measurement or recording errors. Another cause may be due to decay or obsolescence. Continuing with the labor data for example, failure to update the data due to changes in sector or age will result in inaccuracies. Transferring data between different data management or storage formats (including computers and software) may also create errors. Same with transformations that are done incorrectly—for example, converting population data that are in single units into million units but dividing by 1,000 instead of 1,000,000.

Data frequency refers to the regular time intervals that the data are collected and recorded, such as hourly, daily, weekly, monthly, quarterly, annually, decadal, etc. Data on factors that are rapidly changing such as prices may need to be collected more frequently such as hourly, daily or weekly, whereas those that are changing at a moderate pace such as employment may need to be collected less frequently such as monthly or quarterly. Data on factors that are slowly changing such as population structure may need even lower frequency such as annually or even biennially. Often, data may be transformed to obtain either greater or lower frequency data points compared to the frequency at which the original data were collected. Thus, data reported at lower frequencies may be aggregated or transformed from data collected at higher frequencies, commonly by either summing or averaging. In some cases, data reported at a regular frequency, say annual, may be transformed from data collected at a lower frequency, say biennially or every four or years, commonly by extrapolating or assuming a linear trend between two data points for example.

Related to data accuracy are data consistency and data completeness. Consistency, which may be a subcategory of accuracy, refers to data whose value or meaning are not changed during processing or transfers. This is especially crucial to the functioning of applications, databases, programs, and systems. Continuing with the labor data for example, using different definitions of agriculture (e.g. including or excluding forestry) or age categories (e.g. youth) in different locations or different years in the same location will yield inconsistent data. Completeness refers to the extent to which the data are of enough breadth, depth, and scope for the analysis or task that the data are required for. A common way to determine completeness is from the metadata (data about data) or the relevant information or records that are included with the data, such as title, description, location, and names and dates associated with the creation, modification or transformation of the data.

Data transparency and traceability are closely related. Whereas transparency is about the ability to access and work with data no matter where the data are located or the format they are in, traceability is about the ability to trace the history of the data. Thus, traceability will depend on how well the data has been tracked through the data supply chain, which can be recorded with the metadata. Together, data transparency and traceability can help evaluate data accuracy or the extent to which the relevant data are accurate and are derived from the original source.

Validity is about the precision and exactness of the results that are acquired or derived from the relevant data. As such, validity plays a significant role in the policy ecosystem as it ensures the conclusions or implications of the relevant data, either in terms of cause-effect relationship (internal validity) or generalizations to the entire population (external validity). Therefore, the data validation process may perhaps be one of the most critical elements with data in the mutual accountability process, given that the parties must take responsibility for their actions as implied by the data or evidence.

This leads to the issue of data reliability, which encompasses virtually all the traits discussed above. Reliability may be defined as data that are reasonably complete and accurate, meet the intended purposes that the data are required for (relevance), and are not subject to inappropriate alteration (consistency). Usually, reliability is closely linked with the integrity of the data-generating entity in terms of key principles—relevance to policy issues, credibility among data users, trust among data providers, and independence from political and other undue external influence (NASEM 2017). Then, in the policy ecosystem or in the mutual accountability process, there is no way that data that are determined to be unreliable by any one party will be validated. Similarly, since data that are validated by all parties must be treated as reliable, the data validation process becomes critical.

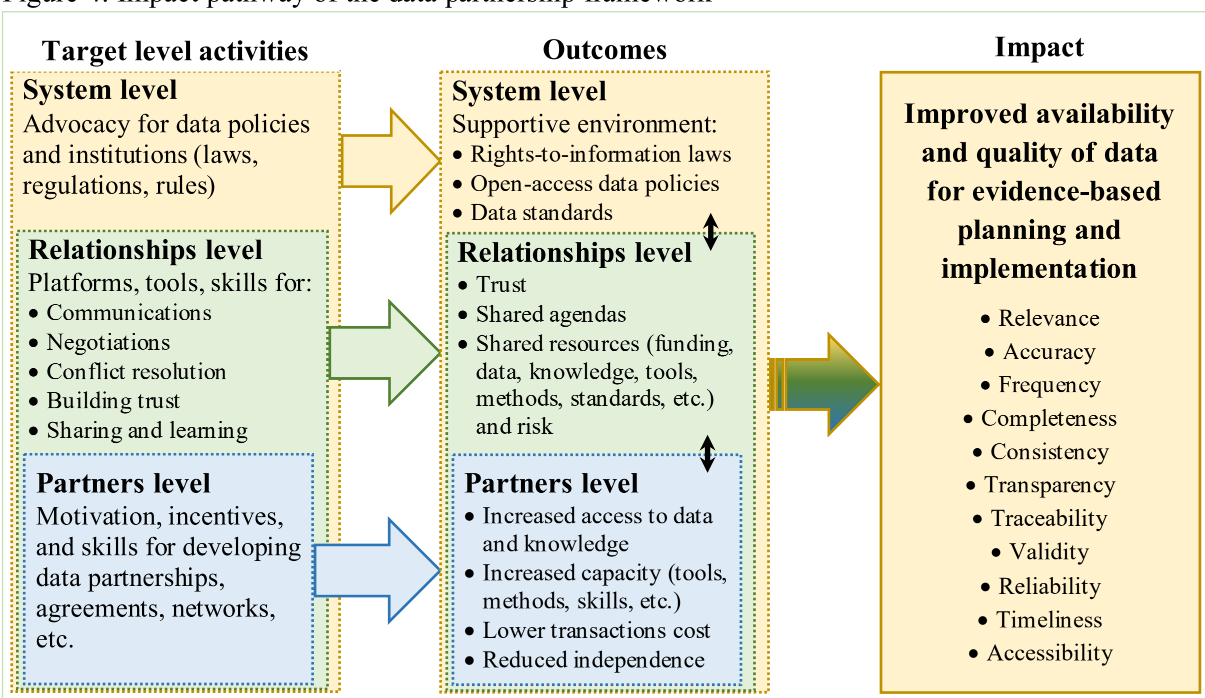
Timeliness is a key factor within the policy ecosystem. If the data are outdated or the results of strategic analysis are late in the mutual accountability process for example, it can make all the difference in terms of the parties making commitments or no commitments (or taking actions or inactions) that reflect the reality on the ground.

Pathways of impact

The pathways of impact illustrated in Figure 4 reflect an integrated and reinforcing set of capacity strengthening activities and outcomes at the partners, relationships, and system levels. The main pathway however derives from the relationships level and involves the partners having a shared agenda and, consequently, sharing resources and risk to generate and use high-quality data for evidence-based planning and implementation. Therefore, trust becomes important. Defined as a firm belief in the reliability, truth, ability, or strength of an entity, trust is expected to invoke the best from each partner toward achieving the shared agenda. Thus, with each partner contributing their best data sources, tools, and methodologies and learning from each other to generate and use data of mutual interest, the availability and quality of the data are expected to increase. In the context of CAADP and the BR, trust between state and non-state organizations and individuals is especially critical for reducing politicization of data generated by state agencies or of lack of use of good data generated by nonstate organizations (think tanks, universities, and researchers) in the policymaking process. Trust is also important for diffusing the pressure that governments may feel with the BR as an evaluation of their policies, which may also contribute to the politicization of data to begin with.

However, the different traits of data quality (relevance, accuracy, frequency, completeness, consistency, transparency, traceability, validity, reliability, timeliness, and access) may derive differently from the different levels or aspects of the partnership. For example, improvement in supportive policies at the system level alone may improve access to data only, other factors remaining unchanged. Similarly, strengthening the capacity of selected data-generating organizations alone may improve some of the traits—such as accuracy, frequency, completeness, consistency, transparency, traceability, and access—for the targeted data only. Issues of availability of other data as well as other traits—such as relevance, reliability, validity, and timeliness—for the targeted data may remain unchanged. Without trust and a shared agenda, the relevance, reliability, validity, and timeliness of the data may be undermined. With the demand and incentives for high-quality data and knowledge deriving from a mutual accountability process, the framework builds on the Strategic Analysis and Knowledge Support System (SAKSS) concept, which is defined as a network of people and institutions that provides timely, credible, and evidence-based knowledge and analysis to inform agricultural and rural development strategies in Africa (Johnson and Flaherty 2011, Johnson 2018).

Figure 4: Impact pathway of the data partnership framework



Source: Authors' illustration.

3. PILOT PROJECT DESCRIPTION AND IMPLMENTATION

Pilot project goal and objectives

The goal of the pilot project is to improve data systems and the quality of data available for policymaking in CAADP implementation and for achieving the Malabo Declaration targets, using the data partnership framework presented above. The overall objective is to improve the accuracy, consistency, traceability, and validation of the data used in the 2020 BR process in five selected countries (Kenya, Malawi, Mozambique, Senegal, and Togo). The specific objectives are:

1. To identify gaps and challenges in the country BR process related to data, methodologies, capacities, and systems;
2. To strengthen human and institutional capacities in the BR process related to data compilation, analysis, management, and reporting;
3. To strengthen capacities in information and communications technology (ICT) for the BR data management and sharing;
4. To improve the quality (accuracy, consistency, traceability, and validation) of the BR data;
5. To support countries to deliver a high quality 2020 BR report;
6. To conduct strategic analyses for achieving key Malabo targets and country development objectives; and
7. To develop a roadmap for the country to fill missing data for future BR reports.

Although the aim of the pilot project is to improve data quality (accuracy, consistency, traceability, and validation) in the 2020 BR process, improvement in data quality is expected to improve the statistical significance and reliability of estimated relationships between policies, investments, and outcomes. Then, policymakers and investors can be more confident in using results of strategic analysis to make policies and investments decisions that are more likely to yield desirable outcomes. This will in turn strengthen the links between policies, investments and the BR scores, so that policymakers can also be confident that selecting policies and investments that lead to desirable outcomes will also lead to higher BR scores. This is the true spirit of the overall BR exercise: that more accurate tracking, review and benchmarking of efforts and outcomes will ensure that African countries continue their progress towards the goals set in Malabo and the Agenda 2063 (AUC 2015).

Country selection

The main criteria for purposely selecting the five countries—Senegal, Kenya, Malawi, Mozambique, and Togo—are three: (1) having a SAKSS or SAKSS-like function; (2) willingness to participate in the pilot project and openly and proactively put all of its data challenges on the table and work with IFPRI-ReSAKSS to identify and address the gaps; and (3) that the effect of the pilot project on the 2020 BR score is irrelevant. With respect to the latter, it is important to clarify that the focus of the BR pilot exercise is on helping the countries to improve data quality for a more accurate BR outcome. And so, while high-quality data may change the BR score, the exercise should not be misunderstood as an effort to raise a country's

2020 BR score over what it was in 2018. This is because the values of the indicators that determine the BR score are a result of policies and investments made by countries, and how these policies and investments impact sectors, firms, households, and individuals.

With respect to the first criteria, Senegal is the most advanced because it has a SAKSS with most of the key elements of the data framework for evidence-based policymaking: a LAN, a data and knowledge management platform, an inclusive platform for policy dialogue and review, and a capacity strengthening component to improve capacities for data collection, analysis, and communications on agricultural policies. Senegal's SAKSS was set up in 2016 as part of the Feed the Future Senegal *Projet d'Appui Aux Politiques Agricoles* (PAPA) that is led by Senegal's Ministry of Agriculture and Rural Equipment in collaboration with Michigan State University and IFPRI. The LAN is made up of nine institutions including the National Agency of Statistics and Demography, Consortium for Economic and Social Research, Senegalese Institute for Agricultural Research/Bureau of Macroeconomic Analysis, Directorate for Analysis, Forecasting and Agricultural Statistics, School of Economics and Management at the University of Cheikh Anta Diop, and the University of Gaston Berger. Therefore, lessons from the Senegal PAPA project (FtF 2020)[2] were also used to design and implement the pilot project in all the countries.

Pilot project activities and implementation

To achieve the goal and objectives of the pilot project, capacity strengthening activities were designed and implemented along the different stages of the 2020 BR process in the countries, starting with a review of the inaugural (2018) process and reports to identify the gaps and challenges in data and capacities (see Figure 5). At each stage, training workshops and hands-on or hand-holding approaches were used. The training workshops were used to: first, go over the materials prepared or procedures put in place by the AU (e.g., the technical guidelines on the BR indicators, the data reporting template, entering data into the eBR, and the technical notes on the BR scoring methodology); and second, focus on addressing the country-specific gaps and challenges identified through the reviews and capacity needs assessment. The hands-on approach involved working closely with the BR team to support the day-to-day activities.

***Capacity needs assessment*:** This involved desk reviews, interviews with stakeholders, and consultation workshops.[3] Desk reviews of the 2018 BR report and potential data sources were used to assess the parameters with missing data as well as those that were entered inaccurately or inconsistently. Other documents (e.g., country BR briefs and lessons, past capacity needs assessments, and JSR assessments) were also reviewed to help identify general data capacity weaknesses. The interviews and consultation workshops were used to corroborate the findings from the desk reviews as well as to obtain the perceptions and experiences of the various CAADP stakeholder groups with the 2018 BR process, especially in the data collection and validation activities. The interviews and consultation workshops were also used to obtain their

---

[2] For details on the PAPA project, see http://www.papa.gouv.sn/categorie_publication/rapports-etudes/.

[3] Annex Table A1 (section on *assessment of the 2018 BR process and report to identify areas for improving the 2020 BR process and data systems*) provides details of the needs assessment in terms the review areas and key questions asked in conducting the reviews. Full project reports on the activities and outputs for each of the five countries are available on request.

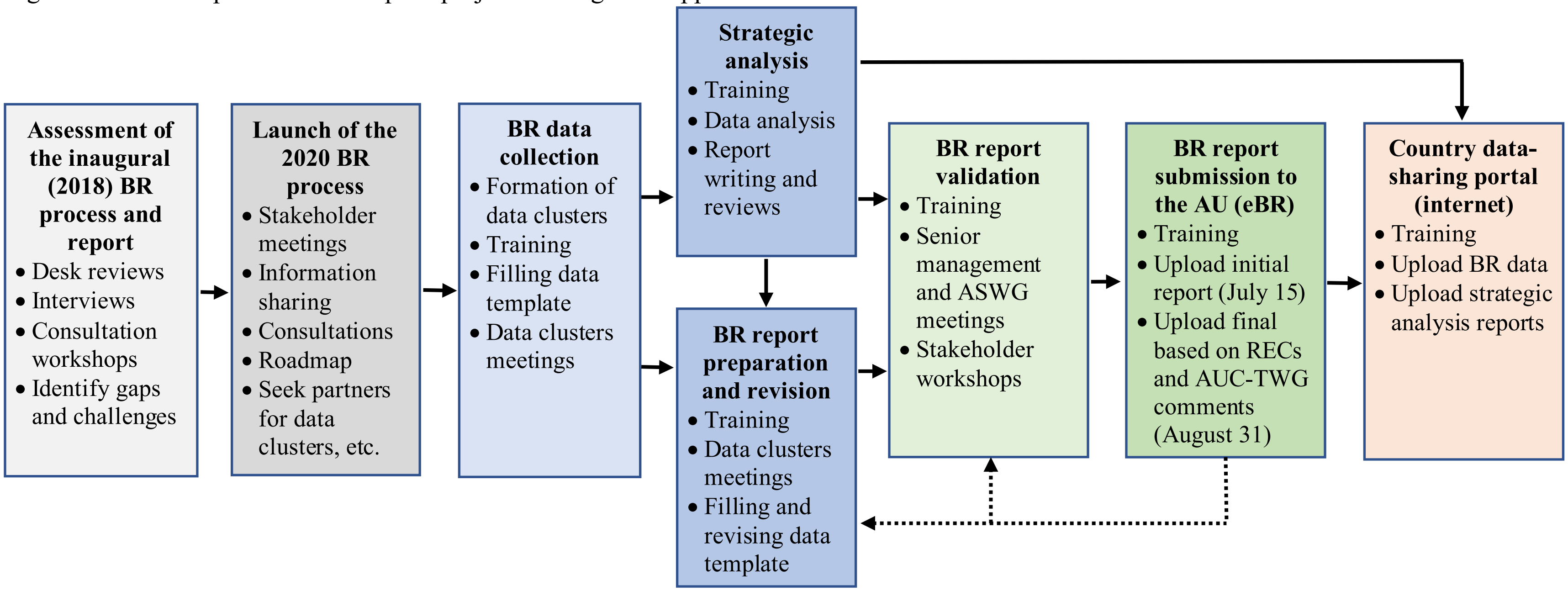
expectations with the 2020 process, including their interest and capacity to participate in it by membership in the data clusters.

Key questions that guided the reviews include:

- For indicators completed, which ones were done correctly or incorrectly?
- For indicators not completed, which ones could other data sources or methods have been used to complete them (either unknown at the time or known but not used)? If known but not used, why?
- For indicators not completed, which ones will the data gap remain?
- Are there any data standards and protocols for collection, management and sharing data? If yes, were they used during inaugural BR process? If no, why not?
- What are the characteristics of the data system used in the compilation process? For example:
  - Which institutions (e.g., research center, consulting company, individual consultants, etc.) are used to collect or compile data?
  - Which institutions give access to their data sets?
  - Which centers of expertise are involved in helping with computation of indicators, completing templates, etc.?
- What are the mechanisms used to assess the quality of data?
  - How does it work?
  - Which institutions or centers of excellence are involved?
- Was a JSR-like process used in the validation process?
  - If yes, evaluate against the AU guidelines in the JSR best practices.
  - If no, why not?

The reviews were also important for understanding the system, organizational and individual level capacities related to the BR process, ICT platforms for data management and sharing, and strategic analysis and mutual accountability mechanisms in support of policymaking processes. In general, the reviews showed that the data used in preparing the inaugural 2018 BR report were compiled from many sources (published and unpublished) that were known to the BR team. In addition to missing indicators, the reviews show that several of the indicators were incorrectly computed. Some of the missing data were due to low frequency of measurement such as for poverty and nutrition, which are measured every four-to-five years and annual data are not available. The most challenging data or indicators were those on gender/women/youth, private investment, resilience, and post-harvest loss. The results of the reviews were used to develop specific training activities to strengthen capacities of the BR country team in understanding the 2020 BR indicators and data requirements, making an inventory of the available data and data sources, compiling the relevant data, making appropriate transformations where needed, validating the data with the relevant CAADP stakeholders, submitting the data to the AU, and making the data and related analysis accessible to the public via the internet (see Figure 5). Therefore, the capacity strengthening activities focused at the partners and relationships level, especially in motivating, providing incentives, and strengthening skills for forming and working in data clusters at the partners level, and strengthening platforms, tools, skills for sharing and learning at the relationships level (see Figure 4).

Figure 5: 2020 BR process and the pilot project training and support activities



Source: Authors' illustration.

Notes: ASWG = agriculture sector working group. eBR = electronic or online BR system. RECs = Regional Economic Communities. AUC-TWG = African Union Commission-Technical Working Group.

***Data clusters*:** As illustrated in Figure 2, data clusters were formed according to the seven thematic areas of the Malabo Declaration, with each cluster being responsible for compiling and checking the data for the BR parameters and indicators under the relevant theme. They were constituted by representatives from various organizations that predominantly generate or use data and statistics. Table 1 shows the composition of the data clusters as well as the characteristics (gender, non-state actors, and sector) of the cluster members. In general, the clusters were dominated by male, state and agriculture-sector actors.

***Training activities*:** The core BR team and data clusters were trained on various aspects of data collection, strategic analysis, data and knowledge management and sharing, and mutual accountability. Specifically, the trainings included:

- The 2020 BR indicators, guidelines, data requirements, and reporting template;
- Methodologies for measuring resilience and women in agriculture empowerment;
- Standards and protocols for data compilation and management (e.g., curation, checking related sources, analyzing trends and correlations, etc.) for improving data consistency and accuracy;
- Standards and protocols for data management and for improving data traceability;
- M&E system that links to the Malabo goals;
- Strategic analysis for mutual accountability;
- Data validation according to the JSR guidelines and best practices
- The AU eBR system for uploading the data;
- Open access data; and
- ICT and country eAtlas for data management and sharing.

Because the trainings on and support to strategic analysis and ICT and country eAtlas for data management were more involving and costly, as well as not being necessary for the BR work in terms of improving the quality of the 2020 BR data, they were restricted to Mozambique and Senegal only.

***Data compilation and validation*:** Using the data collection template provided by the AU, the data clusters collected the available data and sent them to the core BR team. In the process, about two to three cluster-level meetings were held to check data before sending them to the core BR team. Various senior management and working group meetings in the agricultural sector were held to also check and validate the data, before the national-level validation workshop involving all the relevant stakeholders that took place between late June and early July (see Table 1 for the composition of participants at these workshops). The validated data were then uploaded into the eBR system by July 15, 2019. As the effectiveness of the national validation workshops depends on how well the participants know or understand the data, sharing the data with them prior to the workshop would have been ideal. The short timeframe did not allow this. The disadvantages for not doing this likely reduce however with the data cluster approach and having several internal validation meetings.

Table 1: Composition of the core BR team, data clusters, and participants at stakeholder workshops in the five pilot countries

	Kenya				Malawi				Mozambique				Senegal				Togo			
	Total	Female	NSA	NAG	Total	Female	NSA	NAG	Total	Female	NSA	NAG	Total	Female	NSA	NAG	Total	Female	NSA	NAG
Core BR team	12	4	1	2	7	0	1	0	3	1	0	0	2	0	0	0	2	0	0	0
Data clusters:																				
BR theme 1	5	1	1	1	16	3	2	0	3	1	0	0	7	0	3	0	4	0	2	0
BR theme 2	8	2	2	2	4	3	0	1	8	6	0	3	6	0	0	2	6	0	0	2
BR theme 3	9	2	1	2	3	2	0	1	12	4	0	3	14	0	0	3	10	0	0	2
BR theme 4	8	2	2	2	3	0	2	0	12	6	1	2								
BR theme 5	7	2	0	3	2	1	0	0	8	2	0	3	6	0	0	0	7	0	0	5
BR theme 6	5	1	0	3	2	0	0	0	7	3	1	3	6	0	0	0				
BR theme 7	9	2	2	5																
Total	51	12	8	18	30	9	4	2	50	22	2	16	22	0	3	3	25	0	2	8
Workshops:																				
Launch	70	12	12	10	27	6	7	0	50	18	4	12	27	6	3	3	24	2	6	11
Training	33	9	12	6	23	6	2	0	50	22	2	16	27	6	3	3	40	5	8	11
Validation	58	23	24	7	40	7	12	2	60	15	15	16	52	10	6	3	116	18	22	36

Source: Authors' compilation from pilot project

Notes: Notes: NSA = non-state actor. NAG = non-agriculture sector. n.a. = not applicable. In general, the figures refer to representatives of institutions that contributed to more than one cluster. For Malawi, the cluster for theme 1 was also responsible for theme 7. For Mozambique, the core BR team was also responsible for themes 1 and 7. For Senegal, the cluster for theme 1 was also responsible for theme 7 and theme 3 for theme 4. For Togo, the cluster for theme 1 was also responsible for theme 7, theme 3 or theme 4, and theme 5 for theme 6.

Although effort was made to get all the relevant CAADP stakeholder groups to participate in the national validation workshops, participation of NSAs was limited. It was recommended at some of workshops that the NSAs need to better organize themselves and articulate their role in the CAADP processes (e.g. contributing to the data collection or initiating policy dialogues). Also, whereas the involvement of parliamentarians in transformation of agriculture is key, their engagement in BR process remains a challenge. In general, however, there was an overall improvement in the data reported as several of the data gaps and missing indicators identified in the inaugural BR were filled (more on this later). Some inaccuracies or inconsistencies may remain, depending on the timeframe available to the country team to conduct or internalize any post-validation analysis and comments. Leadership of the cluster was found to be critical. Therefore, a key question is how to sustain the clusters to continuously update and manage the BR data? This is taken up later in the last section of the paper on implications.

Strategic analysis and data management and sharing**:** These activities were limited to Mozambique and Senegal. In Mozambique, strategic analysis was conducted on five areas: outcome and lessons of the BR, access to financial services in the agriculture sector, fertilizer value chain, seed value chain, and maize production and productivity.[4] The brief on the BR analyzes Mozambique's performance, discuss best practices and experiences from the implementation of the 2020 BR process, and provides recommendations for strengthening mutual accountability and performance of the agriculture sector in Mozambique. The analysis of access to financial services helps to address the data needs on the indicator in the BR process. The studies on the fertilizer and seed value chains aim at building awareness among policymakers in Mozambique on the key challenges in these value chains. The maize study analyzes the trends in production as well as the challenges and opportunities for improving productivity in the maize subsector.

In Senegal, the analysis included an evaluation of the progress made in reaching the Malabo targets by 2025 by: (i) identifying the performance categories (PCs) of indicators with the lowest rates of achievement; and (ii) assessing the extent to which the country has caught up with or is moving away from the targets for the different PCs, comparing the achievements in 2018 versus 2020. A brief on the first analysis shows that the PCs on food safety, social protection, intra-African trade of agricultural goods and services were the ones with the lowest rates of achievement. For the second analysis for which a brief is underway, the country moved from being "on track" to "not on track" for the PCs on intra-African trade, the CAADP process, and mutual accountability. On the other hand, the country caught up with respect to access in agricultural inputs and technologies and agricultural productivity. The briefs also include recommendations on documenting progress at the PC level in the next rounds of the BR.[5]

---

[4] For Mozambique, the 2018 and 2020 BR reports and the five studies are available for download at the government data portal (http://www.portaldogoverno.gov.mz/por/Cidadao/Agricultura).

[5] For Senegal, the 2020 BR report is available for download at the website of the Directorate for Analysis, Forecasting and Agricultural Statistics (http://www.dapsa.gouv.sn/content/rapport-rb-sénégal-2019-version-finale or Cliquez ici pour télécharger le fichier PDF.).

4. METHODS AND DATA FOR ASSESSING IMPROVEMENTS

We use two indicators of outcome to assess the impact of pilot activities—the data reporting rate (DRR) and the quality of data reported (QDR). Here, DRR is measured by the data parameters reported as percent of total data parameters required, and QDR is measured by the data parameters that have issues (inaccurate, illogical, inconsistent, etc.) as percent of the data parameters reported. Therefore, QDR does not double count missing data that are already factored into DRR. Regarding QDR, some of the main issues identified include: wrong units of measurement (thousand, million, billion, etc.); values that are too high or low due to extra or missing digit(s); spaces in between numbers which makes the value a string or text (e.g., 1 234 567 instead of 1234567); use of decimal point (e.g., 1.234.567) instead of commas (e.g., 1,234,567) which makes the value too low (e.g., 1.234567); illogical responses (e.g., where the response to a subsequent question is supposed to be conditioned by the response to a preceding question, or having a value where the sum of parts is greater or less than the aggregate value); and other inaccuracies, typos, etc.

Then, the improvements made can be assessed by analyzing the change over time (i.e., between 2018 and 2020) in DRR and QDR in the pilot countries. However, since the goal of the pilot project is to help improve data for CAADP in general by extending the results to other African countries in the BR process, it is important to compare the improvements made in DRR and QDR in the pilot countries to changes in the same outcome indicators in the non-pilot countries that are involved in the BR process. Therefore, it is important to have non-pilot countries that are like the pilot countries prior to the project or at the baseline.

Measuring the relative improvement in DRR and QDR

Using the 2018 BR results as the baseline, the relative improvement in DRR in the pilot countries is measured by a difference-in-difference (DID) approach, which is the change in DRR between 2018 and 2020 for the pilot countries relative to those for comparable non-pilot (or like-pilot) countries. Then, the DID in DRR between the pilot and like-pilot countries (denoted by DRR_{DID}) is given by:

$$DRR_{DID} = (DRR_{2020} - DRR_{2018})^{pilots} - (DRR_{2020} - DRR_{2018})^{like\text{-}pilots} \quad ...(1)$$

For the relative improvement in QDR, the measure of QDR is not available in the 2018 BR because the data are not in a format that allows a comparative analysis of the data issues to be done. Thus, we only compare QDR in 2020 for the pilot countries with QDR for the like-pilot countries. Thus, the difference in the QDR between the pilot and like-pilot countries (denoted by QDR_D) is given by:

$$QDR_D = (QDR_{2020})^{pilots} - (QDR_{2020})^{like\text{-}pilots} \quad ...(2)$$

Identifying the like-pilot countries and estimation of DRR_{DID} and QDR_D

Different approaches are used to identify the like-pilot countries and to estimate equations 1 and 2. In general, the like-pilot countries are identified as those with similar characteristics of the three criteria used for selecting the pilot countries—having a SAKSS or SAKSS-like function, willingness to participate in the pilot, and being passive about the 2020 BR score. As we have no data on first criterion for all the non-pilot countries, and the other two criteria are unobservable, we use variables that capture the three criteria and are likely to affect the outcomes, i.e. DRR and QDR. Agricultural Statistics Capacity Index (ASCI) is used to capture the SAKSS-function, as the SAKSS is expected to raise the capacity to generate and use data and statistics for evidence-based policymaking (Johnson and Flaherty 2011, Johnson 2018). For willingness to participate in the pilot, the value of DRR in 2018 seems like a good indicator, where countries with low or moderate values and seeking to improve their performance would be more willing to participate than those that have high values. Also, those that already have high DRR values in 2018 would be less likely to be selected to participate in the pilot, as the room for improving their DRR in 2020 may be small. Although the BR score (BRS) is not a focus of the pilot study, it is also included in the analysis as its value in 2018 is expected to influence either the willingness to participate in the pilot or the likelihood of being selected to participate. Countries with low BRS values in 2018 may be more willing to participate to improve their score in 2020 via improving their DRR, as low DRR (especially due to missing data) was a major reason for countries that low scores in the 2018 BR (Benin et al. 2018). Those that have high BRS values in 2018 may be less willing to participate in the pilot to the extent an improvement in QDR in 2020 may lead to a reduction in BRS in 2020. Similarly, those with high BRS in 2018 may be less likely to be selected to participate in the pilot to the extent that pressure is exerted on the pilot to yield higher BRS in 2020. With these three variables and their measure in 2018 ($ASCI_{2018}$, DRR_{2018}, and BRS_{2018}), we consider three methods for identifying the like-pilot countries and for estimating DRR_{DID} and QDR_D: (1) standard deviation method; (2) propensity score matching method; and (3) two-stage weighted regression method. Because of potential high positive correlation among the three variables, we consider different combinations of them in the analysis.

Standard deviation method (SDM): In this method, the like-pilot countries here are defined as those within 1, 2, or 3 standard deviations of the average values of the variables for the pilot countries. First, only the main variable (DRR_{2018}) is used in the formation of the like-pilot (LP) country groups, denoted by SDM1-LP1, SDM1-LP2, and SDM1-LP3 for countries within 1, 2, and 3 standard deviations, respectively, of the average value of DRR_{2018} for the pilot countries. Then, all the three variables (DRR_{2018}, $ASCI_{2018}$, and BRS_{2018}) together are used, which again results in three LP country groups denoted by SDM3-LP1, SDM3-LP2, and SDM3-LP3. In general, increasing the number of standard deviations may increase the number of countries in the group, which may reduce the likeness of the group to the pilot countries. Contrary, increasing the number of variables may reduce the number of countries in the group, which may reduce the number of countries in the group (indicating difficulty of finding comparable countries) and affect the estimator when there are too few countries in a group. Then, DRR_{DID} and QDR_D using

SDM are obtained from the difference between the average for the pilot and like-pilot groups according to:

$$DRR_{DID\text{-}SDM} = \Sigma_j \Delta DRR_{1j}/n_1 - \Sigma_j \Delta DRR_{0j}/n_0 \quad \ldots(3)$$

$$QDR_{D\text{-}SDM} = \Sigma_j QDR_{2020,1j}/n_1 - \Sigma_j QDR_{2020,0j}/n_0 \quad \ldots(4)$$

where ΔDRR is the change in DRR between 2018 and 2020; subscripts 1 and 0 represent pilot and like-pilot countries, respectively; and n_1 and n_0 are the number of pilot and like-pilot countries, respectively.

Propensity score matching method (PSMM): Here, each country in the pilot is first matched with non-pilot countries that are as similar as possible in terms of the variables that determine selection or affects the outcomes (represented by the vector $\boldsymbol{x}$) using a propensity score, which is the estimated conditional probability (Pr) of participation in the pilot according to:

$$\Pr(\text{pilot}_j = 1 \mid \boldsymbol{x}_j) = \Phi(\boldsymbol{x}_j'\gamma) \quad \ldots(5)$$

where pilot = 1 is for being in the pilot, and 0 otherwise; and γ is the parameter to be estimated. Then, the matched countries (using subscript Mj as the match for each pilot country j) are used to obtain DRR_{DID} and QDR_D according to:

$$DRR_{DID\text{-}PSMM} = \Sigma_j \psi_j (\Delta DRR_{1j} - \Delta DRR_{Mj}) \quad \ldots(6)$$

$$QDR_{D\text{-}PSMM} = \Sigma_j \psi_j (QDR_{2020,1j} - QDR_{2020,Mj}) \quad \ldots(7)$$

where ψ_j is a weight based on the propensity score (Pr).[6] To improve matching, it is common practice to try different variables and different transformations of the variables, including logarithms, higher-order terms, and interaction factors. In addition to following these practices, we use different matching techniques such as nearest neighbor (considering different number neighbors) and kernel density matching. Various balancing tests (Rubin 2001, Dehejia and Wahba 2002) are used to select the best matched sample of treatment and comparison observations. The balancing tests are a check of whether and the extent to which any differences that existed between the two groups prior to the matching have been reduced or eliminated in the matched sample. The best matching result was nearest neighbor matching with 5 or 6 nearest neighbors (more on this later).

Selection bias arising from differences in the willingness or ability to participate in the project is reduced to the extent that the three variables in $\boldsymbol{x}$ and their transformations captures the differences in this dimension as hypothesized above. Given that the entire population of countries available for the analysis is only 45 (5 treated and 40 non-treated), it is difficult to compare the two groups across many pretreatment characteristics.

[6] See Imbens and Wooldridge (2009), Dehejia and Wahba (2002), and Becker and Ichino (2002) for details.

Two-stage weighted regression method (2SWRM): This method combines matching and regression in a two-stage estimation procedure. Here, the predicted probabilities (Pr) or propensity scores from the PSMM are used as weights w_j in a regression of the outcomes on $\boldsymbol{x}$ according to:

$$w_j\Delta\text{DRR} = \alpha_{\text{DRR}} * w_j + \delta_{\text{DRR}} * w_j\text{pilot} + w_j\boldsymbol{x}'\beta_{\text{DRR}} + w_j\varepsilon_{\text{DRR}} \quad \ldots(8)$$

$$w_j\text{QDR}_{2020} = \alpha_{\text{QDR}} * w_j + \delta_{\text{QDR}} * w_j\text{pilot} + w_j\boldsymbol{x}'\beta_{\text{QDR}} + w_j\varepsilon_{\text{QDR}} \quad \ldots(9)$$

where ε is the error term; α, δ, and β are the parameters to be estimated; and δ_{DRR} and δ_{QDR} are estimators of $\text{DRR}_{\text{DID-2SWRM}}$ and $\text{QDR}_{\text{D-2SWRM}}$, respectively. The sample is restricted to the pilot and the matches. The weighting can be interpreted as removing bias due to any correlation between $\boldsymbol{x}$ and selection to be in the pilot, while the regression isolates the effect of $\boldsymbol{x}$ over time.[7] We expect $\text{DRR}_{\text{DID-PSMM}} > \text{DRR}_{\text{DID-2SWRM}}$ and $\text{QDR}_{\text{D-PSMM}} > \text{QDR}_{\text{D-2SWRM}}$ to the extent that $\boldsymbol{x}$ has the same sign and statistically significant effect on the selection (equation 5) and on the outcomes. The estimations using the different methods were conducted in STATA (StataCorp 2019).

Sources of data

The data on all the variables are from the AU eBR system that were reported by the countries in the 2018 and 2020 BR processes (AUC 2018, 2020). For DRR and QDR, using these data sources helps to eliminate possible self-evaluation bias by the implementers of pilot on the improvements made in the pilot countries. For ASCI, the value used to represent the baseline (or ASCI_{2018}) is an average of the values from 2014 to 2016 that are obtained from the 2018 BR. For DRR and BRS, the baseline measures (DRR_{2018} and BRS_{2018}) are those reported in the 2018 BR.

Regarding the quality of data, a detailed analysis of the data submitted by the countries was conducted by the BR technical working group (TWG) at the AU BR writeshop that took place on September 9-13, 2019 in Lusaka, Zambia. The data issues presented earlier were uncovered at this meeting. As a result, there was some data cleaning by the BR TWG following the event to improve the quality. Also, countries were notified to make further corrections before the 2020 BR report was finalized. To avoid contamination of the data quality outcome (or QDR_{2020}) however, the data used in this paper are those prior to the above changes. This is also because the analysis by the TWG to uncover the data issues was informed by the analysis conducted for this pilot study. Therefore, the results presented in this paper on the DRR_{2020} and QDR_{2020}, and consequently change from 2018 as well as differences between the pilot and non-pilot countries, may be different from a similar analysis conducted with the final dataset released by the AU to the extent that the data issues that were uncovered have been corrected.[8]

---

[7] For more on the 2SWRM, see Robins and Rotnitzky (1995), Robins, Rotnitzky, and Zhao (1995), Imbens and Wooldridge (2009), and Benin et al. (2011).
[8] The 2020 BR report was released at the 33rd Ordinary Session of the AU Summit in Addis Ababa, Ethiopia, 9-10 February 2020.

The total number of data parameters required increased by 60 percent from 166 in the 2018 BR to 266 in the 2020 BR (Table 2). The bulk of the increase derives from theme 3, where new indicators on food safety were included in the 2020 BR. This added 3 new indicators and 29 new data parameters. In general, the change derives from the level of the data parameters required in the 2018 BR versus the 2020 BR. In the 2018 BR, the data were at a higher level and countries could do some aggregations before reporting the data. With the introduction of the eBR system in the 2020 process, the data were at a lower level and countries did not have to perform any aggregations before reporting the data. The data parameters required for themes 1 and 7 declined due to simplification of the data needed to compute the respective indicators.

Table 2: Number of indicators and data parameters required in the biennial review (BR)

Theme	2018 BR		2020 BR		% change	
	Indicators	Parameters	Indicators	Parameters	Indicators	Parameters
Total	43	166	46	266	7	60
Theme 1	3	28	3	27	0	-4
Theme 2	6	20	6	28	0	40
Theme 3	18	63	21	153	17	143
Theme 4	8	21	8	29	0	38
Theme 5	3	16	3	16	0	0
Theme 6	3	7	3	8	0	14
Theme 7	2	11	2	5	0	-55

Source: Authors' representation based on the BR data (AUC 2018, 2020).

Notes: In the 2020 BR, 3 new indicators and 29 new indicators on food safety were added under theme 3.

Countries in the sample

The total number of countries in the data for the analysis are 5 pilots and 40 non-pilots, which are the countries that submitted a report and data in both the 2018 and 2020 BRs.[9] The distribution of the non-pilot countries according to the different methods for identifying the like-pilot countries are shown in Table 3. Regarding the standard deviation method based on DRR_{2018} only (SDM1) for example, there are 9, 18, and 24 like-pilot countries within 1 (SDM1-LP1), 2 (SDM1-LP2), and 3 (SDM1-LP3) standard deviations of the mean DRR_{2018} for the pilot countries, respectively. For SDM3, which is based on the three variables—DRR_{2018}, $ASCI_{2018}$, and BRS_{2018}), there are 1, 10, and 19 like-pilot countries within 1 (SDM3-LP1), 2 (SDM3-LP2), and 3 (SDM3-LP3) standard deviations of the mean of each of the variables, respectively. Because there is only one observation for (SDM3-LP1), it is excluded from further analysis. For PSMM and 2SWRM, the best matching results were obtained from using nearest neighbor matching with 5 or 6 nearest neighbors (more on this later when the balancing results are presented). There are 13 like-pilot or matched countries. As the list of countries in Table 3 shows, all of those identified by PSMM are also identified by SDM1 or SDM3. The main

[9] In the 2018 BR, eight countries (Algeria, Comoros, Eritrea, Guinea-Bissau, Libya, Sahrawi, Somalia, and South Sudan) did not submit a report. In the 2020 BR, six countries (Algeria, Comoros, Libya, Sahrawi, Egypt, and Sao Tome and Principe) did not submit a report.

differences are with respect to Burundi, Djibouti, Gabon, Seychelles, and Zimbabwe, which are identified by SDM1 only.

Table 3: List of the non-pilot countries in the sample, by method of identifying the like-pilot countries

Method	Like-pilot (LP) countries		
SDM1	Within +/- 1 s.d. (SDM1-LP1)	Within +/- 2 s.d. (SDM1-LP2)	Within +/- 3 s.d. (SDM1-LP3)
	Botswana, Cote d'Ivoire, Eswatini, Gambia, Lesotho, Madagascar, Mauritius, Namibia, and Tanzania (9)	SDM1-LP1 plus Benin, Burkina Faso, Cabo Verde, Djibouti, Ethiopia, Morocco, South Africa, Uganda, and Zimbabwe (18)	SDM1-LP2 plus Burundi, Gabon, Ghana, Mali, Seychelles, and Zambia (24)
SDM3	Within +/- 1 s.d. (SDM3-LP1)	Within +/- 2 s.d. (SDM3-LP2)	Within +/- 3 s.d. (SDM3-LP3)
	Namibia (1)	SDM3-LP1 plus Benin, Botswana, Burkina Faso, Cabo Verde, Eswatini, Lesotho, Mauritius, South Africa, and Uganda (10)	SDM3-LP2 plus Côte d'Ivoire, Ethiopia, Ghana, Madagascar, Mali, Morocco, Seychelles, Tanzania, and Zambia (19)
PSMM	5 or 6 nearest neighbor matches		
	Botswana, Burkina Faso, Cote d'Ivoire, Eswatini, Ethiopia, Gambia, Lesotho, Madagascar, Mauritius, Morocco, Namibia, Tanzania, and Uganda (13)		

Source: Authors' representation based on the BR data (AUC 2018, 2020).

Notes: SDM1 = standard deviation method based on the data reporting rate (DRR_{2018}), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean DRR_{2018} for the pilot countries, respectively. SDM3 = standard deviation method based on the three explanatory variables (DRR_{2018}, BRS_{2018}, and $ASCI_{2018}$), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean of each of the three variables for the pilot countries, respectively. BRS is biennial review score and ASCI is agricultural statistics capacity index. PSMM = propensity score matching method using nearest neighbor matching. Number in parenthesis is the total number of countries in the group.

5. RESULTS ON THE IMPROVEMENTS MADE IN DRR AND QDR

Relative improvements in DRR and QDR: standard deviation method (SDM1 and SDM3)

First, Table 4 shows how the pilot and non-pilot countries compare at the baseline in the three variables—DRR_{2018}, $ASCI_{2018}$, and BRS_{2018}. Compared to all the non-pilot countries, the average values for the pilot countries are significantly higher. Compared to the like-pilot countries identified by SDM1 and SDM3 however, the differences are no longer statistically significant, except for BRS in the case of SDM1-LP1. However, this result is not surprising since SDM1 was based on DRR_{2018} only. But it does indicate the estimates of DRR_{DID} and QDR_{D} using SDM1-LP1 may be biased to the extent that BRS_{2018} is an important factor in determining selection or affecting the outcomes. Because SDM3-LP1 has only one observation (see Table 3), it is excluded from further analysis.

Table 4: Characteristics of the pilot and non-pilot countries at the baseline, 2018

Country/group/method	DRR	BRS	ASCI
Pilot countries:			
Kenya	88.0	4.77	65.0
Malawi	86.1	4.92	58.5
Mozambique	81.3	4.13	65.6
Senegal	78.9	3.84	58.4
Togo	80.7	4.92	53.1
Average (pilot countries)	83.0	4.52	60.1
Like-pilot (LP) countries:			
SDM1			
Within +/- 1 s.d. (LP1)	81.7	3.78 **	53.9
Within +/- 2 s.d. (LP2)	82.4	4.06	56.4
Within +/- 3 s.d. (LP3)	83.2	4.07	56.4
SDM3			
Within +/- 1 s.d. (LP1)	n.a.	n.a.	n.a.
Within +/- 2 s.d. (LP2)	82.7	4.29	59.8
Within +/- 3 s.d. (LP3)	84.2	4.24	60.0
All non-pilot countries	73.9 ***	3.54 ***	54.4 *

Source: Authors' based on model results.

Notes: DRR = data reporting rate (%). BRS = biennial review score (0 to 10). ASCI = agricultural statistics capacity index (0 to 100). SDM1 = standard deviation method based on the data reporting rate (DRR_{2018}), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean DRR_{2018} for the pilot countries, respectively. SDM3 = standard deviation method based on the three explanatory variables (DRR_{2018}, BRS_{2018}, and $ASCI_{2018}$), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean of each of the three variables for the pilot countries, respectively. *, **, and *** mean the difference between the average for the pilot countries and the reference like-pilot or non-pilot group of countries is statistically significant at the 10%, 5%, and 1% levels, respectively. n.a. = not estimated because of limited number of observations.

The results from estimation of equations 3 and 4 or by SDM1 and SDM3 are shown in Table 5 for DRR_{DID} and QDR_{D} for all the seven themes together.[10] Looking at the five pilot countries alone, Table 5 shows that the improvements in DRR between 2018 and 2020 (or ΔDRR) are

[10] Details on DRR and QDR for each country in 2018 and 2020 are presented in the annex Tables A2 and A3.

largest in Togo (increase of 12.2 percentage points [%pts]) and Senegal (10.9 %pts), followed by Kenya (3.8 %pts) and Malawi (3.0 %pts). The change in Mozambique is very small (0.6 %pts). With respect to QDR in 2020 (or QDR_{2020}), Senegal has the least issues with the data reported (2.3% of the data reported in 2020), followed by Kenya and Malawi (2.8% each), and then Togo (6.1%). Mozambique has the most data problems, with 12% of the data reported having some quality issues. Comparing the performance of pilot countries to all the non-pilot countries, the results in Table 6 show that DRR_{DID} is only 1.0 %pts on average, which is not statistically significant, and DQR_{D} is -4.4 %pts on average, with low statistical significance. Compared with the like-pilot countries using the two standard deviation methods however, the estimated DRR_{DID} is 7.6 to 8.9 %pts on average, which is statistically significant. For QDR_{D}, the estimate is -3.5 to -2 %pts on average, although not statistically significant.

Table 5: Differences in DRR and QDR between the pilot and non-pilot countries (2018 and 2020)

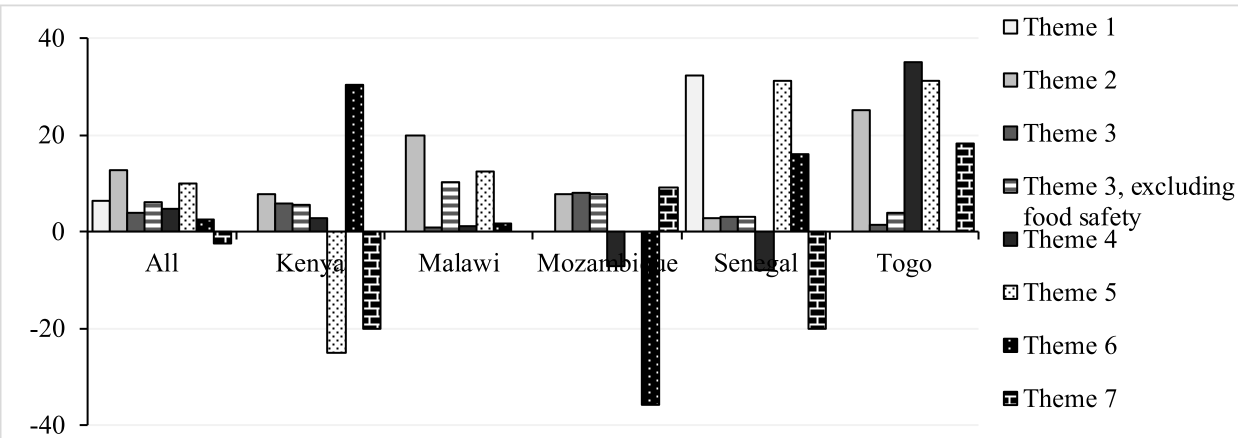
Country or group	DRR						QDR			
	2018	2020	ΔDRR	DRR_{DID}			2020	QDR_{D}		
			%pts.	Est.	p-value	Sig.		Est.	p-value	Sig.
Pilot countries:										
Kenya	88.0	91.7	3.8	n.a.			2.8	n.a.		
Malawi	86.1	89.1	3.0	n.a.			2.8	n.a.		
Mozambique	81.3	81.9	0.6	n.a.			12.0	n.a.		
Senegal	78.9	89.8	10.9	n.a.			2.3	n.a.		
Togo	80.7	92.3	12.2	n.a.			6.1	n.a.		
Average (pilot countries)	83.0	89.1	6.1	n.a.			5.2	n.a.		
Like-pilot (LP) countries:										
SDM1										
Within +/- 1 s.d. (LP1)	81.7	78.9	-2.8	8.9	0.028	**	7.9	-2.7	0.255	
Within +/- 2 s.d. (LP2)	82.4	80.0	-2.5	8.5	0.028	**	8.7	-3.5	0.128	
Within +/- 3 s.d. (LP3)	83.2	81.5	-1.7	7.8	0.019	**	7.8	-2.6	0.219	
SDM3										
Within +/- 2 s.d. (LP2)	82.7	82.1	-0.6	6.7	0.108		8.8	-3.6	0.164	
Within +/- 3 s.d. (LP3)	84.2	83.5	-0.7	6.8	0.030	**	8.1	-2.0	0.177	
All non-pilot countries	73.9	79.0	5.0	1.0	0.769		9.6	-4.4	0.051	*

Source: Authors' based on model results.

Notes: ΔDRR = change in data reporting rate (DRR, %) between 2018 and 2020. QDR = quality of data reported (%). SDM1 = standard deviation method based on the data reporting rate (DRR_{2018}), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean DRR_{2018} for the pilot countries, respectively. SDM3 = standard deviation method based on the three explanatory variables (DRR_{2018}, BRS_{2018}, and $ASCI_{2018}$), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean of each of the three variables for the pilot countries, respectively. BRS is biennial review score and ASCI is agricultural statistics capacity index. DRR_{DID} and QDR_{D} measure the difference between the average for the pilot countries and the reference like-pilot or non-pilot group of countries. *, **, and *** mean DRR_{DID} or QDR_{D} is statistically significant at the 10%, 5%, and 1% levels, respectively. n.a.= not applicable.

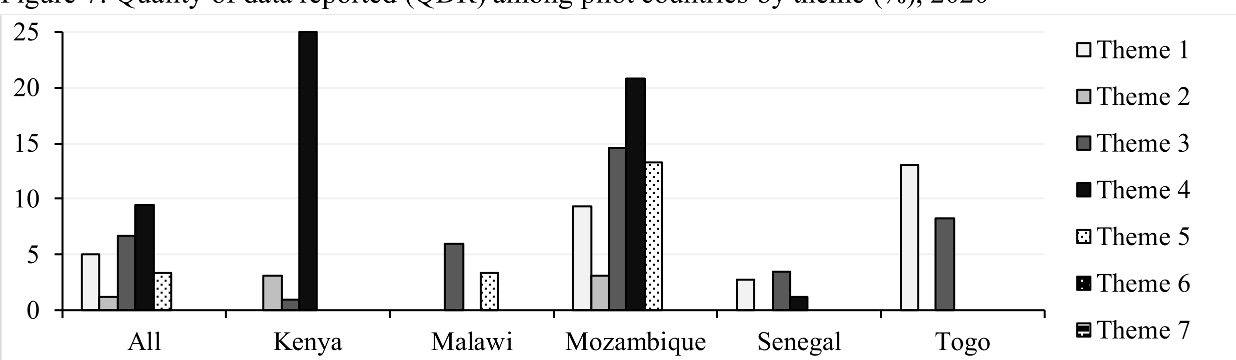
The results by theme are shown in Figures 6 and 7 for differences among the pilot countries alone, and then Tables 6 and 7 for the DRR_{DID} and QDR_{D} estimates, respectively. The results show that they are different for the seven thematic areas of the BR. Looking at the five pilot countries alone and with respect to DRR, the main differences across the countries are in themes 4 through 7, where some countries experienced a decline, compared to themes 1 to 3, where all countries experienced an increase or no change (Figure 6). In Kenya for example, the largest increase is in theme 6, but it also experienced large declines in themes 5 and 7. In Senegal, the largest increases are in themes 1, 5, and 6, and moderate to large declines in themes 4 and 7. The changes in Mozambique are moderate, except in theme 6, where the decline is large. Togo and Malawi are the only countries that did not experience any decline in thematic DRR. For the pilot countries with respect to QDR in 2018, themes 3 and 4 are the common areas with issues in most of the countries (Figure 7). Whereas each country experienced issues in two or three themes at most, Mozambique experienced issues in themes 1 through 5, with a QDR of at least 9% in four of them.

Figure 6: Change in DRR among pilot countries by theme (%pts), 2018 to 2020



Source: Authors' based on model results.
Notes: DRR = data reporting rate.

Figure 7: Quality of data reported (QDR) among pilot countries by theme (%), 2020



Source: Authors' based on model results.
Notes: QDR is measured as percent of data reported that have issues.

Comparing the average changes in DRR by theme for the pilot countries with that for the different like-pilot groups, the statistically significant differences are in theme 3 on food security and nutrition, either with or without the new food safety indicators. The estimated DRR_{DID} in theme 3 is 5.5 to 8.3 %pts on average (Table 6). With respect to QDR_D, the more statistically significant estimates are mostly in theme 2, -8.6 to -7.2 %pts for SDM1-LP2 and SDM1-LP3 (Table 7).

Table 6: Difference-in-difference in DRR between pilot and non-pilot countries, by theme

	Theme 1	Theme 2	Theme 3	Theme 3, excl. food safety	Theme 4	Theme 5	Theme 6	Theme 7
Pilot countries:								
ΔDRR	6.4	12.7	3.8	6.1	4.7	10.0	2.5	-2.5
Like-pilot (LP) countries:								
SDM1								
Within +/- 1 s.d. (LP1)								
ΔDRR	6.3	1.0	-4.5	-0.5	-8.6	-1.4	2.6	3.4
DRR_{DID}	0.1	11.7	8.3	6.6	13.3	11.4	-0.1	-5.9
P-value	0.989	0.043	0.011	0.065	0.253	0.410	0.995	0.559
Significance		**	**	*				
Within +/- 2 s.d. (LP2)								
ΔDRR	2.9	4.9	-3.9	-0.7	-5.4	1.0	-4.6	1.9
DRR_{DID}	3.5	7.8	7.7	6.8	10.1	9.0	7.1	-4.4
P-value	0.620	0.213	0.022	0.057	0.327	0.462	0.588	0.624
Significance			**	*				
Within +/- 3 s.d. (LP3)								
ΔDRR	2.9	3.4	-4	-1.6	-1.2	5.2	-2.9	0.4
DRR_{DID}	3.5	9.3	7.8	7.7	5.9	4.8	5.4	-2.9
P-value	0.606	0.106	0.004	0.009	0.536	0.684	0.659	0.734
Significance			***	***				
SDM3								
Within +/- 2 s.d. (LP2)								
ΔDRR	-0.7	8.9	-0.9	2.9	0.3	-3.1	4.5	-3.4
DRR_{DID}	7.1	3.8	4.7	3.2	4.4	13.1	-2.0	0.9
P-value	0.287	0.605	0.239	0.421	0.699	0.332	0.873	0.922
Significance								
Within +/- 3 s.d. (LP3)								
ΔDRR	3.7	6.0	-2.2	0.6	-1.5	2.0	1.1	-2.6
DRR_{DID}	2.7	6.7	6.0	5.5	6.2	8.0	1.4	0.1
P-value	0.696	0.223	0.020	0.043	0.512	0.501	0.906	0.996
Significance			**	**				
All non-pilot countries								
ΔDRR	7.1	13.0	2.7	5.5	2.7	19.4	4.1	7.0
DRR_{DID}	-0.7	-0.3	1.1	0.6	2.0	-9.4	-1.6	-9.5
P-value	0.927	0.966	0.714	0.849	0.82	0.43	0.895	0.259
Significance								

Source: Authors' based on model results.

Notes: ΔDRR = difference in the data reporting rate (DRR, %) between 2018 and 2020. DRR_{DID} measures the difference in ΔDRR between the average for the pilot countries and the reference like-pilot or non-pilot group of countries. SDM1 = standard deviation method based on the data reporting rate (DRR_{2018}), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean DRR_{2018} for the pilot countries, respectively. SDM3 = standard deviation method based on the three explanatory variables (DRR_{2018}, BRS_{2018}, and $ASCI_{2018}$), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean of each of the three variables for the pilot countries, respectively. BRS is biennial review score and ASCI is agricultural statistics capacity index. *, **, and *** mean DRR_{DID} is statistically significant at the 10%, 5%, and 1% levels, respectively.

Table 7: Difference in QDR between the pilot and non-pilot countries in 2020, by theme

	Theme 1	Theme 2	Theme 3	Theme 4	Theme 5	Theme 6	Theme 7
Pilot countries:							
QDR	5.0	1.2	6.7	9.4	3.3	0.0	0.0
Like-pilot (LP) countries:							
SDM1							
Within +/- 1 s.d. (LP1)							
QDR	3.2	8.1	12.7	7.9	1.3	0.0	0.0
QDR_D	1.8	-6.9	-6.0	1.5	2.0	0.0	0.0
P-value	0.551	0.239	0.063	0.827	0.459	n.a.	n.a.
Significance			*				
Within +/- 2 s.d. (LP2)							
QDR	7.6	9.8	11.2	8.5	2.6	0.0	0.0
QDR_D	-2.6	-8.6	-4.5	0.9	0.7	0.0	0.0
P-value	0.457	0.049	0.123	0.884	0.796	n.a.	n.a.
Significance		**					
Within +/- 3 s.d. (LP3)							
QDR	6.3	8.4	10.1	7.6	3.0	1.7	0.0
QDR_D	-1.3	-7.2	-3.4	1.8	0.3	-1.7	0.0
P-value	0.685	0.039	0.209	0.759	0.911	0.168	n.a.
Significance		**					
SDM3							
Within +/- 2 s.d. (LP2)							
QDR	6.2	7.3	12.8	7.5	2.8	0.0	0.0
QDR_D	-1.2	-6.1	-6.1	1.9	0.5	0.0	0.0
P-value	0.731	0.122	0.096	0.767	0.851	n.a.	n.a.
Significance			*				
Within +/- 3 s.d. (LP3)							
QDR	6.1	8.0	11.1	7.5	3.7	2.2	0.0
QDR_D	-1.1	-6.8	-4.4	1.9	-0.4	-2.2	0.0
P-value	0.728	0.057	0.123	0.755	0.898	0.168	n.a.
Significance		*					
All non-pilot countries							
QDR	5.9	13.0	12.4	9.5	6.4	2.7	0.0
QDR_D	-0.9	-11.8	-5.7	-0.1	-3.1	-2.7	0.0
P-value	0.767	0.001	0.063	0.993	0.405	0.073	n.a.
Significance		***	*			*	

Source: Authors' based on model results.

Notes: QDR = quality of data reported (or data with issues, %). QDR_D measures the difference in QDR between the pilot countries and the reference like-pilot or non-pilot group of countries. SDM1 = standard deviation method based on the data reporting rate (DRR_{2018}), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean DRR_{2018} for the pilot countries, respectively. SDM3 = standard deviation method based on the three explanatory variables (DRR_{2018}, BRS_{2018}, and $ASCI_{2018}$), where within +/- 1, 2, 3 s.d. refers to countries within 1, 2, and 3 standard deviations of the mean of each of the three variables for the pilot countries, respectively. BRS is biennial review score and ASCI is agricultural statistics capacity index. *, **, and *** mean QDR_D is statistically significant at the 10%, 5%, and 1% levels, respectively.

Relative improvements in DRR and QDR: propensity score matching method (PSMM)

Table 8 shows the probit results for selected model specifications that were used to achieve a good match. The probability to be selected into the pilot is mostly determined by DRR_{2018}, which has an inverted U-shaped relationship with the selection (positive for the level term and negative for the squared term). The other variables, BRS_{2018} and $ASCI_{2018}$, also have similar relationships with selection, although that for BRS_{2018} is not statistically significant. The effect of their interactions with DRR_{2018} are not statistically significant. Furthermore, the overall model fit results—chi squared values—when BRS_{2018} and $ASCI_{2018}$ in addition to their interactions with DRR_{2018} are included in the model are weakly or not statistically significant. This indicates multicollinearity. Therefore, the preferred model specification is the one that includes DRR_{2018} and its squared term only (which is Model 1 in Table 8).

Table 8: Probit results of factors affecting the probability of being in the pilot group for selected models

Variable	Model 1			Model 2			Model 3			Model 4		
	Coef.	z	Sig.	Coef.	z	Sig.	Coef.	z	Sig.	Coef.	z	Sig.
DRR_{2018}	4.60	2.43	**	5.00	1.95	*	5.60	1.76	*	8.04	1.85	*
DRR_{2018} squared	-0.03	2.45	**	-0.03	1.95	*	-0.03	1.77	*	-0.05	1.84	*
BRS_{2018}				0.64	1.06		-5.51	0.38		4.62	0.72	
BRS_{2018} squared										-0.39	0.50	
$ASCI_{2018}$				0.01	0.26		0.66	0.70		3.09	1.78	*
$ASCI_{2018}$ squared										-0.03	1.78	*
DRR_{2018}* BRS_{2018}							0.07	0.43				
DRR_{2018}* $ASCI_{2018}$							-0.01	0.69				
Intercept	-192.31	2.43	**	-210.11	1.97	**	-240.45	1.72	*	-437.35	2.16	**
model												
Chi squared	6.12		**	5.26			7.27			11.59		*
Pseudo R^2	0.27			0.34			0.35			0.45		

Source: Authors' based on model results.

Notes: DRR = data reporting rate. BRS = biennial review score. ASCI = agricultural statistics capacity index. Blank cells indicate not applicable. *, **, and *** mean the estimated coefficient is statistically significant at the 10%, 5%, and 1% levels, respectively.

Different matching techniques (nearest neighbor and kernel density) were applied to the different probit model specifications, and then balancing tests were used to identify the best match. Using 5 or 6 nearest neighbor matches on the preferred probit model specification (Model 1 in Table 8) generated the best match results. Table 9 shows the balancing test results for this. In the annex Table A4, the balancing test results for the various probit model specifications using different matching techniques are shown. Basically, using 5 or 6 nearest neighbor matches on the preferred probit model specification are the ones that produce the best balance in terms of being within various recommended limits of differences in the variables between the pilot (treatment) and like-pilot (control) groups (Rubin 2001).

Table 9: Balance between pilot (treatment) and like-pilot (control) groups, nearest neighbor matching

Method/variable	Mean				Variance	Bias	Rubin's		Covariates		
	T	C	t-test		ratio	%	B	R	Good	Concern	Bad
			Est.	p-value							
5 nearest neighbors											
DRR_{2018}											
Unmatched	83.0	73.9	1.36	0.180	0.24 **	84.5					
Matched	83.0	83.2	-0.10	0.924	0.90	-2.2					
DRR_{2018} squared											
Unmatched	6902.8	5678.2	1.31	0.196	0.24 **	80.5					
Matched	6902.8	6943.0	-0.10	0.924	0.90	-2.6					
Overall model											
Unmatched						82.5	83.8 *	0.00 *	0	0	100
Matched						2.4	6.9	1.64	100	0	0
6 nearest neighbors											
DRR_{2018}											
Unmatched	83.0	73.9	1.36	0.180	0.24 **	84.5					
Matched	83.0	83.0	0.00	1.000	1.09	0.0					
DRR_{2018} squared											
Unmatched	6902.8	5678.2	1.31	0.196	0.24 **	80.5					
Matched	6902.8	6901.9	-0.10	0.924	1.09	0.1					
Overall model											
Unmatched						82.5	83.8 *	0.00 *	0	0	100
Matched						0.0	11.7	1.26	100	0	0

Source: Authors' based on model results.

Notes: T = treatment or pilot group. C = control or like-pilot group. DRR = data reporting rate. For the variance ratio, * is for variables "of concern" or where ratio is in [0.5, 0.8) or (1.25, 2] and ** for "bad" variables or ratio <0.5 or >2 (Rubin, 2001). For Rubin's B and R, * indicates values outside the recommend limits of <25 and in [0.5 and 2], respectively (Rubin 2001). Covariates capture the percentage that are orthogonal to the propensity score with the specified variance ratios (% good, % of concern, and % bad). Blank cells indicate not applicable.

Using the 5 and 6 nearest neighbor matches on the preferred probit model specification, the PSMM estimates of DRR_{DID} and QDR_D are presented in Table 10. For DRR_{DID}, the statistically significant estimates are 7.6 to 9.1 %pts for all the themes together, and 6.9 to 9.0 %pts for theme 3, with or without the new food safety indicators. These results are like those obtained using the SDM, although the magnitudes are slightly higher for those obtained using PSMM. The estimates for QDR_D are not statistically significant.

Table 10: Propensity score matching results of differences in DRR and QDR between pilot and like-pilot countries, 2018 to 2020

	All	Theme							
		1	2	3	3, excl. food safety	4	5	6	7
Data reporting rate									
Pilot countries:									
ΔDRR	6.1	6.4	12.7	3.8	6.1	4.7	10.0	2.5	-2.5
Like-pilot (LP) countries:									
5 nearest neighbors									
ΔDRR	-3.0	3.7	3.8	-5.1	-2.0	-2.0	-3.5	3.6	-2.5
DRR_{DID}	9.1	2.7	8.9	9.0	8.1	6.7	13.5	-1.1	-0.1
t-test	2.82	0.36	1.54	3.72	3.00	0.64	1.08	0.09	0.01
Significance	***			***	***				
6 nearest neighbors									
ΔDRR	-1.6	3.1	4.5	-4.2	-0.0	0.4	1.5	3.9	0.1
DRR_{DID}	7.6	3.3	8.2	8.0	6.9	4.4	8.5	-1.4	-2.6
t-test	2.36	0.45	1.42	3.32	2.58	0.42	0.68	0.11	0.29
Significance	***			***	***				
Quality of data reported									
Pilot countries:									
QDR	5.2	5.0	1.25	6.7	n.a.	9.4	3.3	0.0	0.0
Like-pilot (LP) countries:									
5 nearest neighbors									
QDR	7.1	4.7	7.1	10.7	n.a.	6.6	1.9	0.0	0.0
QDR_D	-1.9	0.3	-5.9	-4.0		2.8	1.5	0.0	0.0
t-test	0.90	0.08	1.28	1.36		0.43	0.50	n.a.	n.a.
Significance									
6 nearest neighbors									
QDR	7.0	4.7	5.9	10.7	n.a.	6.7	2.1	0.0	0.0
QDR_D	-1.9	0.3	-4.7	-4.0		2.7	1.3	0.0	0.0
t-test	0.88	0.08	1.02	1.37		0.42	0.44	n.a.	n.a.
Significance									

Source: Authors' based on model results.

Notes: ΔDRR = difference in the data reporting rate (DRR) between 2018 and 2020. DRR_{DID} measures the difference in ΔDRR between the pilot countries and the reference like-pilot or non-pilot group of countries. QRR = quality of data reported (or data with issues, %). QDR_D measures the difference in QDR between the pilot countries and the reference like-pilot or non-pilot group of countries. *, **, and *** mean DRR_{DID} or QDR_D is statistically significant at the 10%, 5%, and 1% levels, respectively. n.a. = not applicable.

Relative improvements in DRR and QDR: two-stage weighted regression method (2SWRM)

The 2SWRM estimates of DRR_{DID} and QDR_D are presented in Table 11. For DRR_{DID}, the statistically significant estimates are like those obtained with PSMM, although of lower magnitude as expected; 7.0 to 7.7 %pts for all the themes together, and 6.1 to 7.6 %pts for theme 3, with or without the new food safety indicators. For QDR_D however, the estimates are statistically significant for theme 3, -4.8 to -3.8 %pts.

Table 11: Two-stage weighted regression results of differences in DRR and QDR between pilot and like-pilot countries

	All	Theme							
		1	2	3	3, excl. food safety	4	5	6	7
Data reporting rate									
Control for DRR_{2018} only									
DRR_{DID}	7.7	3.0	6.3	7.6	6.1	5.7	7.6	5.9	-7.3
t-test	2.43	0.47	1.12	3.14	2.16	0.61	0.62	0.47	0.82
Significance	**			***	**				
Control for DRR_{2018} and its squared value									
DRR_{DID}	7.0	2.0	6.2	7.3	6.1	4.6	6.4	3.2	-4.9
t-test	2.32	0.38	0.99	2.79	2.01	0.47	0.51	0.27	0.63
Significance	**			**	*				
Quality of data reported									
Control for DRR_{2018} only									
QDR_{D}	-2.6	-1.0	-5.3	-4.8	n.a.	2.9	0.6	0.0	0.0
t-test	1.51	0.29	1.47	1.95		0.48	0.25	n.a.	n.a.
Significance				*					
Control for DRR_{2018} and its squared value									
QDR_{D}	-2.1	-1.1	-4.6	-3.8	n.a.	2.8	1.0	0.0	0.0
t-test	1.31	0.28	1.30	2.14		0.45	0.43	n.a.	n.a.
Significance				**					

Source: Authors' based on model results.

Notes: DRR_{DID} measures the difference in the change in the data reporting rate (DRR) between the pilot countries and the reference like-pilot or non-pilot group of countries. QDR_{D} measures the difference in the quality of data reported (QDR) between the pilot countries and the reference like-pilot or non-pilot group of countries. *, **, and *** mean DRR_{DID} or QDR_{D} is statistically significant at the 10%, 5%, and 1% levels, respectively. n.a. = not applicable.

Despite the improvement in DRR, missing data continues to be a problem for many countries. Regarding the pilot countries for example, there only 10 indicators on which none of the five countries had a problem with (see Table A5 in the annex for details). As with all the other reporting countries, the most challenging indicators, in terms of the number of countries that has some missing data parameters, are on the main outcomes under themes 3 (ending hunger) and 4 (halving poverty). With at least one-half of all the countries experiencing some data gaps in several indicators, it is difficult to see how achievement of the goals and targets of the Malabo Declaration overall can be evaluated without measuring several of the outcome indicators.

6. CONCLUSIONS AND IMPLICATIONS

To help improve the quality of data available for policymaking in CAADP implementation and for achieving the Malabo commitments, IFPRI-ReSAKSS, with funding from the Gates Foundation and in collaboration with its implementing partners,[11] initiated a partnership framework for strengthening evidence-based planning and implementation in five selected African countries (Kenya, Malawi, Mozambique, Senegal, and Togo) in 2019 during the second round of the BR process. The capacity-strengthening activities focused on working with the country BR team to: assess the inaugural or 2018 BR process and identify the data gaps; constitute and train members of data clusters to compile and check the data; and then validate and submit the data to the AU.

The results of the initiative are presented in this paper by analyzing the effect of the activities on the BR data reporting rate (DRR) and the quality of data reported (QDR) in the five pilot countries, compared with what was achieved in like-pilot countries. The like-pilot countries are non-pilot countries that have characteristics like the pilot countries at the baseline which affect selection into the pilot or the data reporting and quality outcomes. Different methods (standard deviations, propensity score matching, and two-stage weighted regression) are used to identify the like-pilot countries, and a difference-in-difference method is used to estimate the effect of the pilot activities on the outcomes.

There is an overall improvement in DRR between 2018 and 2020 as several of the data gaps experienced in the 2018 BR have been filled. The capacity-strengthening activities conducted in the five pilot countries seem to have helped those countries to achieve better results. For the pilot countries, the largest improvement in DRR between 2018 and 2020 is in Togo (12.2 %pts) and Senegal (10.9 %pts), followed by Kenya (3.8 %pts) and Malawi (3.0 %pts), and then Mozambique (0.6 %pts). The average increase in DRR between 2018 and 2020 for all the five pilot countries is 6.1 %pts. When compared to the improvements made in like-pilot countries, the improvement in DRR in the pilot countries is higher by about 6.1 to 9.1 %pts.

There are differences in the change in DRR for the seven thematic areas of the BR, with different countries experiencing an increase, no change, or a decline in DRR for different themes. Comparing the average changes for the pilot countries to those for the like-pilot countries shows that the improvements made in DRR derived mostly from improvements in theme 3 on ending hunger, either with or without the new food safety indicators. Overall, more effort is required to fill data gaps in several indicators, especially those on theme 3 (ending hunger) and 4 (halving poverty), to comprehensively assess achievement of the goals and targets of the Malabo Declaration overall.

With respect to the quality of data reported (QDR, measured as the percent of the data reported that have issues in 2020), the best-performing pilot country is Senegal, with the least issues

[11] These are the Africa-based CG centers (IITA, ILRI, and IWMI).

(2.3%), followed by Kenya and Malawi (2.8% each), and then Togo (6.1%). Mozambique has the most data problems, with 12% of the data reported having quality issues, across most of the themes. On average, the pilot countries performed better than the like-pilot countries, but most of the differences are not statistically significant. There are few exceptions in the case of themes 2 on agricultural investment and 3 on ending hunger where, depending on the method used, the relative improvement in the average for the pilot countries are statistically significant.

Some of the data quality issues identified in this study may remain in the final data used for the 2020 BR, depending on the extent to which countries were able to internalize the issues at the time of the data cleaning and make the necessary corrections. In general, the process used, and time allocated for conducting the data checks and doing the cleaning by the AUC TWG (including informing and allowing countries to make corrections) need to be improved.

Looking at the factors that contributed to the higher performance in the pilot countries, the data clusters and their leadership are the most critical. This raises the question of how to sustain the data clusters to strengthen the trust among the members and for them to continuously update the data for the next rounds of the BR and reporting in 2022, 2024, and 2026. One important action may be to broaden the composition of the core BR team and data clusters to include relevant non-state and non-agricultural-sector actors. Although, the composition of stakeholders in the CAADP process in general seems to have improved in the 2020 BR process, the BR core team and the constituted data clusters were dominated by state and agricultural-sector actors.

Given that the data partnership framework and the capacity-strengthening activities conducted in the pilot countries were promoted in all the other countries during different AUC-BR training workshops, the financial support and hands-on technical assistance provided by ReSAKSS seems to have made the difference. These will be important for sustaining the data clusters to continuously update the data for the next rounds of the BR, as well as for successfully extending the work to other countries. This argument may seem flawed with the surprising low performance in Mozambique in both DRR and QDR, which severely drags down the average performance for the pilot countries. This is also because Mozambique, like Senegal, was treated as a full-pilot country where more funding was allocated and more intensive capacity strengthening was conducted, compared to Kenya, Malawi, and Togo, which were treated as light-pilot countries (see annex Table A1 for the differences). The main challenge in Mozambique, compared to the others, is little or ineffective engagement of non-state actors in the process, including engagement of all potential data providers in clusters and using all relevant data sources in compiling the data. A better strategy of engaging non-state actors that is rooted in the mutual accountability framework will be needed.

REFERENCES

AU (African Union). 2014. The Malabo Declaration on accelerated agricultural growth and transformation for shared prosperity and improved livelihoods. African Union Commission, Addis Ababa, Ethiopia.

AUC (African Union Commission). 2015. Agenda 2063: The Africa we want. African Union Commission, Addis Ababa, Ethiopia.

AUC (African Union Commission). 2018. Inaugural Biennial Review Report of the African Union Commission on the Implementation of the Malabo Declaration on Accelerated Agricultural Growth and Transformation for Shared prosperity and Improved Livelihoods. The 2017 progress report to the Assembly. Addis Ababa, Ethiopia: AUC.

AUC (African Union Commission). 2020. Second Biennial Review Report of the African Union Commission on the Implementation of the Malabo Declaration on Accelerated Agricultural Growth and Transformation for Shared prosperity and Improved Livelihoods. The 2019 progress report to the Assembly. Addis Ababa, Ethiopia: AUC.

Becker, S. O., and A. Ichino. 2002. Estimation of Average Treatment Effects Based on Propensity Scores. *Stata Journal* 2 (4): 358–377.

Benin, S., E. Nkonya, G. Okecho, E. Kato, J. Randriamamonjy, G. Lubadde, and M. Kyotalimye. 2011. Returns to Spending on Agricultural Extension: The Case of the National Agricultural Advisory Services (NAADS) Program of Uganda. *Agricultural Economics* 42 (2): 249–267.

Benin, S., Ulimwengu, J.M., and Tefera, W. 2018. The CAADP inaugural Biennial Review and Africa Agricultural Transformation Scorecard: Results and areas for improvement. IFPRI Discussion Paper 1754. Washington, DC: International Food Policy Research Institute (IFPRI).

Brodie, M.L. 2008. Data quality in information systems. *Information and Management* 3: 245–258.

Citro, C.F. 2014. Principles and Practices for a Federal Statistical Agency: Why, What, and to What Effect? *Statistics and Public Policy*, 1(1): 51–59.

Dehejia, R., and S. Wahba. 2002. Propensity Score Matching Methods for Non- experimental Causal Studies. *Review of Economics and Statistics* 84 (1): 151–161.

Fox, C., Levitin, A., and Redman, T. 1994. The notion of data and its quality dimensions. *Information Processing and Management* 30(1): 9–19.

FtF (Feed the Future). 2020. Le Projet d'Appui aux Politiques Agricoles (PAPA). Accessed March 3, 2020 http://www.papa.gouv.sn/categorie_publication/rapports-etudes/

Hartwich, F. Gottret, M.V., Babu, S. and Tola, J. 2007. Building Public–Private Partnerships for Agricultural Innovation in Latin America: Lessons from Capacity Strengthening. IFPRI Discussion Paper 00699. Washington, DC: International Food Policy Research Institute (IFPRI).

Imbens, G. M., and J. M. Wooldridge. 2009. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature* 47 (1): 5–86.

Jerven, M. 2013. *Poor numbers: how we are misled by African development statistics and what to do about it*. Ithaca, NY: Cornell University Press.

Jerven, M. 2014. The Political Economy of Agricultural Statistics and Input Subsidies: Evidence from India, Nigeria and Malawi. *Journal of Agrarian Change* 14(1): 129–45.

Johnson, M., and K. Flaherty. 2011. Strategic analysis and knowledge support systems for agriculture and rural development in Africa: Translating evidence into action. Food Security in Practice technical guide 6. Washington, D.C.: International Food Policy Research Institute.

Johnson M.E. 2018. Strategic Analysis and Knowledge Support Systems (SAKSS): Translating Evidence into Action. In: Henning C., Badiane O., and Krampe E. (eds.), *Development policies and policy processes in Africa: Modeling and evaluation*. Cham, Switzerland: Springer Open. https://doi.org/10.1007/978-3-319-60714-6

Loshin, D. 2011. *The practitioner's guide to data quality improvement*. Burlington, MA, USA: Morgan Kaufmann Publishers.

Matchaya, G., Nhemachena, C., Nhlengethwa, S. and Moyo, T. 2018. The Inaugural CAADP Biennial Review Mechanism (BR) experience: Insights on what worked and what needs improvement. ReSAKSS-SA note prepared for discussion at the CAADP BR policy-learning event, Libreville, Gabon, 23-25 April 2018. Retrieved from

NASEM (National Academies of Sciences, Engineering, and Medicine). 2017. *Principles and Practices for a Federal Statistical Agency: Sixth Edition*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24810.

Roberts, A.S. 2010. A Great and Revolutionary Law? The First Four Years of India's Right to Information Act. *Public Administration Review* 70(6): 1–32.

Robins, J. M., and A. Rotnitzky. 1995. Semi-parametric Efficiency in Multivariate Regression Models with Missing Data. *Journal of the American Statistical Association* 90 (429): 122–129.

Robins, J. M., A. Rotnitzky, and L.-P. Zhao. 1995. Analysis of Semi-parametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association* 90 (429): 106–121.

Rubin, D.B. 2001. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services and Outcomes Research Methodology* 2: 169–188.

Spielman, D., and K. von Grebmer. 2004. Public–Private Partnerships in Agricultural Research: An Analysis of Challenges Facing Industry and the Consultative Group on International Agricultural Research. Environment and Production Technology Division Discussion Paper No. 113. Washington D.C.: International Food Policy Research Institute.

StataCorp. 2019. STATA Version 16.0: Statistics/Data Analysis. College Station, TX, US: StataCorp. www.stata.com.

Steer, L. and Wathne, C. 2009. Mutual Accountability at country level: Emerging good practice. London, UK: Overseas Development Institute (ODI). Retrieved from https://www.odi.org/sites/odi.org.uk/ files/odi-assets/publications-opinion-files/4279.pdf

Vance, A., Lowry, P. and Eggett, D. 2013. Using Accountability to Reduce Access Policy Violations in Information Systems. *Journal of Management Information Systems* 29 (4): 263–289.

Wang, R.Y. and Strong, D.M. 1996. Beyond accuracy: what data quality means to data sonsumers. *Journal of Management Information Systems* 12(4): 5–33.

Yerramareddy, I. and Babu, S.C. 2018. "Knowledge and data: Achieving food and nutrition security through open access," IFPRI book chapters, in: 2018 Global food policy report, chapter 6, pages 46-53, International Food Policy Research Institute (IFPRI).

ANNEXES

Table A1: Pilot activities and outputs and countries implemented in

<table>
<tr><th></th><th>Activity</th><th>Output</th><th>Country</th></tr>
<tr><td rowspan="3">1.</td><td>Assessment of the inaugural or 2018 BR process and report to identify areas for improving the 2020 BR process and data systems</td><td rowspan="3">• Understanding of the data systems and validation processes
• Strengths in data, methodologies, processes, capacities, and systems
• Gaps and challenges in data, methodologies, processes, capacities, and systems
• Data clusters formed
• Inventory of current and potential data sources for each BR parameter/indicator
• Detail plan of activities to strengthen/improve the 2020 BR process and data systems
• Roadmap for the country to collect missing BR data that are not currently available
• Assessment Report based on above</td><td rowspan="3">Kenya
Malawi
Mozambique
Senegal
Togo</td></tr>
<tr><td>Specific activities:
• Review filled reporting template and submitted report
• Review the data collection/compilation process
• Review BR data sources (current and potential)
• Review any data protocols and standards
• Review validation process (platform, stakeholders, etc.)
• Review other reports to help identify data and capacity weaknesses (e.g. country BR briefs/lessons, past Capacity Needs Assessments, joint sector review—JSR—assessments etc.)</td></tr>
<tr><td>Questions:
• For indicators completed, which ones were done correctly?
• For indicators completed, which ones were done incorrectly?
• For indicators not completed, which ones could other data sources or methods have been used to complete them (either unknown at the time or known but not used)? If known but not used, why?
• For indicators not completed, which ones will the data gap remain?
• Does country have any data standards and protocols for collection, management and sharing? If yes, were they used during inaugural BR process?
• What are the characteristics of the data system used in the compilation process? For example:
o Which institutions (e.g., research center, consulting company, individual consultants, etc.) are used to collect or compile data?
o Which institutions give access to their data sets?
o Which centers of expertise are involved in helping with computation of indicators, completing templates, etc.?
• What are the mechanisms used to assess the quality of data?</td></tr>
</table>

	Activity	Output	Country
	○ How does it work? ○ Which institutions or centers of excellence are involved? • Was a JSR-like process used in the validation process? ○ If yes, evaluate against the AU guidelines in the JSR best practices. ○ If no, why not?		
	Methods/approaches: • Desk reviews • Interviews with stakeholders • Consultation workshops		
2.	**Strengthening human & institutional capacities on BR indicators and data compilation, analysis, management, protocols, and M&E**	• Training materials • Number of people trained (disaggregated by institution, topic, gender) • **Report on capacity strengthening activities and results** based on the above and including: i) technical support provided; and ii) perceptions of change in the process, capacities, and systems by institutions and individuals that have received or been potentially impacted by the training	Kenya Malawi Mozambique Senegal Togo
2a.	*Support launch of BR process in country*: • Support the BR team to communicate the BR process and engage all key stakeholders (especially non-state actors—NSAs) in process • Assist BR team to engage a BR Facilitator to facilitate the process from beginning to end *Specific activities related to formation of data clusters*: • Work with BR team to organize data clusters (or working groups) involving relevant stakeholders • Each cluster will be responsible for providing or compiling data and performing consistency checks (first line of data validation) *Specific activities related to training:* • Train BR team and data clusters on 2020 BR indicators, guidelines, and reporting template • Train BR team, data clusters, and stakeholders on standards and protocols for data compilation and management (e.g. curation, checking related sources, analyzing trends and correlations, etc.) to improve *data consistency and accuracy* • Train BR team and data clusters on standards and protocols for data documentation, metadata, etc. to improve *data traceability*		
	Methods/approaches: • Training workshops • Hands-on assistance		

	Activity	Output	Country
2b.	*Specific activities related to training:* • Train BR team and data clusters on strategic analysis for mutual accountability ○ Train on BR data standards and protocols for data sharing ○ Train on country eAtlas for data management ○ Train on eBR ○ Train on M&E system that links to the Malabo goals • Train BR team on JSR and BR best practices Work with others (e.g. Africa Lead) to train potential country BR Facilitators		Mozambique Senegal
	Methods/approaches: • Training workshops Hands-on assistance		
3.	**Strengthening ICT and data management/sharing platforms**	• Tablet with BR reporting template • Upgraded eAtlas/ICT platform with BR indicators	Mozambique Senegal
	Specific activities • Tablet uploaded with BR reporting template and eBR • Upgrade eAtlas (or other existing ICT platform for mutual accountability) to (i) accommodate BR indicators and analytical tools, (ii) ease data entry and management, and (iii) allow for data consistency checks		
4.	**Improving BR data quality (data compilation, validation, revision, and analysis)**	• Documented BR data standards and protocols developed or strengthened • Number (and share of total number) of indicators reported on (*these will be generated by the eBR system*) • Report on improvements made to data accuracy, traceability, consistency, and validation	
4a.	*Data compilation*—work with BR team and data clusters to: • Develop standards and protocols for collecting and managing the BR and related data to improve *data consistency, accuracy, and traceability* • Collect/compile the BR data according to the standards and protocols • Document all data sources for each indicator and (sub)parameters • Check consistency and accuracy of the data and computed indicators • Complete the 2020 BR reporting template (country BR report) *Data validation and revision*—work with BR team to: • Develop standards and protocols for validating the BR and related data (according the AU guidelines on JSR best practices)		Kenya Malawi Mozambique Senegal Togo

	Activity	Output	Country
	• Validate the data via an inclusive BR validation workshop • Revise data, if necessary, based on the outcome of the validations workshop		
4b.	*Strategic analysis*—work with BR team and data clusters to: • Identify strategic analyses needed (e.g. related to achieving a specific Malabo target) • Develop/describe the method(s) to be used and expected outputs • Analyze BR data accordingly and support country with the identified strategic analysis	• Strategic analysis identified and undertaken	Mozambique Senegal
5.	**Data platform and submission of BR data/report**	• Improved 2020 BR report (compared to the 2018 report)	Kenya Malawi Mozambique Senegal Togo
5a.	• Work with BR team to enter data correctly into the eBR system		
5b.	• Work with BR team to upload/link data to eAtlas or existing ICT platform • Work with BR team and data clusters to generate various analyses and reports from eAtlas/ICT platform to meet their own needs	• Strategic analysis reports	Mozambique Senegal

Source: Authors' illustration based on project documents

Notes: All activities (and related outputs) were conducted in all the five countries (Kenya, Malawi, Mozambique, Senegal, and Togo), except where indicated for Mozambique and Senegal only.

Table A2: Data reporting rate (DRR) in 2018 and 2020

Country	2018								2020									
	All	Theme							All	Theme								
		1	2	3	4	5	6	7		1	2	3	3, excl. food safety	4	5	6	7
Angola	52	50	70	60	43	0	71	55	88	100	96	89	94	69	75	100	60
Benin	77	100	80	76	67	44	86	82	95	100	93	99	100	83	69	100	100
Botswana	83	100	90	78	76	75	71	82	73	100	86	76	78	34	38	63	100
Burkina Faso	88	100	60	98	67	75	100	100	92	100	100	90	95	97	75	100	100
Burundi	92	96	90	94	71	100	100	100	95	100	100	97	99	72	100	100	100
Cabo Verde	77	100	60	73	81	38	100	100	80	100	100	81	82	41	69	100	80
Cameroon	55	100	40	62	33	0	43	55	77	100	100	69	72	79	63	100	80
Central Afr. Rep.	70	71	75	87	67	13	71	55	79	100	93	82	87	55	56	50	60
Chad	61	100	85	65	33	13	0	55	94	100	100	93	100	86	94	100	100
Congo	64	100	90	56	62	25	43	55	68	100	93	69	69	38	19	50	80
Côte d'Ivoire	84	100	100	97	57	25	86	82	91	100	100	94	96	72	69	63	100
Djibouti	76	96	80	68	76	38	100	100	37	93	36	35	31	14	19	0	80
DR Congo	54	71	60	60	38	0	29	82	41	100	71	25	31	28	50	63	80
Equatorial Guinea	57	100	50	46	38	31	100	73	25	93	43	12	6	7	19	25	100
Eswatini	80	100	85	83	52	56	71	100	87	100	100	87	95	69	69	100	100
Ethiopia	87	100	90	87	67	75	100	100	88	100	100	89	92	66	75	75	100
Gabon	73	100	100	81	52	0	71	55	76	100	71	74	75	90	56	63	60
Gambia	80	100	90	79	71	50	86	73	83	100	100	79	80	69	69	100	100
Ghana	92	100	90	97	76	75	100	91	96	100	100	93	95	100	100	100	100
Guinea	57	100	50	44	38	31	100	73	80	96	93	80	84	52	75	88	80
Kenya	88	100	85	94	52	100	57	100	92	100	93	99	99	55	75	88	80
Lesotho	80	100	90	79	52	75	43	100	65	100	79	63	67	34	50	75	80
Liberia	32	50	0	38	29	0	43	55	73	100	100	67	72	66	44	63	100
Madagascar	84	71	100	84	86	81	100	82	78	100	93	75	82	52	75	100	100
Malawi	86	100	80	87	71	75	86	100	89	100	100	88	98	72	88	88	100
Mali	93	96	90	92	90	88	100	100	92	100	100	90	89	97	88	100	80
Mauritania	64	100	85	52	52	25	71	82	80	100	100	76	80	48	94	100	100
Mauritius	81	100	90	76	52	69	100	100	70	100	82	63	61	66	56	88	80
Morocco	88	100	100	76	95	75	100	100	79	100	100	72	71	86	50	100	100
Mozambique	81	100	85	73	76	75	86	91	82	100	93	81	81	69	75	50	100
Namibia	83	100	95	86	48	69	86	91	79	93	100	79	86	55	56	75	100
Niger	61	100	30	56	57	38	43	100	72	74	79	78	85	41	69	50	80
Nigeria	57	100	70	52	29	13	43	82	81	100	100	81	83	59	63	50	100
Rwanda	95	100	100	90	90	100	100	100	94	100	100	92	92	97	75	100	100

Country	2018								2020									
	All	Theme							All	Theme								
		1	2	3	4	5	6	7		1	2	3	3, excl. food safety	4	5	6	7
Senegal	79	68	90	94	67	31	71	100	90	100	93	97	97	59	63	88	80
Seychelles	73	71	90	70	57	75	57	100	71	81	96	63	58	69	81	75	100
Sierra Leone	63	50	60	84	67	0	71	55	96	96	100	97	100	86	88	100	100
South Africa	90	100	100	84	81	94	86	100	87	100	100	86	94	83	63	88	80
Sudan	63	100	30	57	86	38	43	73	70	96	79	71	71	48	50	50	60
Tanzania	80	64	90	84	86	63	71	100	85	100	100	89	95	52	69	75	80
Togo	81	100	75	92	48	44	100	82	93	100	100	93	96	83	75	100	100
Tunisia	55	50	100	52	57	13	57	64	87	100	86	87	96	97	56	75	100
Uganda	89	100	100	94	62	56	100	100	93	100	100	95	97	79	75	100	100
Zambia	91	100	100	97	67	75	71	100	85	100	86	88	88	55	94	75	80
Zimbabwe	77	100	80	86	52	44	71	55	77	100	100	69	73	79	75	75	100

Source: Authors' calculation based on the BR data (AUC 2018, 2020).

Notes: Countries not shown (Algeria, Comoros, Eritrea, Guinea-Bissau, Egypt, Libya, Sahrawi, Sao Tome and Principe, Somalia, and South Sudan) indicate they did not submit data in the 2018 or 2020 BR, or both.

Table A3: Quality of data reported (DRR) in 2020

Country	All	Theme						
		1	2	3	4	5	6	7
Angola	6.8	5.6	25.0	7.3	0.0	0.0	0.0	0.0
Benin	4.2	14.8	0.0	2.2	0.0	8.9	0.0	0.0
Botswana	3.9	0.0	3.1	8.3	0.0	0.0	0.0	0.0
Burkina Faso	8.5	0.0	25.0	10.3	14.3	0.0	0.0	0.0
Burundi	4.0	0.0	0.0	4.5	14.3	0.0	0.0	0.0
Cabo Verde	10.8	0.0	0.0	21.8	0.0	9.5	0.0	0.0
Cameroon	19.8	3.7	50.0	25.4	16.7	25.0	0.0	0.0
Central Afr. Rep.	42.3	0.0	87.5	53.5	75.0	100.0	0.0	0.0
Chad	5.8	0.0	25.0	6.1	7.1	0.0	0.0	0.0
Congo	27.8	5.6	50.0	48.7	0.0	0.0	0.0	0.0
Côte d'Ivoire	7.8	8.3	0.0	13.4	0.0	2.2	0.0	0.0
Djibouti	18.6	31.5	50.0	9.1	0.0	0.0	0.0	0.0
DR Congo	8.9	2.8	0.0	26.2	0.0	5.3	0.0	0.0
Equatorial Guinea	11.7	25.9	0.0	0.0	0.0	0.0	0.0	0.0
Eswatini	8.9	3.7	0.0	12.6	17.5	5.3	0.0	0.0
Ethiopia	4.4	2.8	0.0	7.4	0.0	15.0	0.0	0.0
Gabon	5.0	0.0	0.0	11.0	0.0	0.0	0.0	0.0
Gambia	0.7	0.0	0.0	1.8	0.0	0.0	0.0	0.0
Ghana	6.9	0.0	0.0	8.7	14.3	0.0	16.7	0.0
Guinea	6.1	0.0	0.0	8.7	20.0	0.0	0.0	0.0
Kenya	2.8	0.0	3.1	0.9	25.0	0.0	0.0	0.0
Lesotho	7.3	3.7	17.1	12.2	0.0	0.0	0.0	0.0
Liberia	5.9	7.4	0.0	8.0	0.0	22.5	0.0	0.0
Madagascar	14.4	2.8	50.0	15.4	33.3	0.0	0.0	0.0
Malawi	2.8	0.0	0.0	6.0	0.0	3.3	0.0	0.0
Mali	4.5	8.3	0.0	6.2	0.0	4.4	0.0	0.0
Mauritania	12.8	2.8	0.0	19.7	0.0	0.0	50.0	0.0
Mauritius	7.9	0.0	0.0	17.3	0.0	0.0	0.0	0.0
Morocco	8.8	25.9	0.0	9.5	0.0	0.0	0.0	0.0
Mozambique	12.0	9.3	3.1	14.6	20.8	13.3	0.0	0.0
Namibia	11.6	10.4	0.0	21.3	0.0	4.2	0.0	0.0
Niger	15.7	11.1	0.0	25.1	0.0	6.1	0.0	0.0
Nigeria	13.3	2.8	25.0	14.3	33.3	0.0	0.0	0.0
Rwanda	3.1	0.0	0.0	2.0	16.7	3.5	0.0	0.0
Senegal	2.3	2.8	0.0	3.5	1.3	0.0	0.0	0.0
Seychelles	5.4	5.6	25.0	1.3	0.0	21.1	0.0	0.0
Sierra Leone	9.9	7.4	25.0	7.1	14.3	6.3	16.7	0.0
South Africa	21.3	21.3	28.1	21.9	28.6	0.0	0.0	0.0
Sudan	2.8	8.3	6.3	0.0	0.0	5.9	0.0	0.0
Tanzania	8.3	0.0	3.1	12.4	20.0	0.0	0.0	0.0
Togo	6.1	13.0	0.0	8.3	0.0	0.0	0.0	0.0
Tunisia	4.7	0.0	25.0	0.0	14.3	10.0	0.0	0.0
Uganda	3.6	8.3	0.0	0.0	14.3	0.0	0.0	0.0
Zambia	5.2	0.0	0.0	9.2	0.0	0.0	25.0	0.0
Zimbabwe	5.3	3.7	0.0	5.0	25.0	1.9	0.0	0.0

Source: Authors' calculation based on the BR data (AUC 2018, 2020).

Notes: Countries not shown (Algeria, Comoros, Eritrea, Guinea-Bissau, Egypt, Libya, Sahrawi, Sao Tome and Principe, Somalia, and South Sudan) indicate they did not submit data in the 2018 or 2020 BR, or both.

Table A4: Balancing between pilot (treatment) and like-pilot (control) groups using different matching methods and model specifications

Method/variable/model specification	Mean bias	Rubin's		Covariates		
		B	R	Good	Concern	Bad
Unmatched						
DRR_{2018} and DRR_{2018} squared	82.5	83.8 *	0.00 *	0	0	100
DRR_{2018}, DRR_{2018} squared, BRS_{2018}, and $ASCI_{2018}$	84.3	86.7 *	0.00 *	0	0	100
DRR_{2018}, DRR_{2018} squared, BRS_{2018}, $ASCI_{2018}$, DRR_{2018}*BRS_{2018}, DRR_{2018}*$ASCI_{2018}$	88.5	88.7 *	0.00 *	0	0	100
DRR_{2018}, DRR_{2018} squared, BRS_{2018}, BRS_{2018} squared, $ASCI_{2018}$, $ASCI_{2018}$ squared	82.5	101.0 *	0.00 *	0	0	100
Matched						
Nearest neighbor matching						
3 neighbors (DRR_{2018} and DRR_{2018} squared)	9.4	23.7	0.11	0	100	0
5 neighbors (DRR_{2018} and DRR_{2018} squared)	2.4	6.9	1.64	100	0	0
6 neighbors (DRR_{2018} and DRR_{2018} squared)	0.0	11.7	1.26	100	0	0
6 neighbors (DRR_{2018}, DRR_{2018} squared, BRS_{2018}, and $ASCI_{2018}$)	14.3	39.7 *	0.44 *	0	25	75
6 neighbors (DRR_{2018}, DRR_{2018} squared, BRS_{2018}, $ASCI_{2018}$, DRR_{2018}*BRS_{2018}, DRR_{2018}*$ASCI_{2018}$)	9.5	52.5 *	1.46	0	33	67
6 neighbors (DRR_{2018}, DRR_{2018} squared, BRS_{2018}, BRS_{2018} squared, $ASCI_{2018}$, $ASCI_{2018}$ squared)	15.7	116.5 *	0.25 *	0	33	67
Kernel density matching						
DRR_{2018} and DRR_{2018} squared	8.8	22.2	0.96	0	100	0
DRR_{2018}, DRR_{2018} squared, BRS_{2018}, and $ASCI_{2018}$	8.7	25.3 *	2.74 *	0	75	25
DRR_{2018}, DRR_{2018} squared, BRS_{2018}, and $ASCI_{2018}$, DRR_{2018}*BRS_{2018}, DRR_{2018}*$ASCI_{2018}$	12.6	85.8 *	0.60	33	0	67
DRR_{2018}, DRR_{2018} squared, BRS_{2018}, BRS_{2018} squared, $ASCI_{2018}$, $ASCI_{2018}$ squared	15.8	139.7 *	0.97	0	50	50

Source: Authors' based on model results.

Notes: DRR = data reporting rate. BRS = biennial review score. ASCI = agricultural statistics capacity index. For Rubin's B and R, * indicates values outside the recommend limits of <25 and in [0.5 and 2], respectively (Rubin 2001). Covariates capture the percentage that are orthogonal to the propensity score with the specified variance ratios (% good, % of concern, and % bad).

Table A5: CAADP BR indicators with missing data parameters, by pilot versus non-pilot countries

Indicator number and name[1]		Pilot countries (5)			Non-pilot countries (44)			All countries (49)			Number of parameters required
		Total	All[2]	Some[3]	Total	All[2]	Some[3]	Total	All[2]	Some[3]	
3.6iii	Trade aspect of food safety (food safety trade index)	3	1	2	39	28	11	42	29	13	5
5.1	Growth rate of the value of trade of agricultural commodities and services within Africa	4	0	4	38	6	32	42	6	36	8
3.6ii	Health aspect of food safety (food safety health index)	4	1	3	37	11	26	41	12	29	10
4.1v	Reduction rate of the gap between the wholesale price and farmgate price	4	3	1	30	26	4	34	29	5	4
3.5v	Growth rate of the proportion of Minimum Dietary Diversity-Women	4	3	1	29	24	5	33	27	6	2
4.3	Percentage of youth that is engaged in new job opportunities in agriculture value chains	4	1	3	28	12	16	32	13	19	4
4.4	Proportion of rural women that are empowered in agriculture	4	2	2	28	20	8	32	22	10	5
4.2	Number of priority agricultural commodity value chains for which a public-private partnership (PPP) is established with strong linkage to smallholder agriculture	2	1	1	27	21	6	29	22	7	4
3.5vi	Proportion of 6-23 months old children who meet the Minimum Acceptable Diet	2	1	1	25	14	11	27	15	12	2
3.5vii	Reduction in the prevalence (%) of adult individuals (15 years or older) found to be food insecure	2	2	0	25	16	9	27	18	9	2
5.2i	Trade Facilitation Index	2	0	2	24	1	23	26	1	25	6
4.1iii	Reduction rate of poverty headcount ratio, at national poverty line (% of population)	2	0	2	23	0	23	25	0	25	2
3.2i	Growth rate of agriculture value added per agricultural worker	2	0	2	22	3	19	24	3	21	18
6.1i	Percentage of farm, pastoral, and fisher households that are resilient to climate and weather-related shocks	4	0	4	20	11	9	24	11	13	2
6.2	Existence of government budget-lines to respond to spending needs on resilience building initiatives	1	0	1	23	3	20	24	3	21	4
3.4	Budget lines on social protection as % of the total resource requirements for coverage of the vulnerable social groups	1	0	1	22	5	17	23	5	18	20
3.3	Reduction rate of post-harvest losses for (at least) the 5 national priority commodities, and possibly for the 11 AU agriculture priority commodities	1	0	1	21	10	11	22	10	12	12
4.1iv	Reduction rate of poverty headcount ratio at international poverty line (% of population)	2	0	2	20	0	20	22	0	22	2
3.5iv	Prevalence of undernourished (% of the population)	3	0	3	17	9	8	20	9	11	2
7.1	Agricultural Statistics Capacity Index	1	0	1	19	3	16	20	3	17	2
3.2ii	Growth rate of agriculture value added, in constant US dollar, per hectare of agricultural land	0	0	0	19	2	17	19	2	17	18

Indicator number and name[1]		Pilot countries (5)			Non-pilot countries (44)			All countries (49)			Number of parameters required
		Total	All[2]	Some[3]	Total	All[2]	Some[3]	Total	All[2]	Some[3]	
5.2ii	Domestic Food Price Volatility Index	1	0	1	17	12	5	18	12	6	2
2.1iii	Official development assistance for agriculture, disbursement as % of commitment	1	0	1	15	2	13	16	2	14	8
3.1i	Fertilizer consumption (kilogram per hectare of arable land)	1	0	1	14	0	14	15	0	15	12
3.1iii	Growth rate of the ratio of supplied quality agriculture inputs (seed, breed, fingerlings) to the total national inputs requirements for the commodity	0	0	0	15	7	8	15	7	8	4
3.1v	Total agricultural research spending as a share of agricultural value added	1	0	1	13	2	11	14	2	12	8
3.1vi	Proportion of adult agricultural population with ownership or secure land rights over agricultural land	0	0	0	13	4	9	13	4	9	2
6.1ii	Share of agriculture land under sustainable land management practices	1	0	1	12	3	9	13	3	10	2
3.5ii	Prevalence of underweight (% of children under 5 years old)	0	0	0	12	4	8	12	4	8	2
3.6i	Level of improvement of food safety systems (food safety systems index)	0	0	0	12	1	11	12	1	11	14
4.1i	Growth rate of the agriculture value added	0	0	0	12	3	9	12	3	9	8
1.2	Existence of, and quality of multi-sectorial and multi-stakeholder coordination body	0	0	0	11	0	11	11	0	11	14
3.2iii	Growth rate of yields for the 5 national priority commodities, and possibly for the 11 AU agriculture priority commodities	1	0	1	10	1	9	11	1	10	12
3.5i	Prevalence of stunting (% of children under 5 years old)	0	0	0	11	3	8	11	3	8	2
3.5iii	Prevalence of wasting (% of children under 5 old)	1	0	1	9	1	8	10	1	9	2
2.1ii	Government agriculture expenditure as % of agriculture value added	1	0	1	8	1	7	9	1	8	8
2.4	Proportion of men and women engaged in agriculture with access to financial services (%)	1	0	1	8	3	5	9	3	6	4
2.1i	Government agriculture expenditure as % of total government expenditure	1	0	1	7	0	7	8	0	8	8
3.1ii	Growth rate of the size of irrigated areas from the value in 2000	0	0	0	7	0	7	7	0	7	2
3.1iv	Proportion of farmers having access to agricultural advisory services	0	0	0	7	5	2	7	5	2	2
7.2	Existence of inclusive institutionalized mechanisms and platforms for mutual accountability and peer review	1	0	1	6	2	4	7	2	5	3

Source: Authors' based on the 2020 BR report (AUC 2020).

Notes: [1] The table is sorted from the indicator with the greatest number of all countries reporting some missing data parameter (3.6iii) to the least (7.2). [2] All = countries with all the required data parameters missing. [3] Some = countries with some, but not all. of the required data parameters missing.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available **here**

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
www.ifpri.org

IFPRI HEADQUARTERS
1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org

AFRICA REGIONAL OFFICE
Titre 3396, Lot #2
BP 24063
Dakar – Almadies
Senegal
Phone: +221-33-869-9800
Email: ifpri-dakar@cgiar.org