Methodology for Improving Administrative Data for Use in an Integrated Agricultural Statistics System

Name of Author: Agnes M.N. Ssekiboobo (Mrs.)
Organizational Affiliation: School of Statistics and Planning, Makerere University
City, Country: Kampala, Uganda
Email address: agssekiboobo@bams.mak.ac.ug

ABSTRACT

According to Principle 5 of the Fundamental Principles of Official Statistics, data for statistical purposes may be drawn from all types of sources. However statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents. It is known from various country experiences and practices that using administrative data sources is cheaper and therefore will continue to play an important role in any agricultural statistics system as it requires fewer resources and is therefore more sustainable as a source. These sources also provide high frequency data and can better generate small area statistics based on technical and cost considerations. Most of the data for commercial crops, trade data, livestock, fisheries, forestry, water and early warning in developing countries are collected administratively. While administrative data is usually of good quality for executing administrative functions, concerns have been raised regarding the quality of this data for official purposes especially in developing countries. In addition, the assessment of the quality of the administrative data is done subjectively and is not detailed enough to cover the different quality dimensions. For the administrative data to serve a statistical purpose there is need to compile this data following a systematic, objective and standardized approach. A comprehensive analytical framework using a structure, conduct and performance paradigm and a data quality evaluation framework has been used to review the administrative sources in terms of production, quality and use of this data; to identify strengths, weaknesses and suitability of the existing agricultural data systems and review and analyze gaps. Lessons and experiences from developed countries that are relevant for improving the quality and use of these sources in developing countries have been identified. Results of a preliminary gap analysis and possible solutions for filling the gaps have also been presented.

Keywords: Sources, Quality, System, Experiences

==

1.0   Introduction

The Global Strategy to Improve Agriculture and Rural Statistics adopted by the United Nations Statistical Commission in 2010 aims to improve statistics in agriculture, livestock, aquaculture, small-scale fisheries and forestry production in developing countries and ensure the sustainability of their maintenance. One of the key components of the Global Action Plan is its Research Plan whose priorities include "Improving the methodology for using administrative data in agricultural statistics" (World Bank, *2010*).

The major objective of the research is to improve quality and use of administrative data in agricultural statistical systems of developing countries. It aims at developing strategies and methodologies for the improvement of collection and management of agricultural data from administrative sources and of their use in an integrated agricultural statistics system in developing countries. The expected primary products of this research include (i) a technical report that includes a country-tested and validated methodology to improve and make available administrative data for producing agricultural statistics in developing countries and (ii) a proposed strategy on how to use administrative data in cost effective agricultural statistics systems.

A number of authors have come up with definitions of administrative data and these include Brackstone (1987), Pronab (undated), UN (2011). However the working definition of administrative data for this research is *"information collected primarily for administrative (not statistical) purposes by government departments and other organisations usually during the delivery of a service or for the purposes of registration, record keeping or documentation of a transaction (Administrative Data Liaison Service UK, 2015).*

A four-step approach was proposed for this research namely, a thorough review on the quality and use of administrative data to improve agricultural statistics in developed countries; an analysis of the country assessment surveys and other documentation to identify methodological issues in using

administrative data in developing countries; based on the findings in (i) and (ii) and experiences in developed countries, develop a general methodology for integrating administrative data and other sources of auxiliary information with survey data as part of an integrated national statistical system; and select two developing countries for in-country testing to validate and improve the methodology developed in (iii), and produce final guidelines for developing countries to integrate administrative data into agricultural statistics.

2.0   Methodology of analysis

The work so far has involved the analysis of (i) **Country assessments reports** with information on the main sources of core agricultural data (Africa and Asia-Pacific regions); (ii) **Literature review** of research activities, empirical studies and country experiences on the sources, production, quality and use of administrative sources for statistical purposes in developing countries; (iii) **Key informant interviews;** and (iv) Responses from a **questionnaire** administered to the National Statistical Offices (NSOs) in Africa during the Africa Symposium for Statistical Development which was held in January 2015, in Uganda. A review of the quality of administrative data was done using t**he Data Quality Framework** which looked at the different quality dimensions in relation to the core items as stated in the Global Strategy (World Bank, 2010) and in the case of **data use**, a review was done by considering who uses the data, for what, ease of accessibility and frequency of use.

2.1   Proposed Framework for Improved Systems of Administrative Agricultural Statistics

A review of the analytical framework for assessment of Agricultural Market Information Systems developed by Kizito (2011) showed that the framework is comprehensive and can be used for review of other systems' structure, conduct and performance. The framework was modified (See Table 1) for the purpose of assessing the Administrative Data Systems for Agricultural Statistics (ADSAS) in countries. Other frameworks that could inform the process of review of relevant frameworks for agricultural administrative data systems are the Health Matrix Network (HMN) as well as the National Strategies for the Development of Statistics frameworks which are already being implemented in a number of developing countries with some success.

The review also identified systems with potential for providing data which can then be enhanced and improved for use as well as issues related to the role of administrative data in the production of official statistics. In addition, a critical analysis of administrative sources currently used by developing countries was done with an assessment of their strengths, weaknesses and suitability for use in an integrated and cost-effective agricultural statistics system.

Table 1: Structure, Conduct and Performance Design Issues of any ADSAS

Structural Design Issues	Conduct design issues	Performance (Quality of core data items covered)
1. ADSAS's perceived mandate (Aims, objectives, and clientele) • Policy formulation and monitoring • Food security planning and monitoring • Attainment of efficient markets • Clientele (e.g., farmers, traders, consumers, government, donors) 2. Institutional home, organization, and coordination • Public-, private-, farmer organization, or trader and NGO-based ADSAS • Provides complementary services that generate or increase value of information • Geographic coverage and range of commodities • Assuring coordination among stages o Integration of ADSAS Activities o Centralized or decentralized ADSAS activities	1. Information provided • Raw data • Analysis of raw data • Analytical reports 2. ICT used in the collection and dissemination • Traditional ICT (e.g., radio, television, and fax) • Modern ICT (e.g., email, internet, SMS) • PDAs and GPSs 3. Funding strategies 4. Data collection methods used • Structured questionnaire and enumerators • Wiki approach (users	1. Coverage 2. Comprehensiveness 3. Timeliness 4. Punctuality 5. Completeness 6. Relevance 7. Accuracy 8. Reliability 9. Integrity/ Credibility 10. Accessibility to different clientele 11. Clarity/interpretability 12. Comparability 13. Consistency/ Coherence/ 14. Sustainability of ADSAS • Financial support • User support • Cost minimization

o Specialization in ADSAS Products • Design of incentives for ADSAS staff • Profit orientation of the ADSAS 3. Nature of core data items covered (crop items, livestock items, poultry, aquaculture and fisheries products, agro-forestry production, agricultural inputs, land cover, apiary, and horticulture).	SMS or update web) 5. Quality control methods used 6. Feedback mechanism used	

Adapted from Kizito A. (2011)

3.0   Findings
3.1   Benefits of Administrative Data
Benefits of administrative data identified in the review included the fact that administrative data is cheaper to use, reduces response burden and also improves public image, can be collected more frequently and therefore more timely and up to date, can cover the whole target population and sometimes in developing countries, it is the only data available-especially for lower level administrative units. It is important to note that administrative data can help the NSOs cope with their mandate despite budget limitations.

3.2   Sources of Agricultural Administrative Data
Sources of administrative data identified which have application to agricultural statistics, include: Regular returns/reports by agricultural field/extension staff, (for various agricultural items, including crops & livestock), agricultural production and inputs manufacturers and distributors, farmers' associations, private businesses data, meteorological data, parastatals handling the major commercial/cash crops and traceability data (e.g. traceability livestock data). Tax data, land ownership records, farm registers and other registration or licensing systems are potential sources of this data but are rarely used in developing countries because of quality and coverage issues. Most of the institutions producing agricultural administrative data are public/ government but there is also a lot of potential for the private sector, farmer's and traders organizations as well as research organizations.

3.3   Uses of Administrative Data
Uses to which administrative data is put in developed and developing countries included: use for sampling frame construction and sample design, to cover data gaps from surveys and censuses, for forecasting, planning, provision of small area estimates and administrative uses including informing policy and decision making. While developed countries tended to use scientific approaches for adjusting or improving administrative data before use, developing countries tended to use simpler subjective methods like expert opinion, screen surveys, eye estimation etc. The extent of use of administrative data for agricultural statistics in developing countries is quite high especially for cash/commercial crops, crop forecasting/early warning, livestock and poultry, inputs and trade data. In developing countries the agricultural administrative data is mostly used for decision making while in developed countries substantial use is for improvement of statistics.

3.4   Data processing and Accessibility to Administrative Data
A lot of data from this source remain in raw form and are not turned into usable information in developing countries. Apart from exports and imports as well as agricultural price information which are often published widely, a lot of other agricultural administrative data are not widely disseminated for use.

3.5   Data Quality:
Assessment of the quality of administrative data is done subjectively in most developing countries and the assessments are not detailed enough to cover the different quality dimensions. Much of the agricultural administrative data is usually collected and compiled without using standard statistical procedures or personnel with training in statistical methods. Documentation of administrative data collection and processing methodologies is also poorly done. There is also low reliability of crop yield and production forecasts especially for food crops.

3.6   Institutional and Organisational Capacity
As a result of decentralization, at the lower reporting levels, there is delayed information flow, limited capacity, and data is only collected on an intermittent basis. Operational constraints make it difficult for extension staff to go around their areas of jurisdiction to collect data. They have several other

responsibilities and are not, normally, and legally under the statistics authorities. Agricultural returns written by agricultural extension staff are based on data collection forms that lack standardisation and can lead to reporting errors and inconsistencies.

3.7 Potential Limitations

Limitations of administrative data in developing countries include: Changes in administrative processes leading to inconsistencies in estimates across time, or reduced data availability; lack of standardized definitions of variables, units and identifiers which makes synthesis of multiple sources a challenge; under-coverage as in some cases, not all members of the target population participate in the administrative process; reporting error as a result of incentives underlying administrative processes; and the challenge of maintaining confidentiality. Other challenges include: The NSOs using administrative data source for a purpose different from the one for which the data was originally collected; access problems including legal restrictions, policy considerations, organisational arrangements and technical standards; differences in accuracy of the data by purpose (monitoring versus regulatory oversight);and diverse sources for same data with undocumented methods leading to conflicting estimates. Reduced contact of the public with the NSO, timeliness (in time, different reference period), missing data, and resistance to change are the other limitations.

3.8 Gap Analysis

A preliminary gap analysis was carried out (See Table 2) in order to identify areas of methodological improvement and possible solutions for using administrative data in an integrated agricultural statistics system.

Table 2: Gaps Identified and Proposed Possible Solutions to Fill the Gaps

	COMPONENT	GAPS	SOLUTIONS
1.	Collecting & Managing Administrative Data	• Divergence in figures from different sources on the same data item. • Missing data	• A robust Routine Agricultural Administrative data system linked to other agricultural sub systems. • A comprehensive administrative data systems manual covering all aspects of the data collection and management system.
2.	Structure of Organisations Collecting Administrative Agricultural Data	• Lack of qualified skilled staff and low staff retention mainly due to poor working conditions. • Failure to sustain good data collection systems. • Field staff often not well supervised. • Many & frequent changes in the administrative structure that affect data collection and management.	• Training and Support Supervision of the Team on a regular basis • Learning from best practice of other countries. • Developing protocols for metadata documentation and for correcting data inconsistencies. • All Agricultural Administrative units should have at least one statistician as part of their staff team. • Establishment of a Technical Working Group to monitor the process of administrative data production.
3.	Institutional Home	• Institutional Home in the traditional hierarchical organizations such as government departments and ministries give low-powered incentives and often result in staff demoralization and high staff turnover.	• Housing statistical activities in well resourced and semi-**autonomous authorities** with high-powered incentives that lead to better performance outcomes.
4.	Coordination	• Poor coordination or lack of co-ordination between the NSO and the various administrative agricultural data collection and management institutions.	• Better coordination between the NSOs and the various Organizations producing and managing administrative agricultural statistics. • **Best Practice – India:** Well established countrywide infrastructure; a **permanent village**

	COMPONENT	GAPS	SOLUTIONS
			reporting agency.
5.	Human Resource / Incentives to ADSAS staff	• Lack of qualified staff & low staff retention mainly due to poor working conditions. • Poor incentive structures among employees. • High Rate of staff attrition of trained and experienced staff from the government service. • Regular training is not common in most countries.	• **Training Best Practice: Tanzania Agricultural Routine Data System (ARDS) on:** • the common reporting formats • the Village/Ward data collection format • data management • data handling and analysis at district level
6.	Crop Forecasts	• The crop forecasts system is based on subjective appraisals. • Forecasts are prepared based on incomplete data • Changes are made in the production figures between one forecast and another, thus causing users to doubt the credibility of the figures reported.	• The objective/scientific techniques of crop forecasting should be developed. • Remote Sensing technology can also be used to develop reliable estimates of crop area and condition of the crop at various stages of growth for forecast purposes.
7.	Quality Control Procedures	• **Quality assessments** for agricultural administrative data systems in developing countries are **rarely done**. • Most ADSAS in developing countries **do not put emphasis on documenting agricultural data quality parameters**, and where they exist, they are **subjective**.	• Technical Working Group to ensure quality control measures are put in place and adhered to. • Determine the data quality dimensions that are more relevant for assessing quality of administrative data. • Come up with quantitative indicators of the relevant quality dimensions
9.	Institutional capacity	• No interface for dialogue between data producers and users. • Where they exist, channels of communications are not well set up and/or not regularly used as required • One third (30%) of the African countries for example operate below average of the expected level of the primary institutional infrastructure to produce agricultural statistics.	• The interface for dialogue between data producers and users should be set up where they do not exist and strengthened where they do. • Monitoring mechanism should be established to ensure the interface is used on a regular basis. • Institutional infrastructure (physical, statistical, GIS capability, statistical methodologies and classifications) should be strengthened where they are weak. This may require providing technical support to the countries. • Best practices should be drawn from the Asia-Pacific countries of Australia, Japan, Mongolia, New Zealand (APCAS 2012)

10.	Adequacy of inputs	• African Countries have inadequate resources to run the agricultural statistics systems effectively and efficiently. This is in terms of
		  o Finances
		  o Human resource and
		  o Physical infrastructure.
		• These result into late or irregular collection of information, inability to hire well trained staff, failure to ensure sustainability and production of poor quality data.
			The Countries need:
			• Financial support in terms of greater budgetary allocations for Agricultural Statistics in their national budgets as well as external funding where possible.
			• Capacity building in human resources for their agricultural activities.
			• To learn from best practice of other countries; resource mobilization
			• To share technical expertise through staff exchange programmes and study tours
			• To organize inputs dealers into

	COMPONENT	GAPS	SOLUTIONS
			associations that can provide reliable information on their agricultural routine activities.
11.	Data Use	• Limited use of the agricultural administrative data due to quality concerns.	• Use of cross checks and corrections made through survey data. • Combining multiple data sources with complementary strengths and weaknesses. • Develop good identifying variables. • Make the agricultural statistics production process more objective and transparent through digitisation and automation. • Review and revise the legal framework to cover administrative sources.

4.0   Conclusions

The Literature Review clearly shows that statistics agencies in the developed counties have done a lot of research in this area especially with regards to the several uses of administrative records. On the other hand, the Literature Review on Developing Countries also shows that despite questionable administrative data quality, many countries are already collecting and using administrative data in a number of situations. In fact, administrative data are the major source of data in many developing countries, especially for agricultural statistics. There are also new potential sources of data, especially from the private sector. This is partly due to the privatisation of formerly official functions and the growth of the private sector. The major problems are mostly data quality issues. Increased use of administrative data will therefore require quality issues to be addressed and this may mean that the idea of considering administrative data as a "reduced cost" source needs to be revisited. The NSOs have to work more closely with the various administrative data producers and have more influence in the generation of this data from the various sources. More analytical use of this data will then be achieved as users gain more confidence in its credibility.

References:
1. ADLS (2015). *Administrative data introduction*, http://www.adls.ac.uk/adls-resources/guidance/introduction/
2. African Development Bank (2014). *Country Assessment of Agricultural Statistical Systems in Africa: Measuring the Capacity of African Countries to Produce Timely, Reliable and Sustainable Agricultural Statistics.*
3. Brackstone, G.J. (1987). *Statistical Issues of Administrative Data: Issues and Challenges.*
4. Kizito, A. M. (2011). *The Structure, Conduct, and Performance of Agricultural Market Information Systems in Sub-Saharan Africa*. Agricultural, Food, and Resource Economics, East Lansing, Michigan, Michigan State University, Ph.D. Dissertation.
5. Pronab Sen (undated). *Challenges of Using Administrative Data for Statistical Purposes: India* Country paper.
6. United Nations (2011). *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*. Available at http://unstats.un.org/unsd/EconStatKB/Knowledgebase Article10349.aspx."
7. World Bank. (2010). *Global Strategy to Improve Agricultural and Rural Statistics.* Washington, DC. World Bank. https://openknowledge.worldbank.org/handle/10986/12402 License: CC BY 3.0 Unreported.