Agricultural Statistics Requirements and Uses: Discussion

Michael Ward

Development Data Group, The World Bank

The nature of "agricultural statistics" has undergone considerable transformation and conceptual change over the past half-century as the range of requirements and uses of agricultural statistics have grown — features well reflected in the extensive topic coverage of these four papers. By comparison, in an earlier expansionary period of the 50s and 60s, the scope of agricultural statistics was well defined. Young, newly "graduated" but not yet "qualified" statisticians (in the sense of being experienced and tried in the scientific inquiry in the field) from Europe and the USA often first learned the basic craft of their trade carrying out crop surveys in the open expanses of Africa and Asia as cadet statisticians, Peace Corps workers or VSOs. Wherever conducted, however, such apprenticeships in agricultural statistics invariably involved stretching sample rope lines across variegated crop stands, throwing hoops over field plots, pacing ill-defined boundaries, and measuring landholdings using different ingenious procedures. (The latter often only merely reflected the idiosyncratic preferences of supervisors who had their own ways to solve the difficult problems of estimating crop areas where the phenomenon of a square or rectangular field was such an usual event as to arouse excitement!) These traditional roles were performed under the watchful eyes and expert guidance of old hands who knew the pattern of the seasons and the social significance of cropping and grazing practices.

Agriculture, however, was important. In dominantly rural economies, "Economic Development" was synonymous with agricultural progress, which was primarily concerned with raising total agriculture output, increasing crop yields, improving productivity and matching crops to growing conditions, i.e. soil types, land profiles and microclimatic conditions. Agriculturalists believed they could "fix" what was right for farmers to do. In this respect, agricultural policymakers in developing countries followed closely in the footsteps of the missionaries and administrators before them. Agricultural statisticians were among the especially anointed to measure how much output would be necessary to bring tangible progress to people's lives, or at least to alleviate hunger. Along the way, sine qua non, the teachings and elegant technical mastery of survey designs worked out by the statistical Brahmins of the profession were unquestionably accepted.

In those days, however, administrators and agriculturalists tended to look at land purely as a factor of production. Indeed, in the vastness of Africa, India and China, many would have regarded land (and its availability) as the most important factor in the development equation. In holding on to this belief, nevertheless, they usually failed to acknowledge the many other significant and often confusing links between production and people. In particular, the way people and their land were invariably held in a complex genetic vortex of closely knit social bonds was mostly ignored. A related consequence was that the same policymakers were also mostly unconscious of the potentially damaging effects on the environment of different land preparation techniques, land-use methods and land cropping practices. Importantly, then, how official advice to grow "more" might upset the delicate ecological balance or fly in the face of social realities and perceptions of household risk was rarely examined. Nor too, was there much participatory rural appraisal or true community involvement in agricultural decisions which might have alerted policymakers to other dimensions to agricultural policy, and also given more emphasis to needs and demands as opposed to output.

Over time, administrators have learned to better understand the nature of agricultural development, and to be less “supply-side” focused than, perhaps, Ray Bollman suggests. They have become more attentive to the demands of policy, and more definitionally careful and analytically open-minded about development processes. Policymakers are now certainly more aware of the important “upstream” and “downstream” implications of altering land use patterns with their negative environmental effects and population displacements. But, as Ray Bollman explains, the role of agriculture has declined in importance in many countries. Its share in GDP in most industrially developed countries has fallen significantly as well as progressively since the 1950s. Consequently, the rural population that agriculture once supported has declined. People have moved to the towns or taken other non-farm jobs. The basic food and raw material commodities that dominated production are no longer as significant because incomes have risen and synthetic substitutes or replacements for many intermediate raw materials have been found. Correspondingly, relative prices have shifted, because changed demand patterns have interacted with these altered supply conditions. It is estimated that over the past two to three decades, agricultural output overall in developing countries has declined by 1 percent in real terms every year. However, despite this, the incidence — although not the extent — of poverty in low income countries has fallen, and many people’s diets have improved and become more varied. There is a growing demand for different exotic crops and for greater choice. By its very nature, demand globally has become more fickle and fractured and the potential upper limits to the production of mainline crops has been defined at relatively lower levels.

At the same time, however, the demand for livestock products has expanded and grown with higher incomes. Although this has helped to slow down, if only marginally, the relative fall in grain output because of the accompanying need for animal feed, the overall effect of these trends has been to constrain crop production. The desire for more effective control over the vagaries of the weather has also led to the widespread adoption of “production control” measures such as the application of pesticides, insecticides, fertilizers and other protective procedures. Alongside the implementation of deforestation practices, often to expand “ranching” and marginal crop growing as well as land recovery programs, these measures have altered the face of agriculture. They have also contributed to other changes in the agricultural landscape that, as Kevin Parris describes in his paper, have had a significant and continuing adverse impact on biodiversity, wildlife habitats and the overall emission of greenhouse gases.

But it is not only the agriculturalists and farmers who are to blame. The “Driving Force (Pressure) - State - Response” mechanisms Parris refers to take many forms. Changed farm management practices, influenced by taxation and subsidies, and the availability of farm financial resources (at the time they are needed), also play a vitally important role in altering land use patterns and practices. A broader based conceptual understanding is just beginning to take root among policymakers. Here, Kevin Parris’ paper makes a significant contribution to the intellectual debate with practical illustrations of the ways to monitor the interrelations between agriculture and the environment.

Each of these papers takes up a very different topic and it is difficult to deal with them in an integrated overall way. But they suggest that a broader holistic approach could be taken which treats each as a contribution to statistical understanding in the context of the changing emphasis of agricultural statistics in the specific and interrelated areas of data definitions and classifications as more explicitly laid out by Odell Larsen and Pratap Narain in their paper. Such procedures are closely associated with concerns of analytical linkages and impact assessment, and underline the need to correct those informational deficiencies that hinder policy formation. Indeed, the subject of standard classifications is one of the most important for data producers and users as analysts because it raises the whole issue of sectoral coverage and the definition of the appropriate components, structures and linkages of an agricultural information system whose elements relate to operational policy and objectives. Statisticians, nevertheless, need to step back and take a hard strategic look at the direction in which classifications and data systems should go, especially in light of these broader concerns of agricultural statistics. Changing policies, institutions, technology developments and multinational relations, e.g. WTO, European Union, NAFTA, etc., all form an important component of this framework, and hence, review some of the traditional approaches to this process.

The starting point of the issues addressed by Odell Larsen and Pratap Narain refers to "data" in the context of the need for standards. But, data are not just "given facts"; they constitute the core of the problem. The dictionary describes "datum" or "data" as something "given" as a basis for reasoning or inference; it is a standard that can be used, therefore, as the foundation for calculating and measuring. Linked to a classification system and an accompanying analytical framework, data can then define the essentials of the statistical modus operandi. They provide the relevant dimensions to the estimation functions and corresponding aggregation functions that make collected and reported numbers meaningful by converting these statistics into usable information. In agriculture, as compared with "high-tech" industry, analysts at least start with the advantage that, unlike (for example) computer production, the basic nature of the phenomenon to be measured, viz. agricultural final output, is reasonably well defined and remains relatively the same over time, and "quality" improvements can be relatively easily identified.

Probably the best book that has been written on the interrelationship between statistical classifications and analysis (unintentionally, I suspect) is *Zen and the Art of Motorcycle Maintenance*. The work demonstrates that it is possible to completely take apart a motorcycle and put all the various bits into clearly marked different tins and boxes (with separate compartments) but not yet learn from this (a) what a motor cycle itself looks like, or, even more important, (b) understand anything about how it works. It could be added that a wide variety of ways exists in which such parts could have been classified, quite apart from depending on their common (physically alike) characteristics. For example, the parts could have been separated by their material composition — iron, steel, chrome, plastic, glass, etc. — or by their size, or whether they were relevant to some chain of mechanical, structural or electrical functions, etc. In other words, each form of classification or grouping chosen can be designed to serve a different specific purpose. But, once formed for that "mutually exclusive" purpose, it is not always easy to relate a single grouping to some other purpose. The fact that a large number of elements can be meaningfully classified in several equally meaningful ways *at the same time* is also lost.

In particular, showing "plural" characteristics becomes an essential requirement for understanding processes of adaptation and change in agriculture. With multiple features, different groups, orders and classifications can be constructed using the same essential "building block" and the overall informational edifice can be erected on the basis of the various core elements intrinsic to the basic units. The relevance of such an approach and classification schema to the understanding of specific socioeconomic problems, particularly say in the area of farm household benefit-incidence appraisals related to policy support for rural development, in general, is of critical policy interest.

If the idea is to identify in a more innovative and radical way currently unperceived conceptual and practical linkages between phenomena and their classification so as to strengthen vertical and horizontal networks, then a traditional taxonomist view that favors a unique, mutually exclusive and exhaustive process of classifying all entities in an unambiguous, identifiable procedure that is linked to a more general but usually unspecified objective serves little purpose. The opponents of standardized typologies prefer to see classifications more closely related to conceptual models as to how the system is believed to work and that reflect how things change in a dynamic way. With the advantage for the huge advances made in computer technology and data processing in recent years, it is now much easier to move to more fluid "free form" compilation procedures using core building blocks, allowing a more flexible treatment of variables and their identifiable components. This is a feasible proposition if information is properly captured at the basic level to reflect the different facets and relationships applicable to each of these fundamental entities. Clearly, institutions and economic arrangements have changed significantly since the frameworks inherent in most current international classifications were first formulated.

So far, however, attempts to launch new initiatives to develop more fluid and unstructured but interactive data grouping systems alongside existing conventional frameworks have floundered. These are designed to facilitate integration whilst preserving established administrative taxonomies. The elements of interest to policymakers are embedded in the units of observation and their identifiable relationships with other units. It is therefore possible to "create" operationally relevant "bespoke" classifications for specific purposes and to categorize units according to more than one characteristic — the latter constraint being one of the fundamental limitations of most traditional classifications systems. This would help policy analysts to derive the "rural" characteristics sought by Ray Bollman.

Thus, the relative merits propounded by Odell Larsen and Pratap Narain of applying existing well defined and universally recognized classifications to structure information perhaps need to be weighed against some alternative unstructured freestanding system of flexible policy-linked statistical building blocks. Such blocks could be formed around defined units of observation and measurement which already possess the intrinsic components of potential classifications of socioeconomic interest. Despite the huge advances made in recent years in electronic processing that mean data can now be left in a primordial form and have less need to be formally structured and reconfigured in conventional dimensions and a single dimensional way, there are still many complexities to overcome before such a radical approach could be adopted. It has probably to be acknowledged here, that a principal components methodology or factor analysis path to classifications and policy review might have useful research applications in selected socioeconomic contexts, but perhaps cannot be applied in a more general context, even in parallel with existing systems, at the present time.

The bottom line, however, is that too much is invested in the historical legacy of traditional categorization and institutional nomenclatures and that no one wants to change the present way of doing things. New "conventional" classification systems seem to sprout up overnight and, driven by national priorities and characteristics, arbitrary extensions to existing structures are constantly being constructed. Few organizations have the right incentives to introduce fundamental changes to classifications because those in existence serve established standard administrative functions.

There are, however, other related problems in dealing with data ordering to consider. One concern is that "event time" is not analogous to, nor in phase with, the discrete events of chronological time; nor is it necessarily complementary to the agricultural calendar. For centuries, people in China, Russia, Thailand and India have lived with eras and dynasties that marked the passage of time but, for statistical purposes and making valid comparisons, events have to be placed into very distinctly defined and even time periods.

Devising parallel "official" procedures may only create confusion. Across the range of statistical issues that have been raised, analysts are separately dealing with plots, farm holdings, farmers, households and communities, and there are often important mismatches between the selected units of classification and the actual units of observation, or at least the actual reporting units. In some cases, these may not even be the units of policy relevance or interest, but the predetermined limitations of the sampling frame constrain choice.

Finally, how else can the uses of agricultural data be improved and interests of policymakers be better served? One important way, as implied in the approach of Parris, is to refine the linkages between inputs and outputs to try to define the impacts of actions and activities more exactly. Restructuring the "pressure-state-response" relationships into a multidimensional matrix showing the varying impact of nutrient flows, farming practices, resource situations and cultural and institutional behavior on soil, water, wildlife, landscape, greenhouse gases, etc. would certainly help to identify better what is analytically important.

Each of these papers contributes to the present understanding of the role of agricultural statistics. But what perspectives do they give for future agricultural and development policy and for supporting progress into the 21st Century? How will they influence and refine thinking and help define suitable strategies with respect to the future collection and analysis of agricultural statistics? It may be too much, perhaps, to ask of them to provide definitive answers to all these questions, but the papers do draw attention to some interesting broader issues, and provide indications, observations and hints of some of several important, if not critical, questions that should attract policymakers' attention.