Collecting and processing livestock statistics: operational experience from Nigeria

Table of Contents

by
W. O. Ajayi[1]
Gil Rodriguez[2]
John Bennett[3]

1 W. O. Ajayi is Head of Data Collection and Market Information Unit, Federal Livestock Department (FLD).

2 Gil Rodriguez was formerly Economist with The Second Livestock Development Project (SLDP).

3 John Bennett is Information Systems Specialist with SLDP.

alpan - AFRICAN LIVESTOCK POLICY ANALYSIS NETWORK

Network Paper No. 18 October 1988

INTERNATIONAL LIVESTOCK CENTRE FOR AFRICA (ILCA)

P.O. Box 5689, Addis Ababa, Ethiopia • Tel: 18 32 15 • Telex: ADDIS 21207

Table of Contents

Introduction

1. Accurate and timely data sets are needed by various participants in the livestock subsector. Farm gate prices are needed by producers to determine their herd structure and off-take and the corresponding input levels. Farm-wholesale-retail price margins are required by the relevant distributors to determine the type and levels of livestock products to offer at the retail end. Government planners are interested in domestic livestock supply-demand gaps to obtain an idea of the potential foreign exchange pressures from imports.

2. Apart from monitoring the main economic activities, data sets are generated to audit particular government functions. Statistics on livestock movements at established control posts are collected to provide benchmark information of animal related tax revenues and levels of live animal imports from neighbouring countries. Statistics on disease outbreaks are gathered to evaluate the effectiveness of vaccination campaigns.

3. Since a livestock information set is a crucial input in livestock policy planning, implementation, and monitoring, it is important that we focus attention on who manages the collection and processing of key livestock statistics. In the case of Nigeria, the Federal Livestock Department (FLD) has, since 1981, been given the responsibility for the development of information support to the livestock subsector. Presently, the data collection mandate of the Federal Livestock Department is being enhanced through a technical support package provided by the Second Livestock Development Project (SLDP). SLDP is funded by a World Bank loan. The total external loan extended to Nigeria under the project is $82 million.

4. It is within the FLD context that this paper aims to:

(i) provide a perspective of the data collection system used in generating livestock statistics;

(ii) present a set of information priorities required by the livestock subsector for planning purposes;

(iii) describe a data processing structure which will facilitate the management of a livestock data bank.

Data issues

5. Data sets have several characteristics. Firstly, a given information will have traits which can be conflicting. For example, a given livestock price such as for beef can be reported with a pre-determined frequency (giving it a semblance of timeliness) but it may not be accurate. The most precise information may not be relevant any more when one makes a decision. Hence, general livestock statistics cannot satisfy all the data quality demands of the various users.

6. Secondly, the accuracy of different data inputs is dependent on both the quality of the estimates of the components as well as the interrelationship existing between them. Consider the simple relationship given by the following: S = P + M + I

Where

S is livestock supply level
P is domestic livestock production
M is livestock imports
I is the herd inventory.

Assume further that the following ratios are true:

(i) The relative errors (defined as the difference between the actual data and its estimate divided by the actual data) are: P (20%); M (30%); and I (10%).

(ii) The ratios P/S, M/S, I/S are 0.4, 0.4, 0.2 respectively.

7. The relative error of "S" will be 22% (20%* 0.4 + 30%* 0.4 + 10%* 0.2). This means that the estimate of S will be partly influenced by the quality of the estimates of its components. The nature of the relationship between given data also affect the accuracy of estimates. For example, however accurate data measurement is for "S", "P" or "M", "I" (i.e. inventory) will be subject to large errors because the I/S ratio is relatively small.

8. Thirdly, data use makes sense if one selects the appropriate statistical index. For example, causal analysis always leads one to assume that an average measures the representativeness of the population being scrutinized. If the population is the liveweight price of cattle with the following values (Naira/head): 1,000, 2,000, 2,000, and 5,000, the average price is 2,500 which is not even the price of a single animal. A more representative value in the example cited is the mode which is 2,000 (i.e., the price observation with the largest number of frequency count).

9. Fourthly, data cost money. An aerial national livestock census for Nigeria can easily cost US$ 2-3 million. If one is undertaking a survey (i.e. analysing the subsector of a given population), then if the accuracy of the sample statistical estimates (as measured by the standard error) is to be increased by 50%, then the sample size will need to be increased by a multiple of four. This simply means that instead of interviewing 100 farmers, 400 will need to be covered. It is obvious that the additional 300 respondents will cost additional money (e.g. higher traveling cost). The magnitude of data activity will partly be determined by its value relative to its cost. It should further be noted that one should not seek comfort in a layer sample size in terms of assured statistical accuracy for the estimate. Measurement errors can easily creep in with a large sample size - the interviewers can be afflicted with boredom in

conducting the survey and hence become careless in accepting large sample responses.

10. Fifth, the end-use of a data set can influence its quality. A case in point are livestock trade data. If a country has stringent tariff and non-tariff sanctions, then animal imports will likely be understated. It is natural for an entrepreneur to dodge government regulations if it will affect his profits. A livestock trade policy analyst must use his ingenuity in data validation through corresponding major trade partners and other information correlated with live animals traded to detect the magnitude of the bias in his data.

Data collection modes

11. Data can be gathered in various ways. The principal means of collecting data are:

1. Census - involves complete enumeration of the population (code 0);

2. Probability and purposive sampling - a subset of the population is covered in accordance with some probability weights or in terms of accessibility of the respondents (code 1);

3. Subjective and recorded estimates of government personnel involved in implementing a fiscal function and extension services, parastatal activities, and other law enforcement duties (code 2);

4. Subjective estimates of commodity specialists (code 3);

5. Data derived from international documents published by other countries or by an international agency (code 4).

The codes corresponding to items (1) to (5) are not reflective of the accuracy of the data but rather yield a profile of the objectivity of the information and an implicit measure of the efforts of the government to build a sound livestock data base

12. The data collection schemes mentioned previously are applied to the information collated through the Nigerian Livestock Information Service 1984 annual report. The results are:

Date Required	Code
Domestic slaughter cattle, goats, sheep and pigs	3, 2
Live cattle imports	3, 4, 2
Immigrant slaughter cattle, sheep, goats	3, 2
Importation of animal feedstuffs	3, 4
Importation of milk and dairy products	3, 4
Transportation of cattle by road, rail, hoof	3, 2
Transportation of sheep, goats, camels, etc.	3, 2
Average prepared feed prices	1
Average slaughter cattle, sheep, goat, etc. price	1
Average price for beef, mutton, pork eggs, etc.	1, 2
Average red Sokoto goat skin prices (heavy type)	1

13. Several remarks can be made regarding the preceding classification exercise. Firstly, a large portion of the livestock data base is created through the subjective opinions of livestock specialists. The basis of their estimates is seldom documented. Since different specialists will tend to report different levels within and across time periods, it will be difficult to separate systematic variations arising from a given set of factors of interest to policy makers from variations arising from educated guesses. Such a situation is illustrated by the data on cattle movements for the period 1970-84.

(a) Animal numbers moved by hoof rose from 171897 (1970) to 344883 (1971), an

increase of 100%;
(b) Animals transported by road went up from 49461 (1970) to 842461 head (1984);
(c) Animals moved by rail declined from 96308 (1970) to 46412 (1984);
(d) Cattle numbers moved on hoof plummeted from 171897 (1970) to 24406 head (1984).

14. Sharp changes are noted on the said data set. A valid inquiry is: are such patterns of live cattle transport reflective of changing transport costs or marketing institutions or simply data aberrations arising from the temperament and innate livestock knowledge of the person in charge of the collection of the animal movement statistics?

15. A second major point to be made is the absence of any standard sampling scheme at the field level. For example, while only retail livestock prices are reported, the policy adopted by the livestock officer at the state level with respect to market coverage and selection of respondents is highly arbitrary. As a result, the average livestock prices reported may not be truly representative of market conditions if respondents are limited to specific market areas. In the current reporting form, livestock prices are already difficult to use for policy analysis purposes since they are not correlated with specific livestock qualities due to the absence of carcass grading schemes and other modes of market standardization.

16. A possible offshoot of the price measurement errors is a distorted price margin between supply and final consumption points. For example, market reports indicate at times higher retail beef prices in areas like Ibadan relative to Kaduna resulting from the use of the averaging process.

17. Thirdly, the present livestock data collection scheme is oriented more towards moderate monitoring of disease outbreaks and tax collection rather than generating information needed to evaluate overall policy impacts within the livestock subsector and on other related sectors. As a result, data quality checks are seldom undertaken since the utilization of the data inputs for quantitative policy modelling is very low.

Agenda for data prioritization and collection

18. Data ranking is not a simple task because data demand is not unique. Each sector of the economy will not only require specific data forms but will also manifest shifting demand arising from changing market and political conditions.

19. The method adopted in this section with respect to livestock statistics prioritization is a pragmatic one. It classifies the livestock information set based on its strategic role in policy planning.

20. Foremost in the livestock data set which must be given immediate attention is that pertaining to livestock population. Acceptable cattle herd inventories are needed to evaluate livestock self-sufficiency status, livestock incomes, livestock import levels, regional and national technological changes and protein deficiencies. Current cattle herd estimates are based partly on special studies (e.g. ILCA data on herd composition and productivity indices for specific sites) and subjective estimates based on government vaccination records.

21. The existing 11 to 13 million heads of cattle reported for Nigeria by FLD have been partly used to analyse future beef self-sufficiency levels. A difficulty for such an analytic exercise is that informal trading in live animals between Nigeria and its bordering neighbours can range from 350 to 500 thousand head per year. Hence, if one adopts a stance of beef self-sufficiency, this implies foregoing the live animal import option. This will then require that a lower cattle herd inventory level will have to be used in making beef demand-supply gap estimates.

22. In the absence of good data sets on variables highly correlated with cattle population, the only option to estimate objectively cattle numbers will be to undertake a nationwide livestock survey. There is currently a proposal to generate livestock population data by species, by seasons, and by regions. The proposal is part of the Second Livestock Development Project. Already, an aerial data collection strategy has been proposed. While the aerial survey can generate quick estimates, it will be subject to potentially serious reporting errors in thick vegetative areas. It will also not be able to generate the necessary sampling frame which will be needed for future livestock herd surveys. Hence, ground surveys headed by an able team leader (preferably a person with substantial experience in applied statistics, farm surveys, and animal science) will have to be implemented. The team leader will have the difficult task of balancing non-sampling errors (survey inaccuracies arising from faulty questionnaires, uncooperative respondents, and dishonest data collectors) and sampling errors within his allotted budget.

23. Together with the herd inventory surveys, information on major farm livestock inputs (e.g. feeds) needs to be collected. Animal carrying capacities of communal grazing lands are needed information in order to be able to formulate technological intervention involving leguminous and other types of fodder. Feed and labor use levels are also necessary data inputs in looking into the comparative advantage of specific livestock production relative to cropping activities for particular regions.

24. Livestock price data collection at the farm, wholesale, and retail levels must be improved and expanded relative to the existing system. For the price data set to be useful to entrepreneurs and other users, it must have a strong correspondence with quality, seasonal, locational, and intermediate processing variations. It has already been proposed to undertake

a pilot survey on livestock prices in a major wholesale market in which reported prices will be correlated with the above attributes of the animals sold. The aim of the proposed exercise is to test questionnaire designs and the operational capabilities of data collectors at field level. To avoid the danger of taxing scarce resources at the state level, only major livestock prices will be collected at carefully designated frequencies and only for specific livestock farmers and marketing agents. The list of respondents will have to be determined a priori through a logical and practical statistical framework jointly formulated by personnel of the state and FLD Market Information Unit.

25. Animal movements by hoof, rail, and road should still be collected by month and by states through secondary data sources to minimize costs. Such data will provide information on the major marketing channels used by livestock producers and traders. A tedious way to improve the data quality on animal movements will be to validate data reported at the state level through relevant intra-regional trade documents. An alternative data check will be for the FLD Market Information Unit team leader to visit major cattle trading posts on a regular basis to obtain field impressions. The animal movements data could also integrate reports on animals with severe and moderate health problems. This will alert FLD's animal health personnel on a potential geographic source of disease outbreak.

26. Feed quality and prices must be generated through regular periodic small surveys of feed producers at designated sites. Feed inputs account for a large portion of the cost of commercial poultry and pig producers. Information on feed inputs is essential to monitor the overall economic protection extended by the government to the poultry and pig industries.

27. A reasonable set of time series data on world prices and import levels of major livestock commodities and inputs must be maintained. This can be easily obtained from secondary sources such as the World Bank Commodity and Price Outlook reports and the FAO Trade Yearbooks. The world price data are usually used in computing the nominal rates of protection (NPC) - an index which indicates how government policies will affect domestic consumer and producer sectors.

28. Site specific animal recording schemes can be implemented at a modest scale to evaluate through time the effect of various types of economic and technological interventions administered to a particular animal specie. It is suggested that detailed to a particular animal specie. It is suggested that detailed animal recording schemes be limited to technological evaluations at the micro level because of the potential geometric increase in costs once the former is started on a grand scale. The animal records for a dairy project will include, for example, information on sire/progeny, breed, calving intervals, lactation period, daily milk output, mortality rates, feed intakes, etc.

29. Small household consumption surveys must be undertaken to provide some information on the response of consumers to changes in livestock prices, incomes, and quality. For urban consumers, panels can be established to minimize survey costs. A major issue related to this will be the degree of cooperation of the panel members and the representativeness of the panel.

30. The list of data categories covered is no doubt non-exhaustive. It is likely that if the proposed data collection activities are implemented efficiently, more demands will be imposed on FLD. It has been recommended that data functions of the Federal Livestock Department be part of the prime mandate of the four key specialist areas of a proposed FLD Planning Unit (i.e. planning, production, marketing, and information). This will assure that the data set collected are at least used for livestock policy planning within FLD. FLD should also discourage data collection projects which will generate information which will not be used productively. After all, data inputs are just like any other commodity - they have a price tag and that can be quite high considering the severe resource constraints of FLD.

Livestock data computerisation

(a) The role of computers

31. The scope of livestock information required and the various methods of obtaining that information represent a complex scenario. The volume of information being collected by the FLD at present is considerable and continuous. In undertaking this task by manual means it has not been possible for the information to be presented to its intended audience in a timely manner. Indeed, the information although collected by teams of enumerators and State officials throughout the country, will often remain for periods of time in files or on the desk of the person whose task it is to compile and eventually analyse the information. The reporting of partially complete data sets by manual means is prohibitive due to the burden of report preparation, printing and distribution. Innumerable institutions engaged in the production of livestock or livestock related statistics lag behind in the preparation of annual reports while the distribution of quarterly or monthly bulletins represents an almost impossible task. Indeed the actual preparation of such reports, relying as they do on the skills of available secretarial staff and the capabilities of the local printing firms, is in itself a laborious, thankless and time consuming task. Information is stored in its original form over long periods of time under inadequate conditions. The possibilities of damage or loss are considerable.

32. This section of the paper attempts to describe how, with the introduction of improved data processing techniques information can be stored, analysed and reported faster and easier. In examining the role of computers in livestock data collection, storage, editing, analysis and report writing a specific mode of data collection (principally a livestock survey) will be used to identify key tasks where computerization enjoys a comparative advantage.

33. In the following flowchart a typical survey (such as the livestock market price survey) has been considered to illustrate each of the tasks that were undertaken in designing, conducting and eventually reporting the results of the survey. It can be seen that computers can have an input in the design of the survey data forms to satisfy specific in-house requirements (examples are included at the end of this paper), data entry of the collected information into a database management package, statistical analysis of the data, organization of the analysed information in the form of tables and the preparation of the final report.

(b) Sample applications of computer systems in FLD's Market Information Unit (MIU)

(i) Hardware

34. Based on the livestock data priorities discussed earlier it is necessary for the Market Information Unit of FLD to introduce a number of microcomputer systems capable of storing and analysing the large volume of information that has been collected to date. In presenting these recommendations for the future computerization for livestock information processing and

reporting emphasis has been placed on the purchase of single user systems that are each self-contained. The advantage of such data processing systems are flexibility (it can be sited at the data source) and independence (a potential breakdown in one component will not affect the other units). Typically each system will comprise:

80286/80386 Processor
640-1024k bytes of Random Access Memory
40M bytes hard disk drive
Two serial interface ports
One parallel interface port
One floppy diskette drive
One cartridge tape spooler
One high resolution colour monitor
One high speed graphics printer.

Systems as described above are manufactured by a large number of firms throughout the world many of which are represented either through direct agencies or by small organizations in Nigeria.

35. It should be noted that the inclusion of both high resolution monitors and printers will facilitate both ease of use and provide a graphics capability to the systems. The ability for each of the units to produce reports that can be photocopied and distributed facilitates reporting for the entire system. The cartridge tape spooler will ensure that the user has an effective and easy means of ensuring data security against loss, theft, or accidental erasure.

36. In addition to this "standard" configuration provision has been made for the purchase of two sets of specialized equipment for report preparation and graphics. This equipment which includes document scanners and plotters will enable FLD to establish an in-house desk-top-publishing capability. The selected printers, for each unit, will be capable of printing in a variety of text size, font style and print and intensities. The laser scanner will enable the large volume of printed information presently stored by FLD in report form to be transferred automatically into the computer's selected word-processing package. The inclusion of a high quality (A2-size) multi-pen graph-plotter will enable the unit to perform its analysis of collected information and to produce the results of that analysis in a readily understandable form for reporting.

37. The computer industry is noted for its continual improvement in product capability matched by ever falling prices. It is thus recommended for organizations requiring a number of computer systems that the purchase of the systems should be planned over a period of time. In the case of FLD the recommended computer systems are to be purchased over a period of five years. This phasing of the introduction also enables the organization to conduct the high level of training required without the necessity of having units left idle.

38. In the case of FLD the planned purchase schedule is:

		<----Project Year--->					**Total**
		1	**2**	**3**	**4**	**5**	**Order**
A	Micro-computer Systems						
	IBM/PC or compatible (80386 based)	0	2	2	2	0	6
	IBM/PC or compatible (80286 based)	4	2	0	0	0	6
B	Peripherals						
	High Quality Printer	4	4	2	2	0	12
	Document Scanner	1	1	0	0	0	1
	Plotter	1	1	0	0	0	1

	Digitizer	1	0	0	0	0	1
C	Power Supply System						
	2000 va UPS	1	1	0	0	0	2
	20 Minute battery backup	1	1	0	0	0	2
	500 va UPS (with 10 minute battery)	2	0	2	0	0	4

(ii) Software

39. Applications software utilized for livestock data computerization falls into two categories: that purchased from software vendors; and that developed in-house. In general it should be noted that the majority of computer applications can be solved using purchased packages. However, when data volumes become excessive, the types of analysis become complex necessitating the use of low-level personnel in the entry or analysis of the data. Given these needs, there is a justification for developing specific packages to simplify programming routines used.

40. A sample of commercial softwares which can meet the data processing requirements of FLD are:

- The dBase III+ database management system which is used for data storage, editing and elementary reporting in both an interactive and programmed environment;

- The SPSS statistics system which enables a detailed statistical analysis of low volume data sets;

- "Smart" word processing system for report writing, form design and elementary desk-top-publishing;

- "Smart" spreadsheet system for the entry and calculation of tables and the preparation of graphics. This package is integrated with the one described above to facilitate the movement of tabular information for reporting;

- A desk-top-publishing package (possibly Aldus Pagemaker) will be used to assist in compiling final reports from narratives, tables and figures prepared from the word processing, spreadsheet and graphics packages;

- A time-series analysis package (e.g. SORITECH) will be used to examine the trends of livestock variables over time;

- The Turbo Pascal programming language development system has been obtained for in-house applications programming requirements for those systems for which the data handling requirements cannot be easily satisfied by dBase III.

41. Software packages that have been developed specifically to satisfy in-house data peculiarities include the following:

(1) The livestock movements information storage, analysis and reporting system obtains information from pre-printed livestock movement registers which have been distributed to State Chief Veterinary Officers (CVO's) throughout the Federation for use at their livestock control posts. This information is entered and stored in the computer systems through "user friendly" programs (those able to be operated by personnel with little computer training). Both summary and detail files of the information entered are maintained. Reports examining movements of livestock by checkpoint, species, origin, destination and period can be produced

by the system.

(2) The project planning and monitoring system is used internally to define livestock projects and to monitor their performance according to user-specified targets. Developed initially as a means of preparing detailed and summary cost tables for the fifth National Development Plan, this system is presently being expanded to be able to provide an ongoing, target oriented, project monitoring tool.

(3) The livestock and commodity price survey system enables the monitoring of both prices and quantitative information related to livestock and commodities essential to the livestock sub-sector. Much of the field work undertaken by personnel assigned to the MIU relates to the collection of livestock, livestock products and commodity prices. This information, which is collected on forms developed in-house, will now be entered on receipt directly into the system for immediate summary and analysis.

(iii) Maintenance and support

42. In order to ensure the continuing availability of the described systems to the livestock data collection effort attention must be paid to hardware maintenance and software/systems support.

43. In general the type of hardwares currently available are recognized to be robust and have a long service life. However, it is recommended to:

- use reputable local suppliers who have qualified engineering personnel available;

- train in-house personnel in the performance of routine maintenance activities;
- establish a planned cycle of system replacement.

Machines may have to be purchased in year 5 of the project to either enhance the system or replace defective components (see para 38).

44. The modular design of the micro-computer systems also provides the capability of purchasing a number of defined spares particularly those with moving parts.

(iv) Infrastructure

45. In order to operate effectively a computer system it is necessary to be able to provide the system with a reliable power supply that matches the manufacturer's specified voltage and frequency requirements within reasonable tolerance intervals. For the systems used for livestock data computerization this will be achieved by purchasing a battery - based uninterrupted power supply (UPS) system. Such systems, which can be allocated to a single user or shared between users, provide both a stabilizing effect on the incoming power supply in addition to assurance that the supply to the computer systems will not be discontinued. Batteries available for the UPS will provide an additional ten minutes of power after failure of the mains supply. This battery time will either enable users to close down the system in an orderly manner or provide sufficient time for a generator to be started. A three-phase 13.5 KVA has been installed to ensure that the full power requirements for the system including air-conditioning and normal office lighting are met.

(v) Training requirements

46. Computer systems and applications are generally unknown by personnel involved in the area of livestock data computerization. This is generally due to the fact that computers are

outside their particular areas of expertise, but also due to the rapidly changing computer environment. Knowledge gained with computing several years ago will not necessarily allow the user to understand the types of hardware and software being currently employed. For this reason a detailed schedule of training, covering all aspects of the described systems, has been developed and will be implemented among those staff who will be either directly or indirectly interacting with the systems.

Conclusion

47. The paper has attempted to define within the operational perspective of FLD the information needs and priorities for the livestock subsector. We also presented the various modes used in generating livestock statistics. In the latter part of the paper, the computer resources required in data storage, retrieval, editing, validation, and analysis were extensively discussed within the current environmental conditions in Nigeria. The final message we wish to impart is that a data set will command some value only if it is used efficiently. Livestock statistics, for example, are used to monitor and measure herd productivity; to trace the causes of a sluggish growth in domestic livestock supply; to analyse time-series trends of changes in outputs and prices; and to forecast the probable outcomes of certain given conditions. If livestock statistics are aimlessly generated, then scarce resources (e.g. skilled technician working hours) are also drained.