AGENCY FOR INTERNATIONAL DEVELOPMENT
WASHINGTON, D. C. 20523

BIBLIOGRAPHIC INPUT SHEET

FOR AID USE ONLY
BATCH #14

1. SUBJECT CLASSIFICATION

A. PRIMARY: Agriculture — AA00-0000-0000

B. SECONDARY: General

2. TITLE AND SUBTITLE
Improving the data and information base

3. AUTHOR(S)
Olayide,S.O.

4. DOCUMENT DATE: 1972

5. NUMBER OF PAGES: 15p.

6. ARC NUMBER: ARC

7. REFERENCE ORGANIZATION NAME AND ADDRESS
Mich.State

8. SUPPLEMENTARY NOTES *(Sponsoring Organization, Publishers, Availability)*
(Presented at Int.Sem.on the General Systems Analysis Approach for Agr.Sector Analysis and Planning,Airlie House,Va.,USA)

9. ABSTRACT

10. CONTROL NUMBER
PN-RAA-749

11. PRICE OF DOCUMENT

12. DESCRIPTORS
Data processing
Information systems
Simulation

13. PROJECT NUMBER

14. CONTRACT NUMBER
CSD-2975 Res.

15. TYPE OF DOCUMENT

AID 590-1 (4-74)

Not Proofread

IMPROVING THE DATA AND INFORMATION BASE

by

Dr. Samson Olajuwon Olayide

Discussion Paper Presented at the International Seminar
On "The General Systems Analysis Approach for
Agricultural Sector Analysis and Planning," 1-3 May 1972

The Airlie Foundation
Airlie House, Virginia 22186
U.S.A.

IMPROVING THE DATA AND INFORMATION BASE*

by

Samson Olajuwon Olayide**

I. Introduction

The simulation model of the agricultural sector, as of any other sectors of an economy, should be capable of serving as a successful mimic of the real world system if it is to adequately meet the needs of sectoral analysis, and planning component policies, progress and projects.[1/] This constraint, of necessity makes the model be very complicated and really involved. Usually a breakdown of such a complicated model shows that we have micro level simulation models which may deal with individual enterprises such as tobacco, broiler, and beef cattle feeding as well as aggregate all farm enterprises into farm-firm model. On the other hand, we have macro level simulation models of the agricultural sector which may deal with aggregation of farm-firm models into industry models at the provincial, state, regional, and national levels.

Simulation models of the agricultural sector either at the micro or macro levels require the full use of prior knowledge if a useful model is to be built. Also, there is the need for good experimental strategies and/or conjecture. Synthesizing and conjectural operations in themselves call for meaningfully lucid identification of the major variables that affect responses and thus stimulate/necessitate screening experimentation problems. A successful

---

* Paper presented at the International Seminar on "The General Systems Analysis Approach for Agricultural Sector Analysis and Planning," at Airlie Foundation, Airlie, Virginia 22186, U.S.A., 1-3 May 1972.

** Dr. S. Olajuwon Olayide is Senior Lecturer, Department of Agricultural Economics, University of Ibadan, Ibadan W/S, Nigeria.

identification of the major variables leads on to efforts to evolve terse empirical relationships that are designed to associate the leading variables with the responses.[2/]

All these are a necessary prelude to the uses for which any given simulation model of the agricultural sector are designed. In less developed countries (LDC's), the systems approach can serve four major uses. First, it is capable of helping us to fully map out and/or understand the managerial, organizational, administrative, socio-cultural and productive processes of the agricultural landscape with respect to structure and performance. This understanding is necessary if we are to successfully transform the farming economy from a primitive partly monetized peasant system to a modernized, fully monetized and corporate framework.[3/] Second, it facilitates an adaptive as well as engender the evolvement of development of normative, non-normative and prescriptive theories and knowledge which are of profound advantages in establishing priorities in agricultural modernization processes under economic development. Third, it provides us with a framework for the establishment of preconditions for optimization and/or maximization processes based on a set of alternatives.[4/] Fourth, it enables us to design, develop and evaluate policies, programs and projects for modernizing the agricultural sector so as to enable it to provide the food and fiber needs of right quality and quantity for the growing economy. The user dimensions enumerated above further attest the complexities of the systems-analysis approach in agricutural sector analysis, emphasize the intricate multidisciplinary interrelationships of the underlining the various socio-economic systems[5/] and stress the rigor of the multidisciplinary research and governmental policy framework required to build, operate and continually improve any simulation model of the agricultural sector.

II. Data Requirements of Simulation Models

To develop simulation models of the agricultural sector which are capable of successfully serving the essential uses mentioned above, there is the need for dependable and reliable data. Naylor, et al.[6/] list give prominent reasons that attest the importance of efficient data in successfully accomplishing computer simulation experiments either at micro-level, macro-level, or both. First, is the need for descriptive and quantitative information or data which deal with the particular system so as to facilitate a meaningful formulation of the problem. In this connection, it must be stressed that correct problem formulation is a necessary conditions for the derivation of meaningful solutions in empirical analysis.

Second, tenable hypotheses which can be usefully employed in formulating mathematical models of agricultural sector analysis depend for successful testing on the availability of data that have been reduced to a meaningful form. Third, good data may direct attention of modelers to needed improvements and refinements in existing mathematical models of agricultural sector systems and thus improve the simulated results. Fourth, the estimation of parameters of the operating characteristics relating the system's exogenous, status and endogenous variables can be made possible only when we have useful data. Such data are useful only in the sense that they have been reduced to final and dependably stable forms. Fifth, the validity of a simulation model of the agricultural sector as well as any other sectors of the economy can only be tested at the micro and macro levels if and only if we have "accurate" data.

The five uses of data in systems analytic models of the agricultural sector as well as any other sectors of the economy, strongly emphasize the important roles of data processing in building mathematic models for computer simulation. In data processing, as the job of obtaining useful data is

called, six important tasks must be carefully performed if we are to improve the data and information base for purposes of computer simulations.[7/] First, the task of collecting relevant data must be carefully organized since the variables involved in modelling have multidisciplinary interrelationships. The second task, which is data recording, has to be undertaken simultaneously with that of collecting. The two tasks normally prove very expensive and time-consuming, especially in LDC's since they involve identification, auditing, editing, coding, transcription and verification. These functions can only be successfully performed by trained competent and transparently honest staff which, to be frank and candid, may be non-existent in most cases. Third, the task of data conversions plays a very crucial role in the efficiency of data processing. Fourth is the task of data transmission which involves movement of information from one place to where they will be processed. Here shoddy treatment and/or handling will frequently lead to loss of much valuable information. Fifth, the task of data manipulation, calls for such meticulously undertaken functions as sorting, collating, merging, information retrieval and quantitatively logical operations in deriving parameters of the systems variables. Sixth, the task of outputing data is the end-result of the preceding five tasks. The stability and usability of the output data will depend on the efficiency and rigor with which the previous tasks have been undertaken. They will also depend on men, money and material resources allocated to theperformance of these tasks, especially in LDC's where the high rate of illiteracy calls for the use of cross-sectional information in obtaining relevant data.

III. Sources of Data for Simulated Sector Analysis

The general systems-analysis approach for agricultural sector analysis and planning basically places emphasis on a number of burning issues whether the model is micro or macro. Some of these burning issues may be itemized.

1. What is produced and what needs to be added to production, based on observed market prices and national requirements for growth?
2. How much should be produced of the various farm products, based on some "shadow prices" generated by the model or on observed market prices or on pencil and paper considerations of alternatives?
3. Which commodities cannot be produced internally and how could these be obtained based essentially on analysis of imports?
4. How much is paid for the goods obtained from outside the country for intersectoral development, based on the analysis of import-export trade, trade balances, and the levels of capital investments both foreign, public and private domestic capital formation?
5. What are ecological and edaphic components that delimit farm production and hence conducive to optimal intrastate, interstate and interregional specialization of production and optimization of internal trade, without seriously altering or dangerously destroying the ecosystems?
6. How can the agricultural sector be closely related to and carefully integrated with the other sectors as a means of generating employment opportunities, enhancing better incomes and improving general standards of living?
7. What is the resulting structure of GDP components, value added of the varous sectors, income, net foreign balance, etc. under economic growth?

To provide valid and useful answers to these burning issues in systems analysis model of the agricultural sector will make serious demands on data obtained from many sources. Some of these sources will be itemized without detailed discussion since our interest lies in how to improve these sources, especially in LDC's.

The first and readily available source is that of recorded time series data. In LDC's, this source is utterly inadequate except in the foreign trade sector of the economy where careful recording and organized customs operations have led to compilation of consistent data. Even here, the volatile nature of trade restrictions tend to bias the results obtained from the utilization of such data in quantitative analysis.[8/] All other time series data on outputs, inputs, and other variables of the agricultural sector of LDC's are mere guesstimates based on non-representative samples.

A second source consists of experimental data. Experiments in various fields and disciplines in the agriculture of many LDC's have been conducted in haphazard fashion with each experimenter or researcher doing what interests him in isolation of others in allied fields. Consequently, non-composite, imprecise and non-definitive enough for systems-analysis models of the sector.[9/]

Third, we make use of cross-sections data obtained from farm sample surveys, which constitute an acceptably convenient technique for measuring specific economic and social situations. Based essentially on past conditions, they are not too applicable in long-term projections of future trends, and may even tend to conceal the impact of dynamic components of the system. According to Jacoby, survey data are generally limited by their inflexibility in view of the fact that the survey design has to be fixed in advance and cannot be adjusted to new elements in the situation as they are discovered in the course of the investigation.[10/]

Fourth, there are engineering-observation data sources, which constitute precise quantifications of variables observed in performance processes. The time and motion study[11/] or work-study[12/] as well as carefully quantified

observations constitute a highly dependable and scientific approach for the evaluation of parameters.

The fifth source of data included judgemental-institutional-prescriptive information based on attitudinal evaluations. This source constitutes a hitherto unused information from the social sciences such as sociology, psychology, etc. Based essentially on examinations of factors and selections of those factors which load very highly on performance, behavior and adoption processes, they could constitute parametric valuations for setting bounds and/or examining changing reaction patterns on structure and performance.

Sixth source is the set of normative data obtained from optimization programs such as linear programming[13/] and its various extensions.[14/] Such data are largely constrained sets of prescriptive information which must be used with extreme caution not in building the model but in examining optimization components of generated simulation results.

A seventh source of data is the set of primary and secondary information obtainable from various documents and reports of governments as well as other organizations. Governmental sources of primary and secondary data may be obtained from local, provincial, state and federal institutions. Such information may also be obtained from international organizations such as the United Nations and FAO, accounting records of businesses and private research organizations. Such data must be subjected to judgemental and sensitivity tests when used in systems-analysis model.[15/]

The last source consists of generated data.[16,17/] The generated data should satisfy some criteria which include the fact that the numbers are (a) uniformly distributed, (b) statistically independent, (c) reproducible, and (d) non-repeating.[18/] The pseudo random numbers can be generated by the multiplicative method, the mixed method and the combination method.[19/]

Data generated by any of these methods mucst be subjected to such statistical tests as (a) mements test, (b) the run test, (c) X^2 test for uniform distribution and independence of sequences, (d) the serial test, and (e) tests of derived vector sequences.[20,21] Whatever method of testing generated data used must be guided by the particular application of the results as well as by the criterion of minimizing the bias in the data.[22,23]

The usability of the data collected from these and other sources will depend on the efficiency of performing the six functions of data processing discussed earlier. It will also depend on careful performance of sensitivity tests that seek to identify those parameters to which the simulation model is most sensitive or for which the model's behavior responds the most as parameter values are changed. Information culled from sensitivity tests can be very valuable in the tuning and validation processes. They also provide policy makers with the crucial means of identifying the most profitable areas of investments in research and data collection operations.[24]

IV. Data Situation in LDC's

In most LDC's as in Nigeria, the method and institutional arrangements for data processing tasks have left much to be desired. The Nigerian institutions for data processing tasks include the Federal Office of Statistics (F.O.S.) with its branches in various states capitals, the Federal and states research institutes or centers, universities and various foundations. An assessment of these known institutions and available state of sectoral information or data will only indicate the serious gaps in such a body of knowledge and/or information required for meaningful model building and sectoral planning. In cases where the F.O.S. and other researchers have conducted surveys, the sample spaces have been, to all intents and purposes, too small for any meaningful and definitive derivation of parametric coefficients required in simulation modelling.

Most of the data gathering studies undertaken to date have generally not been planned and carried out in the hollistic framework required for agricultural sector modelling. The information so collected are purely administrative and narrowly streamlined data that can only present a completely misleading and/or blurred caricature of real sectoral picture when used in simulation models. This contention stresses the need for a comprehensive and integrated enterprise research approach to information gathering based on ecological and edaphically delimited production regions of the rural landscape in any given country. Such a data collecting research should be carried out on all crops and other integrated rural enterprises and trades undertaken by sampled families.[25/] With careful design, implementation, collation, and coordination it is possible to gradually build up a data library which is subject to periodic additions of new information that are specially collected for use in simulation models.

In addition, there is the need for a multidisciplinary research designed to study the structure and conduct of the various enterprises as well as their problems in the agricultural sector.[26/] Such a research should be with a view to obtaining technical, engineering, judgemental, prescriptive, normative and attitudinal data required for improvising and modifying the simulation model of the sector as well as for incorporating micro and macro level policy issues designed to test a set of alternative projects.

All these call for new institutional arrangements for data collection, core modellers, interacting policy makers and considerable financial investments if our systems analysis approach to agricultural sector planning is to have any operational and/or practical usefulness.

V. Suggested Recommendations for Improvement

The systems analysis model of the Nigerian agricultural sector so far developed is fraught with some limitations which the authors rightly and explicitly recognized. These include (i) the problems of validation and verification, (ii) the need to minimize arbitrary assumptions as well as unnecessary rigidities, (iii) the need to improve the model's realism, applicability and operational usefulness (iv) the inability as well as the need to interact with policy makers and planners at the ministerial level as well as feed back mechanism for incorporating responses of farm-family units, (v) the omission of such vital developmental issues as irrigation, mechanization, income distribution, employment generation, etc., (vi) inadequate treatment of investment-disinvestment in durables, on-farm/off-farm capital generation, changes in technology institutional changes, and oversimplification of the behavioristic capabilities of the human agent, and (vii) inadequate research base for information inputs.

To make the model operational, the following recommendations on data gathering improvements are therefore in order.

1. A core modeller/simulator should be established within the Planning Division of the FMANR and based at Lagos.
2. This core would work with and be member of National Quantitative Planning Model Committee whose other members will include policy makers, consultants and State Planning Officers.
3. The core modellers working with the other committee members would
   a. Study the present model and work out needed expansions.
   b. Enumerate the type and magnitude of data required for implementation and revision.

c. Prepare research projects designed to collect such information for submission to the National Agricultural Research Council who will commission researchers to work on them.

d. Redesign, with full consultations, the questionaire, sampling procedures methods of surveys, etc., of the F.O.S. for obtaining micro level data.

e. Use the existing model to generate macro series and test the results, using regression analyses, spectral analysis and other non-parametric tests the macro-level data[27/] so obtained and compare results with guesstimates as well as aggregated sample data.

4. Study the structure of the F.O.S. and suggest new institutional arrangements with coordinated decentralized branches in the states for purposes of improved data gathering.

5. Explore the possibilities of processing all sources of data mentioned earlier through collecting institutions to build up data library required for modifying expanding and improving the model.

6. Establish meaningful training programs and incentives for staff participating in the whole exercise, as a means of creating the necessary manpower with the relevant skills.

VI. Summary and Conclusion

In this short discussion paper, I have attempted to examine very broadly the component structures of the systems-analysis model of an agricultural sector with respect to the leading variables and four major uses which it can serve in LDC's. I have also sketched the data requirements of such models with respect to the five important reasons why reliable data are needed as well as the six important tasks of data processing for computer simulation.

The issues involved in systems simulation model of the agricultural sector and the eight major sources of data were sketched as well as the usability of data from these sources. A brief account of data situations in LDC's with special references to Nigeria was followed by an examination of the limitations of the current Nigerian model as well as suggested recommendations for improving the data base both at the micro-macro level and for purposes of sensitivity analysis. These recommendations constitute necessary conditions for making the model dynamic[28/] and operationally validated.[29/] The sufficient condition is essentially the financial implications of implementing the recommendations. This I hope should stimulate dialogue and contractual arrangements between USAID and the Nigerian or any other interested government.

In conclusion, quantitative economic analyses and planning models designed to generate numerical results have been torn between two conflicting issues. The first may be expressed as: "Give me the data and I will search for and find the theory as well as the model that will fit them." The second is expressed by the statement: "Give me the theory and model and I will find data that will be employed in building it to generate usable results that constitute simplified images of the real world." The second statement is much more salutary and really compliments the outstanding job undertaken in the current systems-analysis model of the Nigerian agricultural sector. Rather than discredit it as garbage in and garbage out on the issue of not too accurate data and the current limitations, I am of the opinion that initial policy runs have attested its usefulness. It therefore throws a big challenge to modellers and user country to seek ways and means of improving the data base to continually improve and expand the model to operationalize it for planning, projection, and evaluation purposes required in modernizing the country's agricultural sector under national economic growth.[30/]

References

1. Guetzkow, H., (ed.) Simulation in Social Science, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

2. Fromm, G., and P. Tunbman., Policy Simulations with an Econometric Model, The Brookings Institution, Washington, D. C., 1968.

3. Meir, R. C., W. T. Newell, and H. L. Pazer, Simulation in Business and Economics, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969.

4. Orcutt, G. H., "Simulation of Economic Systems," American Economic Review, Vol. 50, Dec., 1960, pp. 843-907.

5. Orcutt, G. H., et al., Microanalysis of Socioeconomic Systems: A Simulation Study, Harper & Brothers, New York, 1961.

6. Naylor, T. H., et al., Computer Simulation Techniques, John Wiley and Sons, Inc., New York, 1966.

7. Tocher, K. D., The Art of Simulation, Van Nostrand, Inc., Princeton, N.J., 1963.

8. Olayide, S. O., "Some Estimates of Supply and Demand Elasticities for Selected Agricultural Commodities in Nigeria's Foreign Trade," Journal of Business and Social Studies, Vol. 1, No. 1, 1968, pp. 1-8.

9. Olayide, S. O., Towards Realistic Agricultural Economics Research in Nigeria. (mimeo) Paper presented at the International Seminar on "Employment and Agriculture in Developing Countries" held at Michigan State University's Kellogg Center, East Lansing, USA 14-17 December 1970, (revised) 15 pp.

10. Jacoby, Enrich H., Evaluation of Agrarian Structures and Agrarian Reform Programmes: A Methodological Survey, F.A.O., Rome 1966, p. 37.

11. Barnes, R. M., Motion and Time Study, John Wiley & Sons, Inc., New York, 1937, pp. 17-62 [285 pp.].

12. Ogunfowora, O., "Application of Work-Study Techniques In Improving Piggery Enterprise on the University Farm," Bulletin of Rural Economics and Sociology, Vol. 4, No. 1, 1969.

13. Dorfman, R., P. A. Samuelson, and R. M. Solow, Linear Programming and Economic Analysis, McGraw-Hill, New York, 1958, 512 pp.

14. Arrow, K. J. et al. (ed.) Studies in Linear and Nonlinear Programming, Stanford University Press, California, 1958, pp. 1-209.

15. Olayide, S. O., Quantitative Methods in Economic Planning: A Survey of Research Methodology and Tools. (mimeo) Technical Report, 1968, 23 pp.

16. Yagil, S. "Generation of Input Data for Simulation," IBM Systems Journal, Vol. II, 1963, pp. 288-296.

17. Chambers, R. P., "Random Number Generator on Digital Computers," IEEE Spectrum, Vol. IV, Feb., 1967, pp. 48-56.

18. Clark, C. E., "The Utility of Statistics of Random Numbers," Operations Research, Vol. VIII, March-April 1960, pp. 185-195.

19. Franklin, J. N., "Deterministic Simulation of Random Processes," Math. Comput., Vol. 18, 1963, pp. 28-59.

20. Gorenstern, S., "Testing a Random Number Generator," Communications of the ACM, Vol. 10, Feb., 1967, pp. 111-118.

21. MacLaren, M. D., and G. Marsaglia, "Univorm Random Number Generators," Journal of the ACM, Vol. 12, 1965, pp. 83-89.

22. Gelder, A. Van., "Some Results in Pseudo-Random Number Generation," Journal of the ACM, Vol. 14, Oct., 1967, pp. 785-792.

23. Peach, P., "Bias in Pseudo-Random Numbers," Journal of American Statistical Association, Vol. 56, 1961, pp. 610-618.

24. Manetsch, T. J., et al. A Generalized Simulation Approach to Agricultural Sector Analysis with Special Reference to Nigeria, Michigan State University, East Lansing, Nov. 30, 1971, pp. 167-172.

25. Olayide, S. O., Research on Rural Integrated Development, Food Production, and Employment Generation in the Guinea Savannah Zone of Nigeria, (mimeo) Paper presented at the Seminar on Rural Employment, Michigan State University, East Lansing, November, 1971, 26 pp.

26. Olayide, S. O., and Earl O. Heady, Improving and Implementing Agricultural Planning Machinery to Achieve National Goals in Nigeria, (mimeo) Document (to FMANR, Lagos, Nigeria) Iowa State University, Ames, December, 1971, 50 pp.

27. Naylor, T. H., Computer Simulation Experiments with Models of Economic Systems, John Wiley and Sons, Inc., New York, 1971, pp. 159-163.

28. Cohen, K. J., and R. M. Cyert, "Computer Models in Dynamic Economics," The Quarterly Journal of Economics, Vol. 75, Feb., 1961, pp. 112-127.

29. Naylor, T. H., and J. M. Finger, "Verification of Computer Simulation Models," Management Science, Vol. 14, Oct., 1967, pp. 92-101.

30. Theil, H., Applied Economic Forecasting, Rand McNally, Chicago, 1966.