

*Saldru copy*

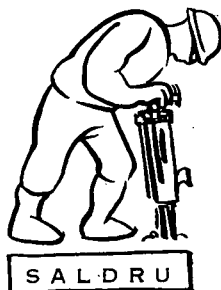
SALDRU FARM LABOUR CONFERENCE

SEPTEMBER 1976

Paper No. 38

Agricultural Statistics - A need for Reform

Delia Hendrie



Preliminary Draft : No portion of this paper  
may be quoted without permission of Saldru,  
School of Economics, University of Cape Town.

AGRICULTURAL STATISTICS - A NEED FOR REFORM

CONTENTS

- (1) Introduction
- (2) Sources of Agricultural Data
- (3) The Frequency of Publication of Agricultural Data
- (4) The Availability and Reliability of Agricultural Data
- (5) Problems of Definition
- (6) Data Gaps
- (7) A Suggested Method of Collecting Agricultural Data

## (1) INTRODUCTION

Statistics and their users have always been viewed with a certain amount of sceptism by the general public. I think here of aphorisms such as : 'There are lies, damned lies and statistics', and : 'He uses statistics as a drunken man uses a lamp-post - for support rather than illumination'. The former is levelled at the actual statistics, the latter at statisticians.<sup>(1)</sup> It is not necessary nonetheless, to emphasize the importance of statistics. They are an integral part of most reaserch, extension work and policy formulation and it is essential that they be both available and reliable. The statistician has at his disposal principles and methods developed for handling numerical data: the significance of any results he obtains however, depends implicitly on the quality of the data he is using, whilst the direction of his research is often determined by the availability of data.

Although a certain amount of data is obtained from individual efforts of private firms and researchers, the bulk of descriptive statistics pertaining to a country is collected by a central statistics office - in South Africa, the Department of Statistics. In the following discussion, focusing on the availability and reliability of South African agricultural statistics, it should be noted that these statistics constitute merely one section of the total data requirements of the economy as a whole. Demand for data has escalated in the past few decades and interest, whether popular or academic, is continually shifting from one field to another. This places an unenviable burden on the Department of Statistics, whose task in meeting these data needs is formidable.

## (2) Sources of Agricultural Data

Data relating to the agricultural sector is obtained from the agricultural census which is conducted annually. Questionnaires are sent to all Asian, 'coloured' and White farmers and submission of returns is compulsory. The results are published in three

---

(1) Throughout this paper, the term 'statistician' means anyone who uses statistics, and not merely professionally - qualified statisticians.

publications, namely:

- (1) Statistical News Release;
- (2) Agricultural Census; and
- (3) Abstract of Agricultural Statistics.

The Statistical News Release is released prior to the publication of the full census results and gives a 'preview' of a few selected principal statistics. The Agricultural Census contains data on farm holders and holdings, employment and earnings, machinery, production figures, livestock numbers and farm expenditure by magisterial district for the year in question; while the Abstract contains similar time series data for the country as a whole.

(3) Frequency of Publication of Agricultural Data

The most recent published Agricultural Census is that containing information for the year 1972/3. Apart from production estimates supplied by the Department of Agricultural Economics and Marketing and the limited statistics presented in the Statistical News Release (1973/4), little is known about what has been happening in the agricultural sector during the last three years. While an analysis of wage patterns in 1973 may be an interesting exercise, the results are of no current significance. Fragmentary evidence suggests that farm wages have increased considerably as a consequence of higher wages being offered in other sectors. African wages in the manufacturing sector have increased by 62% in money terms since 1973 while in the mining sector this increase has been 189%. What effect has this had on farm wages? Have increased labour costs led to a greater degree of mechanisation? What effect has this had on farm employment? Answers to these, and many other questions will probably only be available in 1978, by when, they again, will be of purely historical interest.

Table 1 : Time Lag in the Publication of Agricultural Census'

See overleaf.

Table 1 : Time Lag in the Publication of Agricultural Census'

Date of Census	Agricultural Census No.	Date of Publication
1964/5	39 - Part 1	September 1972
	- Part 2	October 1972
	- Part 3	December 1972
	- Part 4	November 1973
1970/1	44 - Part 1	December 1974
	- Part 2	December 1974
1971/2	45	October 1975
1972/3	46	August 1975

Source: Department of Statistics.

Note : These dates are approximate.

Table 1 reveals that there has been some acceleration in the publication of agricultural census', but there still remains a time lag of more than two years before census results are made available. This lag impairs the usefulness of the census results and severely hampers the work of agricultural statisticians.

#### (4) The Availability of Data

The data available on farm products and their prices normally suffices for most purposes, but information on other aspects of agriculture are far from adequate. One field for which this is particularly true is labour. Table 2 shows the statistical coverage in this area.

Table 2 : Statistical Coverage of Labour

See overleaf

Table 2 : Statistical Coverage of Labour

Subject	Coverage
Employment	Number by type of employee (regular, casual and domestic) and race as at 31 August.
Earnings	(i) Total cash wages paid during the twelve months ended 28 February or 30 June.  (ii) Total estimated value of rations and other goods and services supplied to employees during the twelve months ended 28 February or 30 June.
	Note: (i) and (ii) are presented by type of employee and race.

Source: Department of Statistics Agricultural Census No. 46  
Report No. 06-01-10.

Superficially coverage might seem fairly comprehensive, but a few examples will illustrate some shortcomings in these statistics.

Example 1 - Employment

There are a number of problems which arise on this issue when official statistics are used. The first arises from the lack of time series data on numbers employed in agriculture.

Table 3 : Summary of Employment in Agriculture 1960 - 1974

Year	Total No. of Farm Employees	No. of Regular Employees	No. of Casual Employees	No. of Domestic Employees
1960	1 505 227	756 397	591 882	156 948
1961	1 693 084	853 070	688 417	151 597
1962	1 794 864	872 784	769 789	152 291
1963	..	781 539	..	134 625
1964	1 253 661	756 581	365 557	131 523
1965	1 296 192	833 884	328 263	134 045
1966	..	..	..	..
1967	..	..	..	..
1968	1 640 295	..	..	..
1969	1 738 392	830 068	778 966	129 358
1970	..	..	..	..
1971	1 638 761	745 748	770 265	122 748
1972	1 505 791	736 236	652 991	116 564
1973	1 468 105	726 768	627 463	113 874
1974	1 454 183	712 892	631 976	109 315

- Sources: (1) Dept. of Statistics, Report on Agricultural and Pastoral Production 1970/1, Agricultural Census No. 44.
- (2) Dept. of Statistics, Report on Agricultural and Pastoral Production 1972/3, Agricultural Census No. 46.
- (3) Dept. of Statistics, Statistical News Release, 15 April 1976.

Notes : (1) Employment is shown as at June for the years 1960-1964, and thereafter is as at the 31 August.

Table 3 shows there is no employment data for the years 1966, 1967 and 1970 : in 1963 and 1968 there is no full breakdown of employment by type of employee; and the figures for 1964 and 1965 must be treated with caution because such large fluctuations seem unlikely. There are therefore, only eight years in the fifteen year period 1960 - 1974 when some indication of the numbers employed in each region by type of employee can be obtained. This discontinuity in time series presents difficulties for the statistician, especially for those interested in constructing econometric models.

What of the statistics that are available? The two major faults centre around loose definitions of the categories of employment and the statistics referring to only one date in the year.

Employment data being given at a fixed date prevents any meaningful conclusions being drawn about agriculture's reliance on labour. The demand for labour fluctuates enormously from one month to another. For certain areas this August count of employment reflects their 'minimum' labour requirements; for others, their 'normal' labour requirement. While knowledge of type of production in each area allows one to establish whether the August count is likely to be above or below the 'normal' employment for that area, it is not possible to gauge the extent of this difference.

A more frequent count of employment is essential if any meaningful estimate of the number of farm workers is to be obtained.

The census distinguishes between the following categories of employment:

- (a) Regular employees - that is, employees normally engaged in farm work in agriculture and forestry and sugarcane plantations and include managers, foremen and other regular employees;
- (b) Casual employees - that is, seasonal and occasional employees, excluding contractors and their employees;
- (c) Domestic servants - employees who are mainly or exclusively engaged in domestic work.

The definitions of regular and casual employees are not satisfactory. It is not clear for example, whether the migrant farm worker employed for nine months of the year should be classified as a regular or casual employee. He is not regular inasmuch as he is not in wage employment for the whole year, yet his labour input cannot be compared with that of a purely seasonal worker, employed for only two months of the year. There is need for a more meaningful distinction between these two categories of employment to be made. Casual employees should be subdivided into groups according to their length of employment, and some distinction made between full-time and part-time employees.

This last point is also relevant when considering the enumeration of domestic workers. A domestic employee who works for three mornings a week is not fully employed and should not be classified thus. That this is being done, is apparent if the census data on employment and earnings is used to arrive at an estimate of average monthly earnings of domestic employees. This wage in 1973 works out at R6,86 for African employees, and supposedly includes total cash earnings plus the imputed value of all payments in kind, excluding housing. It is difficult to believe that such appallingly low wages are being paid for full-time work - this figure is more likely the result of not distinguish the average hours worked by domestic employees.



### Example 2 - Renumeration

No data is given specifically on wage rates paid to farm employees although it should be possible to make crude estimates using annual earnings and employment data. As already mentioned these estimates are severely distorted by the poor quality of data on employment. To obtain these estimates, total annual payments during the year ended 28 February or 30 June must be divided by the August count of the numbers employed to obtain an average annual wage. It is obvious that, with no data on length of employment of casual employees, no estimate of wages paid to these workers can be obtained. Similar, but less serious, less difficulties arise in estimating wages paid to regular and domestic employees, since no distinction is made between full-time and part-time employees.

A further problem exists concerning the reliability of the imputed value of wages paid in kind. The farmer is required to estimate the value of rations and other goods and services supplied during the year. No guidance is given in the census questionnaire as to how the farmer should value these wages, with the result that these estimates appear arbitrary. This point is important since a large proportion of farm workers' wages are paid in kind, and should these be inaccurately reported total wages (cash and kind) will also be inaccurate.

### Example 3 - Data Gaps

Some important data gaps have already emerged in the above examples. Other aspects of labour about which nothing is known at present, but which do warrant inclusion in our agricultural statistics, include a breakdown of employees by age, sex, level of education, job category and number of dependants.

There are other faults in the data - both on labour as well as in other fields. These examples nevertheless, suffice to illustrate

that our agricultural statistics are in need of revision to make them more meaningful and useful. Some of the difficulties encountered result from problems of definition while others are data gaps which need to be filled if our knowledge of what is happening in South African agriculture is to be improved. The Commission of Enquiry into Agriculture (Du Plessis Commission) made reference to this lack of data and was compelled to undertake a great deal of primary research themselves, and to make use of unpublished material in the files of the Department of Agricultural Economics and Marketing. The fact that the Commission's findings are still so often quoted in agricultural studies - despite their being almost 10 years out of data - is sufficient evidence that serious data gaps do exist.

Table 4 gives some of the available labour statistics in the agricultural sector in the U.S.A. and shows that this request for more comprehensive data is not unreasonable.

Table 4 : Coverage of Labour Statistics in the Agricultural Sector in the U.S.A.

- |     |   |
|-----|---|
| (1) | <u>Farm Population</u> (published annually)   |
|     | (A) Number by region, race, age and sex.  |
|     | (B) Number by type of employment, i.e. agriculture or non-agriculture, self-employed, wage and salary worker, unpaid family worker, unemployed and not economically active. |
| (2) | <u>Farm Employment</u> (published annually)   |
|     | (A) Number by type of employee i.e. family, hired, etc.   |
|     | (B) Wage rates per day and per month by type of employee.   |
| (3) | <u>Farm Wage Workers</u> (published monthly)  |
|     | Number by region, race, age, sex, median earnings per day, migratory or non-migratory and length of time in farm wage work  |

Note: This is not complete and does not include data obtained from the quinquennial agricultural census. Despite this, it serves to illustrate the wider coverage of the U.S.A.'s labour statistics.

Source:

(5) Problems of Definition

Problems of definition and classification can best be tackled by a joint effort on the part of the users and collectors of agricultural statistics. Statisticians should focus greater attention on the definition of data requirements in relation to the purpose for which they need this data. The difficulties entailed in establishing 'working definitions' are often not appreciated. Without these definitions it is impossible to construct good questionnaires - and the questions in a survey are as important as the answers. It is essential for statisticians to assist in ironing out these theoretical and practical problems. It serves no purpose if the statistician merely criticizes the available data. Criticism of this kind must be heard by the collector of this data, and will only be heeded if close communication is established and maintained between the collector and user of data.

(6) Data Gaps

On this front, also, the statistician must express his data needs 'in terms of detailed content, frequency of collection, tabulation specifications and accuracy requirements.'<sup>(1)</sup> Once this has been established, the onus falls on the Department of Statistics to meet these additional data demands where they are justified.

It goes without saying that these demands must be met without antagonizing farmers. People resent being compelled to supply seemingly endless information about themselves and their activities to government departments, since this process is seen both as a waste of time and an invasion of privacy. One should remember that there is a two-way flow between statistics and people. The flow from statistics to people arises because decisions made on statistical evidence affect the lives of people - the reverse flow arises because the basic information on which these conclusions

---

(1) Houseman, E.: Sources for Econometric Models. Jnl. of Farm Economics 46.

are based is derived from people. The collector of statistics has therefore a dual responsibility - one to the user of his product; the other to the people who supply the raw material.

This means that the burden of form-filling must be kept to the minimum necessary to ensure the co-operation of the farmer. There is a sharp conflict between this consideration and the need to fill the present data gaps which prevail in our agricultural statistics, and this contradiction cannot be resolved if the present method of census enumeration is maintained as the sole source of agricultural data.

The 1974 report of the Department of Statistics mentioned that a proposal to make use of sample surveys to obtain agricultural data was rejected on the grounds that:

'.....the census is undertaken with the aid of the South African Police with little effort, and in these circumstances no need exists for a change in this regard. Should the South African Police be unable to assist with the census in future, consideration would be given to other statistical methods.' (1)

Although ease of execution of a census or survey is a necessary consideration, there appears to be no alternative but to replace the annual agricultural census with a less frequent census, additional information being obtained from quarterly sample surveys. The reasons for advocating this have been discussed above. To re-iterate:

- (a) The time lag before census reports are published detracts from their usefulness. This lag is partly attributable to the limited resources available to the Department of Statistics. The use of sample surveys would considerably lessen the bulk of data preparation and hence eliminate this lag.
- (b) There is need for knowledge on certain variables at more frequent intervals than the annual census allows. This

---

(1) Department of Statistics. Annual Report of the Statistics Council and the Secretary of Statistics for 1974. Pg. (IV).

need cannot be met without the use of sampling.

- (c) There is need to widen the coverage of agricultural data. To maintain a high quality while limiting the amount of form-filling by farmers, questions in questionnaires should be kept to a minimum. This apparent contradiction can only be resolved by the use of samples.

(7) Suggested Method of Collection of Agricultural Data

A less frequent census supplemented by quarterly sample surveys would allow both the frequency of the agricultural data to be increased and the coverage to be extended. Proper sampling permits generalisation from its findings within acceptable limits of doubt, and offers a practical compromise between certainty and expediency. Advantages of sampling are saving in time and cost, while disadvantages result from sampling error.

Sampling error can be considerably reduced by using a stratified sample in place of a simple sample. The population to be sampled is divided into strata which are as homogenous within themselves as possible, and a random sample is drawn from each stratum. In this way all strata are assured of representation in the sample. It is not necessary for the sampling fractions to be applied to each stratum-this can be determined by the particular aspects being investigated.

The population of sampling elements - in this case farming units - would be known from the most recent census. Farming units could be classified into strata according to their predominant type of production (or mix of products) and their size (based either on acreage or standard labour requirements). These strata could then be sampled using sampling fractions so as to ensure that sampling error is kept within acceptable limits. The frequency of the census would be determined by the rate of change in the factors by which the population had been classified into strata. For example, if the size of farming units changed rapidly, it would be necessary to conduct a biennial census; if this remained fairly constant a quinquennial census would suffice.

Each quarterly survey need not contain identical questions; each questionnaire could consist of a core of questions requiring answers at frequent intervals while the remaining questions could concentrate on one particular aspect of agriculture. For example, one quarterly survey could contain all core questions plus specific questions on cropping intentions; in this survey the sampling fraction in the cropping strata could be marginally increased at the expense of other strata. The number of times a farmer would be required to supply information would depend on the sampling fraction - provided this was less than 1 in 4, then theoretically any one farmer would not be sampled more than once a year. A problem that might arise is that random sampling might lead to certain farmers having to complete many more forms than others depending how often their number was chosen. This could be overcome with the use of a 'historic tape', which is a computer tape of the surveys in which each holding had participated during that year. Holdings which had been sampled more often than deemed acceptable in a certain period could be excluded from the sampling population for the next survey. This, of course, means that the sample would no longer be strictly random, which conflicts with statistical theory. In drawing up the design this conflict would have to be off-set against the desire to be 'fair' to all farmers.

It is however desirable to maintain a certain amount of controlled overlap of sampling elements in consecutive samples. This is known as 'rotational sampling' and allows for sensitive measurements of change in the core questions to be obtained; these sensitive measurements of change being as important as measurements of absolute levels.

The sampling method outlined above, together with research into problems of definition, would result in both the quality and quantity of our agricultural statistics being improved. This change from an annual census to a less frequent census supplemented

with quarterly sample surveys would need initial research done into the required frequency of the census, acceptable limits of sampling error, optimal strata classification, and suitable sampling fractions. Once this had been done, the execution of the new method would be routine and entail little or no extra cost - either in manpower resources or finance-to the Department.

The Department of Statistics is no doubt fully aware of the improvements which could and should be made to the South African agricultural statistics. However as Kendall so aptly says:

'In most of the pedestrian activities of the statistical world the problem is not what to do but to find someone to do it.'

Bibliography

- (1) Allen, R. (1970). On Official Statistics and Official Statisticians. Jnl. R. Statist. Soc. A. 133.
- (2) Britton, D.K., and Hunt, K.E. (1951). The Sources and Nature of Statistical Information in Special Fields of Statistics: The Agricultural Statistics of the U.K. Jnl. R. Statist. Soc. A. 114
- (3) Cox, K.K. (ed.). Readings in Market Research. (New York: Appleton-Century-Crofts, 1967).
- (4) Department of Statistics. Agricultural Census No. 46. Report No. 06-01-10.
- (5) Department of Statistics. Annual Report 1974.
- (6) Ferber, R. (ed.). Handbook of Marketing Research. (New York: McGraw-Hill, 1974)
- (7) Houseman, E.E. (1964). Sources of Data for Econometric Models. Jnl. of Farm Economics. 46.
- (8) Jones, H.J.M., Lawson, H.B., and Newman, D. (1973). Some Recent British Developments in Methodology. Jnl. R. Statist. Soc. A. 136
- (9) Nieuwoudt, W.L. (1973). Data Problems in Agricultural Economics Research. Agrekon. 46.
- (10) Orton, C.R. (1972). The Development of Stratified Sampling Methods for the Agricultural Census of England and Wales. Jnl. R. Statist. Soc. A. 135.