Measuring the size of the rural population in Malawi

A contribution to the 1999-2000 Starter Pack Evaluation Programme

Peter K. Wingfield Digby

Statistical consultant*

August 2000

The preparation of this report was carried out under DFID contract number MAL/31/2000. It formed part of Module 5 of the 1999/2000 starter pack evaluation programme. The programme was managed by the Statistical Services Centre (SSC) of the University of Reading, UK.

* P.O. Box 308 Chiang Mai, Chiang Mai 50000, Thailand. Tel. +66 53 895 136 Fax +66 53 895 271
e-mail: pwdigby@loxinfo.co.th

Contents

Page

Executive summary

Acronyms

1. Background information
   - 1.1 The starter pack programme 1
   - 1.2 The 1999/2000 evaluation 1
   - 1.3 Module 5 1
   - 1.4 The structure of this report 2

2. Review of data sources
   - 2.1 The population census 3
   - 2.2 SPLU data 4
   - 2.3 Issues of definition and classification 7

3. Methodology of the 'Ground Truth Investigation Study' 9

4. Reconciling population data from different sources
   - 4.1 Micro-level comparisons 11
   - 4.2 Macro-level comparisons 12

5. Starter pack estimates based on the GTIS
   - 5.1 Main indicators 14
   - 5.2 Deletions 16
   - 5.3 The influence of household size 17
   - 4.3 Other aspects of registration and distribution 19

6. Conclusions and recommendations 20

Annexes:

A. Sampling and grossing up factors 22
B. Questionnaire used in the GTIS 25
C. Documents consulted 27

Executive Summary

This report is part of an evaluation being carried out into the 1999/2000 starter pack programme. The evaluation is being managed by the Statistical Services Centre of the University of Reading, U.K.

There are five evaluation modules, each one looking at a different aspect of the starter pack programme. This module – Module 5 – is concerned with trying to reconcile different estimates of the number of rural households in Malawi available from a variety of sources. The focus is particularly on data on households and population available from the National Statistical Office (NSO) and on the data on farm families available from the Starter Pack Logistics Unit (SPLU). A main component of Module 5 was the Ground Truth Investigation Study (GTIS), which was designed to cover 60 villages spread out around Malawi. It was hoped that this study would yield independent estimates of the size of the rural population which could then be matched against the data from the other sources.

Initial work on Module 5 was undertaken by a team at Bunda College of Agriculture, but the main fieldwork was carried out by a team based at MoAI's Agricultural Development Division headquarters in Mzuzu. The team prepared detailed maps for 54 of the 60 villages selected for this study, and then carried out a village census, listing each household. Interviews were conducted with every household concerning the starter pack programme, and focus group discussions were held with representatives of each village. The survey was carried out to a very tight timetable, and the data checked and entered onto a computer. This consultant assisted with the initial design of the survey, and with the analysis of the data.

The results of the GTIS were surprising, in that they suggested that the size of the rural population was considerably larger than the figures obtained in the 1998 population census. Unfortunately it has not proved possible to obtain from NSO the detailed census counts for the 60 villages, which might have provided some crosscheck on the numbers for each village.The survey did however confirm a fairly close relationship between the number of rural 'households' and the number of 'farm families' as defined by the SPLU.

Besides giving the main results of the survey, this report contains a brief review of the population census and the SPLU database. The SPLU database provides a valuable source of data for planning purposes, but its value is limited at present, because of the different definitions and classification systems used by SPLU.

A particular problem surrounds the main concept of 'farm family'. It is suggested that it might be more appropriate to focus on a subset of households: those in which at least one member cultivates their own garden. Any database of information on these 'smallholder households' could also contain information about the various people within the household engaged in cultivation of their own gardens.

Acronyms

ADD	Agricultural Development Division
DFID	Department for International Development, UK
EPA	Extension Planning Area
FA	Field Assistant
FEWS	Famine Early Warning System
FFMIS	Farm Family Management Information System
GTIS	Ground Truth Investigation Study
MoAI	Ministry of Agriculture and Irrigation
NSO	National Statistical Office
ODI	Overseas Development Institute, UK
PRA	Participatory Rural Appraisal
RDP	Rural Development Project
SP	Starter Pack
SPIS	Starter Pack Initiative Scheme
SPLU	Starter Pack Logistics Unit
SSC	Statistical Services Centre, University of Reading, UK
TA	Traditional Authority

1. Background Information

1.1 The starter pack programme

The Starter Pack Initiative scheme (SP for short) was implemented by Malawi's Ministry of Agriculture and Irrigation (MoAI) during two cropping seasons (SP1 in 1998/99 and SP2 in 1999/2000) as a method of increasing maize production within the country. The scheme was supported by several donor agencies. The UK Department for International Development (DFID) was a major partner in the scheme.

The overall goal of the scheme was to contribute to food security and poverty alleviation amongst smallholder farmers. More specifically, it aimed to increase food productivity through the promotion of hybrid seed and fertilizer directed at the majority of resource-poor farmers. This was to be achieved by providing every smallholder household in Malawi with a pack containing sufficient seed and fertilizer for planting 0.1 hectares of land. The actual contents of the pack varied slightly in some parts of the country, depending on the agro-economic conditions.

1.2 The 1999/2000 evaluation

The evaluation exercise for SP2 was managed by the Statistical Services Centre (SSC) at the University of Reading in the UK.[1] There were five modules:

Module 1: An agronomic survey, carried out by the National Statistical Office (NSO);

Module 2: A study on the microeconomic impact of the scheme and people's willingness to pay for the scheme in future, carried out by the evaluation unit of MoAI's Agricultural Development Division in Mzuzu;

Module 3: A qualitative study on the likely impact of the scheme on sustainable livelihoods and poverty reduction, as well as a study of intra-household distribution and gender-related aspects of the scheme, carried out by a team from Chancellor College, University of Malawi;

Module 4: A study on sustainable agriculture and biodiversity, carried out by a consultant from the Overseas Development Institute (ODI) in the UK;

Module 5: A comparison of data from NSO with data available from the Starter Pack Logistics Unit (SPLU) in MoAI and from a 'ground truth investigation study' carried out under this module. The fifth module is the subject of this present report.

1.3 Module 5

Module 5 is concerned with trying to reconcile different estimates of the number of rural households available from a variety of sources. The SPLU had used Village Registers of heads of households as a basis for pack distribution. It was felt that this system might have introduced bias by encouraging the registration of more individuals as heads of household than would be found using any conventional household definition. It had been reported that the SPLU figures were far out of line with the corresponding information available from NSO. Module 5 is concerned with examining these discrepancies, and trying to come up with improved estimates of the appropriate numbers.

[1] A team from ODI carried out an evaluation of the scheme in the first season (SP1), and their report (Longley et al., 1999) is available from the DFID office in Lilongwe.

Some reconciliation could probably be achieved through careful desk research on the available data, by paying particular attention to the definitions and methodologies used in collecting each set of data. But a major component of Module 5 was to be a small ground truth investigation study (GTIS). The idea was to select a fairly small sample of villages, which would be visited and completely enumerated by teams which were independent from those who had carried out the original SP work. These village figures could then be used to provide new independent estimates of the rural population. The survey might also help to shed light on other aspects of the registration process.

Responsibility for Module 5 was initially assigned to a team from the Rural Development Department at Bunda College of Agriculture, University of Malawi. The author of this present report was also contracted to provide guidance, and he undertook a short consultancy visit to Malawi at the end of January 2000 for this purpose (Wingfield Digby, 2000). During this visit he offered suggestions on the approaches to be used for assessing data from different sources on rural households and population, and gave some advice on the methodology for the GTIS survey. Unfortunately the contractual arrangements with Bunda College proved unsatisfactory, and the project managers decided to reassign responsibility for the project after the initial work on pretesting the questionnaire had been completed.

Responsibility for Phase 2 (the GTIS) was therefore reassigned to a team from the Monitoring and Evaluation Unit of the MoAI's Agricultural Development Division (ADD) in Mzuzu. Responsibility for Phase 3, covering the desk study, the analysis of the results from the GTIS, and the preparation of a report, was reassigned to the author of this present report.[2]

1.4 The structure of this report

The main focus of this report is on estimating the number and geographical distribution of rural households and farm families in Malawi. Section 2 provides an introduction to two key data sources: the population census, carried out by NSO every 10 years or so; and the Farm Family Management Information System, developed by SPLU as a database for the starter pack scheme. The figures from these two sources are not directly comparable, because the population census uses the 'household' as its basic unit, whereas the SPLU database provides a record of 'farm families'. There are also difficulties in reconciling geographical classifications. Section 2 ends with a discussion of some of these definitional and classification problems.

Section 3 describes the methodology of the Ground Truth Investigation Study (GTIS). Section 4 compares the figures on estimated population obtained from the GTIS with data from other sources. Finally, Section 5 provides some other information at national and regional level on different aspects of the starter pack scheme, based on the results of the GTIS. Conclusions and recommendations are presented in Section 6.

[2] In addition, the technical inputs from Mr Carlos Barahona, a representative of the project managers, in various key phases of the project are gratefully acknowledged.

2. Review of Data Sources

2.1 The population census

Population censuses form a key source of statistical information. The 1998 Population and Housing Census was the fourth post-independence census in Malawi. The earlier ones had been conducted in 1966, 1977 and 1987.

Table 2.1 shows the number of households, population living in households, and average household size, for the three regions of Malawi, at the time of the 1987 and 1998 censuses. The 1998 figures were supplied by the NSO in July 2000 and are the latest available figures. The rural figures for 1998 were obtained by subtracting the urban figures for each region from the corresponding total figures for that region. The 1998 figures include the institutional population, but the size of this population is not known. Since the institutional population is likely to be located mainly in urban areas, its inclusion in the 1998 figures is likely to inflate slightly the average household size and intercensal growth rates of the population in urban areas, but to have little effect on the estimates for rural areas.

According to the census estimates, the total population of Malawi has increased from 7.9 million in 1987 to 9.9 million in 1998. Out of this total, the population living in rural areas has increased from 7.1 million in 1987 to 8.5 million in 1998. The rural population has therefore increases by 20 percent between the two censuses, representing an annual growth rate of 1.6 percent.

Table 2.1 Households, population, and average household size in urban and rural areas, by region, 1987 and 1998, and intercensal growth rates of household and population

	Households			Population			Av. household size		
	Total	Urban	Rural	Total	Urban	Rural	Malawi	Urban	Rural
1987 Census									
Malawi	1,859,572	198,939	1,660,633	7,928,792	828,928	7,099,864	4.3	4.2	4.3
Northern	190,283	18,891	171,392	905,452	87,150	818,302	4.8	4.6	4.8
Central	716,648	67,321	649,327	3,086,611	291,625	2,794,986	4.3	4.3	4.3
Southern	952,641	112,727	839,914	3,936,729	450,153	3,486,576	4.1	4.0	4.2
1998 Census									
Malawi	2,273,846	325,895	1,947,951	9,933,881	1,435,437	8,498,444	4.4	4.4	4.4
Northern	243,060	33,378	209,682	1,233,560	159,671	1,073,889	5.1	4.8	5.1
Central	908,942	125,637	783,305	4,066,347	567,604	3,498,743	4.5	4.5	4.5
Southern	1,121,844	166,880	954,964	4,633,974	708,162	3,925,812	4.1	4.2	4.1
Intercensal annual growth rate									
Malawi	1.8	4.5	1.5	2.0	5.0	1.6			
Northern	2.2	5.2	1.8	2.8	5.5	2.5			
Central	2.2	5.7	1.7	2.5	6.1	2.0			
Southern	1.5	3.6	1.2	1.5	4.1	1.1			

Note: In 1998 the figures for rural areas have been obtained by subtracting the urban figures from the total figures. The figures for the two censuses are not directly comparable because the 1998 figures include the institutional population. The 1998 average household size figures, and the estimated population growth rates, are therefore slightly inflated, particularly for urban areas .

Sources: 1987 Census: NSO, *Malawi Population and Housing Census 1987, vol. 1: Summary of final results,* Table 1.5
1998 Census: Household and population data supplied by NSO, July 2000

There has been a corresponding change in the number of households. For Malawi as a whole, the number of households increased from 1.86 to 2.27 million. Rural households increased from 1.66 to 1.95 million. There has been a very slight rise in average household size, for Malawi as a whole and for rural areas, from 4.3 to 4.4, but this apparent rise may be due to the inclusion of the institutional population in the 1998 figures. Obviously the way in which a household is defined for the purposes of the census will have an important effect on the resulting count of households, and this issue is discussed further in Section 2.3 below.

If the census results are taken at face value, they suggest that the three regions (northern, central and southern) have experienced different growth rates over the intercensal period, with the rural population growing much faster (2.5 percent per annum) in the northern region than in the central and southern regions (2.0 and 1.1 percent respectively). However, the population growth figures are strongly influenced by the effects of migration. Migration was an important factor in the 1987 census, which was conducted at a time when Malawi was sheltering a large number of refugees from Mozambique. These people will have been included in the 1987 population figures, but by the time of the 1998 census most of them had been repatriated. As a result, the annual intercensal growth rates for many of the districts which share a border with Mozambique are extremely low. The growth rates for several of the districts in the southern region, and for Ntcheu district in the central region, are particularly affected by this movement of refugees. The different growth rates for the three regions can in part be explained by the effects of this migration.

Some measure of the size of this migrant population in the 1987 census can be seen by looking at the number of people reported as having been born abroad. Back in the 1977 census, just over a quarter of a million people (256,215) in rural areas were born abroad. By the 1987 census the number in rural areas born abroad was approaching half a million (458,548). The corresponding figure for the 1998 census is not yet known, but it is likely that the figure will be very much lower than the 1987 figure. A figure of 400,000 has been mentioned as the possible size of the refugee population at the time of the 1987 census.

In addition to migration, other factors may well have played a part in reducing the overall growth in the population. In particular, the AIDS epidemic is likely to have had (and to continue to have) a major impact on the population projections. The impact is not only on mortality rates, but also on fertility rates, since many of the deaths from AIDS occur to people who have not completed their reproductive cycle.

2.2 SPLU data

In order to understand how the SP estimates came about, it is necessary to describe in some detail the process involved in generating the lists of farm families to whom SPs would be distributed.

For the first season 1998/99 (SP1), a registration of all rural households was carried out with the help of the Field Assistants (FAs) in MoAI. The FAs listed the names of all farm families in their sections, and these names were then transferred into distribution registers, each one containing 250 names. Later these registers were passed on to the appropriate Non-Governmental Organization (NGO) responsible for SP distribution in that area. In the planning stage the initial estimate of rural households had been 1.7 million, but this figure was later increased to 2.3 million. In the end the registration produced a final total of 2.86 million farm families. The issue of the definition of household, farm family, etc. is clearly important and is discussed below. The work of registration for SP1 was carried out to a very tight timetable, during August/September 1998, and the starter packs were all distributed by January 1999.

In April 1999 it was agreed that a database of farm families should be prepared from these registers. The SPLU therefore began to collect in the registers, and data entry commenced, using a special program written in Access. A large number of registers were successfully recovered and the data entered, but some registers could not be located. Eventually by the end of July 1999 the register contained about 2.2 million farm families, or just over three-quarters of the farm families covered in SP1.

Despite being incomplete, the database provided a valuable starting point for launching SP2. Information from the database was printed out, showing the farm families recorded for each section. Starting in June 1999, these printed registers were distributed to the appropriate FA as they became available, with a request that the listing be brought up to date. It was hoped that the final listing obtained from this updating exercise would be a valuable management tool for use in a range of agricultural projects, rather than being seen solely as a tool for SP distribution.

Unfortunately the distribution of the registers took place during the period running up to a national election, and it was almost inevitable that the process of updating would be influenced by those engaged in political activities. The evaluation of SP1 had indicated that 14 percent of those targeted to receive starter packs had actually received more than one pack, while only 4 percent had not received a pack at all. The SPLU team were therefore expecting that the new registration process would yield a total of rather less than the 2.86 million farm families registered under SP1. In fact, the returns from the field suggested a total projected final figure of as many as 3.7 million farm families. Strenuous efforts were made by SPLU to get their field staff to produce more realistic estimates. These attempts initially led to a new total of 3.2 million farm families. Only after further draconian measures were taken did it prove possible to reduce the total to about 2.9 million, so that it was in line with the figures from SP1. The final updating of the database was completed, not without difficulty, by November 1999.

As indicated above, the SP1 data is incomplete, and is therefore not worth reporting here. Only 2.21 million farm families are listed in the database, compared with the 2.86 million beneficiaries of starter packs in SP1. The SP2 data, on the other hand, is a very rich source of statistical information, despite any problems there might be over the use of the 'farm family' as the unit of analysis.

The database is on CD-ROM, and is fairly simple to use, but care needs to be taken, because of the way in which the database has been structured. The villages are listed according to Section, EPA, RDP and ADD. But Region and District have also been added to the database. Whilst this is an extremely useful addition, it has the problem that EPAs and Sections sometimes cut across the Regions and Districts. Where this happens, the EPA or Section is listed twice in the database, and there is a risk of double-counting if care is not taken to check for duplicates.

Table 2.2 provides a detailed SP2 listing of the number of farm families within villages, sections and EPAs, classified simultaneously by district and RDP, and by ADD and region. The total number of farm families shown in the database is 2,890,039. Although care has been taken to avoid EPA duplicates, it is possible that the listing might contain some duplication of sections and possibly even villages, though this will have no effect on the count of farm families.[3]

Despite some problems of classification, the database provides a valuable source of information, not only on 'farm families', but on the general structure of the agricultural sector. The CD-ROM could have many potential applications as a planning tool, for instance in determining the optimal allocation of field assistants around the country, bearing in mind the distribution of the population that depends on agriculture for its livelihood. Other information of an agricultural nature could presumably be added to the database if needed, so as to enhance its usefulness for particular users.

[3] In dealing with the SP2 listing, certain anomalies were noted, and these are reported here for illustrative purposes. In Kasungu ADD, there appears to be some problem with Chioshya EPA. It is shown as containing two sections (Kamwendo East with two villages and Zulu East with three) which should have been assigned to Mchinji district rather than to Kasungu. In Lilongwe ADD, there appears to be a problem with Chigonthi EPA, where one village (Chiweza) has been assigned to Lilongwe West RDP rather than Lilongwe East. In Salima ADD, one section (Thonje Centre) in Chinguluwe EPA has been assigned to Dowa district, whereas the rest have been put in Salima district. A particular classification problem concerns two EPAs (Chilipa with 67 villages and Nankumba with 40 villages), both in Bwanje Valley RDP which are shown as being in Mangochi district. They should therefore be classified as being in the Southern Region, as shown in Table 2.2, but in fact for the purposes of the GTIS survey it has been convenient to allocate the whole of Salima ADD to the Central Region. There also appears to be some problem with Chisi section in Karonga ADD, with some villages being assigned to Karonga RDP and others to Chitipa RDP.

Table 2.3 **Data from Starter Pack Logistics Unit 1999/00 (SP2)**

Region	ADD	RDP	District	EPAs	Sections	Villages	Families
MALAWI				**183**	**2,049**	**25,540**	**2,890,039**
Northern				**42**	**448**	**4,002**	**278,689**
	Karonga			**9**	**131**	**655**	**77,218**
		Chitipa	Chitipa	5	60	318	28,316
		Karonga	Karonga	4	71	337	48,902
	Mzuzu			**33**	**317**	**3,347**	**201,471**
		Rumphi North	Rumphi	6	50	617	26,259
			Mzimba	5	49	644	32,148
		Central Mzimba	Mzimba	8	88	842	55,145
		South Mzimba	Mzimba	6	72	817	40,082
		Nkhata Bay	Nkhata Bay	8	58	427	47,837
Central				**69**	**748**	**13,317**	**1,112,346**
	Kasungu			**25**	**250**	**4,981**	**360,878**
		Kasungu	Kasungu	7	54	1,718	111,404
		Mchinji	Kasungu	-	2	5	408
			Mchinji	6	55	1,248	97,542
		Dowa East	Dowa	3	34	468	44,803
		Dowa West	Dowa	5	63	912	69,844
		Ntchisi	Ntchisi	4	42	630	36,877
	Lilongwe			**32**	**376**	**6,369**	**541,761**
		Lilongwe East	Lilongwe	7	93	1,955	141,694
		Lilongwe West	Lilongwe	12	138	2,504	174,353
		Dedza Hills	Dedza	4	49	409	57,735
		Thiwi Lifidzi	Dedza	4	45	864	70,761
		Ntcheu	Ntcheu	5	51	637	97,218
	Salima*			**12**	**122**	**1,967**	**209,707**
		Salima	Salima	4	46	875	90,898
			Dowa	-	1	12	512
		Nkhotakota	Nkhotakota	4	44	710	67,829
		Bwanje Valley	Dedza	2	16	185	24,183
			Ntcheu	2	15	185	26,285
Southern				**72**	**853**	**8,221**	**1,499,004**
	Salima*	Bwanje Valley	Mangochi	2	14	101	17,577
	Machinga			**32**	**350**	**3,763**	**565,079**
		Mangochi	Mangochi	6	63	434	105,895
		Kawinga	Machinga	6	56	875	102,556
		Namwera	Mangochi	3	48	291	62,731
		Balaka	Balaka	6	61	556	88,097
		Zomba	Zomba	11	122	1,607	205,800
	Blantyre			**27**	**314**	**3,119**	**758,757**
		Phalombe	Phalombe	6	57	470	89,186
		Shire Highlands	Blantyre	4	63	615	143,833
			Chiradzulu	2	42	794	115,178
		Mulanje	Mulanje	5	46	546	186,959
		Mwanza	Mwanza	4	43	257	51,488
		Thyolo	Thyolo	6	63	437	172,113
	Shire Valley			**11**	**175**	**1,238**	**157,591**
		Chikwawa	Chikwawa	6	110	623	85,556
		Nsanje	Nsanje	5	65	615	72,035

* Note: A small part of Salima ADD (comprising two EPAs) falls within the Southern region.

Source: Farm Family Management Information System, 1999/2000 data (on CD-ROM)

2.5 Issues of definition and classification

Any attempt to compare data from MoAI and NSO is handicapped by the fact that the two agencies use different systems of classification. For our purposes these differences are particularly important in three key areas

- the use of different administrative boundaries
- different definitions of urban and rural
- different units of study (household vs. farm family)

Administrative boundaries

Data from NSO is collected and presented in terms of the standard official classifications, as laid down by the Office of the President and Cabinet (District Administration).

The country is divided into three regions (northern, central and southern). Each region is further divided into districts. Back at the time of the 1987 population census there were 24 districts in all: 5 districts in the northern region, 9 in the central region, and 10 in the southern region. By the time of the 1998 census two new districts had been created, both of the them in the southern region: Phalombe was created out of part of Mulanje, and Balaka was created out of part of Machinga. In the time since the 1998 census one further district has been created in the northern region, with the small island of Likoma, previously in Nkhata Bay district, becoming a district in its own right.[4] Below the district comes the Traditional Authority (TA) which is responsible for several villages; at the time of the 1998 census there were 112 TAs.

In the case of MoAI, the country is divided into Agricultural Development Divisions (ADD). Below the ADD comes the Rural Development Project (RDP). Below the RDP comes the Extension Planning Area (EPA). The classification used for selecting the sample for the ground truth investigation study (GTIS) reported on later in this report used a figure of 154 EPAs, but it appears that some EPAs may have been split, so that the total number of EPAs is now at least 175. Below the EPA comes the Section. There are usually 8 to 10 Sections in an EPA, and each Section has a Field Assistant assigned to it. Within each Section there are 8 to 10 blocks. Each block consists of 1 or 2 villages, though a large village is likely to be split between two blocks.

The difficulties in using the two systems of classification will be seen at first hand by anyone who uses the CD-ROM of farm families. There would be advantages, particularly in a planning context, if the two systems could be brought into line, though it would require some fairly major changes. Given the increasing political and administrative importance of the district as a unit of government, there would seem to be an urgent need for MoAI to redesign their classification system so that comparable statistics can be easily prepared on the basis of districts.

Urban *versus* **rural**

In the standard official classifications, as used for the census, there appear to be three different types of area that count as urban. First, there are the four 'cities', consisting of Lilongwe, Blantyre, Mzuzu, and Zomba Municipality. Then there are a few gazetted townships (an example is Liwonde Township in Machinga district). Finally there are the Bomas, which are the administrative centres of each district. Each district has a Boma, except for a few where a city or township takes its place.

This classification of areas as being urban is based on administrative criteria, without any specific allowance being made for the degree of services being provided in the 'urban' area. In the case of the new district of Likoma, for instance, which is being formed out of an island in Nkhata Bay district, the new district will contain a Boma, which does not exist at present. The creation of new districts therefore has a tendency to lead to the creation of more urban areas by administrative diktat. On the other hand, some urban areas will increase in size through incorporation of adjoining areas as they become more urbanized.

[4] Since the focus here is on 1998 census data, Likoma will be treated as part of Nkhata Bay, giving a total of 26 districts in all.

Rural areas, although still forming a very large and important part of Malawi, are nevertheless a residual. Rural areas, by definition, comprise those parts of Malawi that are not urban.

In the case of MoAI, agricultural information appears to relate to the areas covered by all the EPAs. Thus, all these areas are treated as rural, even though some of them (or at least parts of them) might have characteristics and facilities that would seem to make them more urban than rural.

Farm family *versus* household

The SPLU data and the census data use different units for classifying the population into groups. The starter pack scheme was originally targetted at all 'smallholder households' in Malawi. Field assistants were instructed to register all smallholder households in their section. They did this by going from household to household and registering the name of the households at their home. According to the instructions issued by SPLU, a smallholder and a household were defined as follows for the purposes of the SP project:

> Smallholder (SPLU): one who cultivates customary land within the section and derives his or her livelihood from smallholder agriculture
>
> Household (SPLU): the family or extended household which "eats from the same pot"

One obvious difficulty with both the definitions shown above is that they are circular. The name of the thing being defined (smallholder and household) is in each case included as part of the definition.

The register of names obtained by applying these definitions was known as the farm family register. It is often stated that a major problem with the SPLU data relates to this use of the farm family as the unit of classification. The final SPLU report on SP2 readily acknowledges the problem (page 8):

> " The major overall weakness of the data base at present is centred around the grey area of the definition of a 'farm family'. The inability to provide a firm direction on this has resulted in what many see as inflated numbers emerging in the reports. The lack of a logical link between the farm family figures produced by the standard reports and the figures emerging from the National Census and other demographic studies adds to the doubt cast on the validity of the agricultural figures."

In contrast, the population census and most other national surveys use the 'household' as the main unit of enumeration (NSO, 1998, p.7):

> Household (census): A household consists of one or more persons, related or unrelated, who make common provision for food and regularly take their food from the same pot and/or share the same grain store (*nkhokwe*) or pool incomes for the purpose of purchasing food."

Although this appears to be the definition that is frequently used, the form of words used here (and especially the use of the words 'or' and 'and' so many times) makes it difficult to interpret its true meaning. Nevertheless the definition does contain elements (especially the sentence up as far as the word 'pot') that would seem to be common to censuses and surveys in many other countries in Africa and elsewhere.

3. Methodology of the 'Ground Truth Investigation Study'

After the initial groundwork had been laid by the team from Bunda College of Agriculture, the main fieldwork for the ground truth investigation study (GTIS) was carried out at very short notice by a team from ADD, Mzuzu. The team had had considerable previous experience of this type of exercise. The sample design is described in Annex A. The survey covered 60 villages, spread across 30 EPAs. Three field teams were established, one for each region. Each team consisted of four enumerators and a supervisor. The enumerators were recruited through public advertisement. A three-day training course (including some practical fieldwork exercises) was held at Mponela Residential Training Centre, and the fieldwork commenced immediately after that. Since only 12 villages had been picked in the northern region, compared with 24 villages in the central region and the same number in the southern region, the northern team moved on to help the other teams once they had completed their villages. The questionnaire used for this study is reproduced in Annex B.

Unfortunately fieldwork could not be carried out in the time available in three EPAs. As a result no data has been collected for six villages, all of them in the southern region (see Table A.2 in Annex A for further details). The existence of this complete non-response is particularly unfortunate from a statistical point of view, since it may introduce bias into the results. This is because the reasons for the non-response for these villages could be directly related to one of the main subjects of this whole study: the issue of who did and did not receive starter packs. The team advised us that the main reason for not visiting four of these villages was the difficult topography existing in these areas (villages 55 and 56 in Livunzu EPA and villages 59 and 60 in Nyachilenda EPA). In the case of the other two villages (5 and 6 in Ntonda EPA), the reason given was that the team considered the villages to be too close to Blantyre City, and they felt this might lead to false information being collected. In any future exercise of this kind, every effort must be made to reach and cover all selected villages, whatever their topographical locality.

On arriving in each village, the team did a very detailed mapping of the village and all its structures, and these details were then entered on special maps prepared for each village. These maps have been retained by the team, and provide a valuable record. The information on the maps is supported by a detailed description for each village, obtained through focus group interviews with key informants. (Copies of these descriptions have already been made available to the project managers.) These descriptions cover the registration process for SP1 and SP2, and describe some of the problems encountered during the registration of beneficiaries and during the distribution of the packs. There are also useful comments about how villagers define households, who they think should receive starter packs, and about how the process of deregistration of potential beneficiaries was managed when it was discovered that an excessive number of beneficiaries had been registered for SP2. The main findings, based on this qualitative information from key village informants, have been summarized in a draft report prepared by the ADD team in Mzuzu (Msowoya, 2000).

In all, then, the survey covered 54 villages. The listing work produced a total of 6326 households.[5] In carrying out the GTIS, the survey team used a definition of household that corresponded closely to the NSO definition. Although the training manual omitted specific reference to taking 'food from the same pot', the concept seems to have been applied during the fieldwork. The draft report from the Mzuzu survey team (Msowoya, 2000) contains some comments from the focus group interviews on how villagers perceive the idea of a household.

Attempts were then made to interview every household, but information for some 72 households could not be collected in the time available. A variety of reasons were given for this non-response; these reasons included the absence of the household head, perhaps through sickness (going to hospital) or for work-related reasons (such as tobacco farmers who have taken their produce to market). In addition, some households were removed during data processing, perhaps because the interview was not complete.

[5] At a very late stage it was discovered that the listing shown for one village was incorrect. The number listed in village 45 should be 81 and not the 121 shown in Table A2. The effect of this correction would be to reduce the estimated number of households in Mzuzu ADD, in the northern region, and in Malawi as a whole by about 20,000. The corresponding reduction in population would be about 70,000.

The final computer analysis is based on 6134 records, but allowance is made in the estimation process for this non-response (and that from the six villages mentioned above). Data entry was carried out by clerical staff recruited specifically for the task. All information from the questionnaires has been entered directly onto the computer. This means that analysis is difficult for those questions which allowed for open responses. For instance, while information is available on whether field assistants, village headmen and members of parliament participated in the registration process, there were two additional spaces for recording details of others who may have been involved. Where an answer was given for these other categories, the response has been entered verbatim onto the computer. It is recommended that in future a special effort is made before data entry to produce a coding frame for these responses and get the clerical staff to enter the codes not words onto the computer. There is a similar problem with the question on main occupation, where the data entry staff have simply entered what they found on the questionnaire, without any attempt being made first to code the information into meaningful codes. Some cleaning of the data has been done prior to analysis, but in view of the time constraints it has not been possible to carry out extensive data cleaning.

Table 3.1 shows the distribution of the achieved sample by region and ADD, and the rest of the analysis in this report is based on these households and persons.

Table 3.1 Distribution of the achieved sample of rural households and persons, by region and ADD

	EPAs visited	Villages visited	Households	Persons
Malawi	27	54	6,134	25,281
Region				
Northern	6	12	1,340	5,846
Central	12	24	2,026	8,545
Southern	9	18	2,768	10,890
ADD				
Karonga	3	6	796	3,701
Mzuzu	3	6	544	2,145
Kasungu	4	8	359	1,578
Lilongwe	5	10	501	2,070
Salima	3	6	1,166	4,897
Machinga	4	8	795	3,113
Blantyre	4	8	1,500	5,793
Shire Valley	1	2	473	1,984

4. Reconciling Data from Different Sources

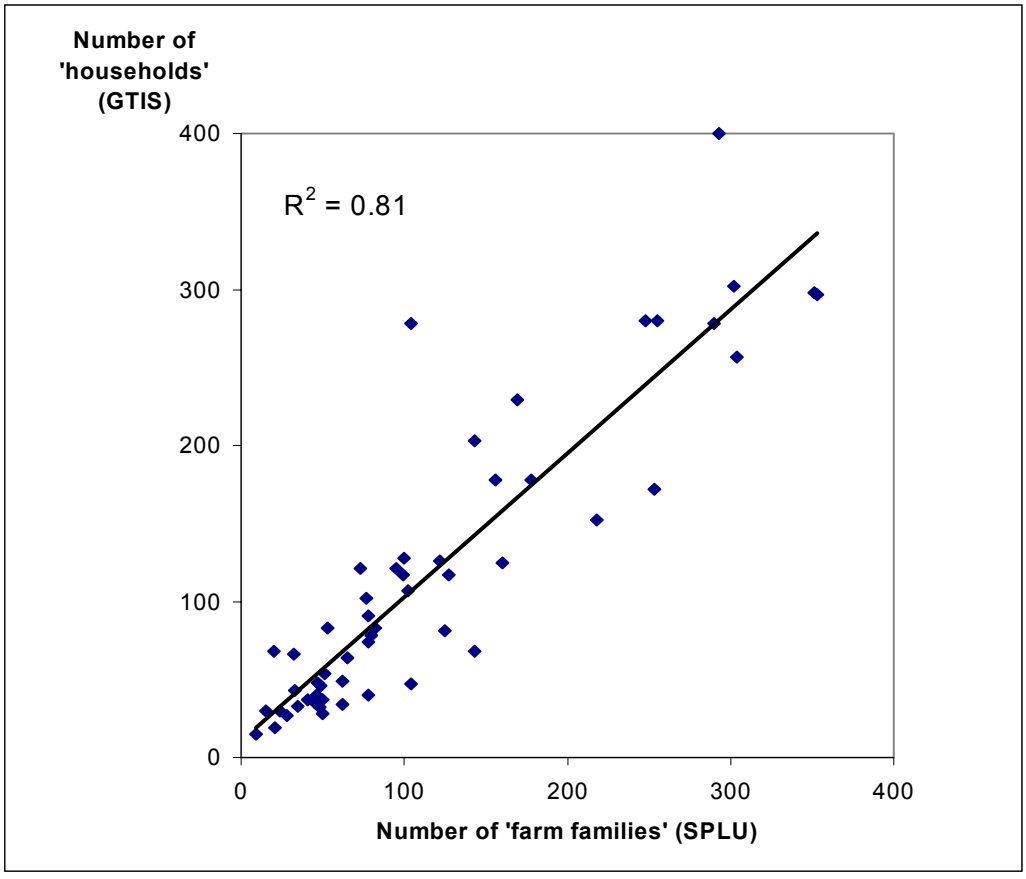
Comparison between the GTIS data and similar data from other sources can be done at two levels. At the micro level, we can compare the individual villages with village data from other sources. At the macro level, we can gross up our survey data to get national estimates of the rural population, which can then be compared with other sources.

4.1 Micro-level comparisons

Besides attempting a reconciliation between the data from the GTIS and the data available in the SPLU register, we had also hoped to be able to compare the results from the 60 villages with the corresponding data from the 1998 census. NSO was approached and agreed to provide the information, but unfortunately there has been a delay in getting the figures. Discussion here therefore focuses on the differences between the SPLU count of 'farm families' and the GTIS count of 'households'.

Figure 4.1 is a scatter diagram showing the relationship between the 'household' count (from the GTIS) and the 'farm family' count (from the SPLU). It can be seen that in general the linear relationship between the two counts is fairly close. The R^2 value is 0.81, indicating a fairly high degree of correlation (0.9) between the two counts. We can also see that the trend line is at about 45 degrees through the origin, suggesting a direct link between the two counts. If we wanted to estimate the number of households in a village, the number of farm families shown on the register would provide a good indicator.

Figure 4.1 Sample villages classified by number of 'households' in them and by number of 'farm families' registered for SP2



Two villages stand out as outliers. The one with 400 households is in Blantyre ADD, and the other with 278 households is in Salima ADD. It is probably not a coincidence that these two villages reported more deletions from the register than any others of the 54 villages covered in the GTIS.

4.2 Macro-level comparisons

As indicated in Annex A, special weights have been calculated, so that the sample figures can be grossed up to give national estimates for all rural areas. Inevitably with a small survey such as this, any estimates produced from the survey are subject to fairly wide sampling error fluctuations (though sampling errors have not been calculated for this survey). One problem with the weighting procedure is that some of the weights turn out to be extremely large. As mentioned in footnote 7 of Annex A, this happens in particular when two very small villages are selected in one EPA, and the data from these two villages are then used to estimate the EPA total.

Although the estimates are the result of the particular weighting procedure used, fairly similar results would be obtained if much simpler weighting systems were adopted. For instance, if we treated all villages as though they were selected directly from the ADD, and made no special allowances for the villages that were not covered in the survey, we would get national estimates of 2.8 million households and 11.6 million people. The estimate for the central region would drop by about 100,000 households and half a million people, but there would be a corresponding increase in the southern region. The estimate for the northern region would stay the same.

Table 4.1 shows how the household and population estimates from the GTIS compare with the 1998 census figures for rural areas. The SPLU 'farm family' figures are also shown for comparison, at both the regional and ADD level.

Table 4.1 Comparison of GTIS figures for households and persons with data from other sources

	Farm families	Households		Persons		Average household size	
	SPLU	GTIS	1998 Census (rural)	GTIS	1998 Census (rural)	GTIS	1988 Census (rural)
	Millions	Millions		Millions			
Malawi	2.89	2.78	1.95	11.52	8.50	4.1	4.4
Region							
Northern	0.28	0.34	0.21	1.41	1.07	4.2	5.1
Central	1.11	1.32	0.78	5.67	3.50	4.3	4.5
Southern	1.50	1.13	0.95	4.44	3.93	3.9	4.1
ADD	Thousands	Thousands		Thousands			
Karonga	77	89		431		4.9	
Mzuzu	201	249		975		3.9	
Kasungu	361	519		2291		4.4	
Lilongwe	542	476		2015		4.2	
Salima	210	325		1363		4.2	
Machinga	565	384		1591		4.1	
Blantyre	759	606		2273		3.8	
Shire Valley	158	138		578		4.2	

Table 4.1 also includes estimates of average household size for each region and ADD. On the basis of the survey results, the national mean household size for rural areas is 4.1, which is slightly lower than the figure of 4.4 found in the 1998 census and shown in Table 2.1. At the ADD level, the average household size for Karonga appears to be rather high (4.9), but it is known that households tend to be larger in the north of the country. Of more concern is the low average size for Mzuzu (3.9), which seems to be rather out of line, and may be a reflection of the fact that the entire sample of villages selected in that ADD were unfortunately situated close to each other (see footnote 6 in Annex A). Blantyre ADD also has a low average household size in the survey (3.9), but this is of less concern, since households are known to be smaller on average in the south of the country.

It was clear from Figure 4.1 that the GTIS and SPLU counts were likely to match up fairly well at the household level. However, Table 4.1 suggests that there are differences at the subnational level. The GTIS estimate of households in the southern region (1.13 million) is very much lower than the SPLU count of farm families (1.50 million), whereas in the other two regions the GTIS estimate is higher than the SPLU count. These figures are reflected also in the differences at the ADD level. Machinga and Blantyre have much higher SPLU counts of farm families than the corresponding GTIS estimates of households. In contrast, in the central region both Kasungu and Salima have much higher GTIS estimates than SPLU counts. These differences are partly a reflection of the different units being compared (farm families and households), but it needs to be emphasized also that the GTIS estimates are based on a relatively small sample size.

It can be seen that the GTIS produced very much larger estimates of households and population in Malawi than the counts provided by the 1998 census. On the basis of the GTIS, it is estimated that there are 11.5 million people in rural areas, living in 2.8 million households. The comparable figures from the census are 8.5 million people in 1.9 million households. The greatest disparity between the two data sources is in the central region, where the GTIS produces much higher estimates than those from the census.

These GTIS figures are indicative only. A much larger sample survey would be required to confirm whether there really has been substantial undercounting by the census of the population in rural areas.

5. Starter Pack Estimates Based on the GTIS

5.1 Main indicators

In addition to providing population estimates, the GTIS also collected some useful information on various aspects of starter pack registration and delivery. Table 5.1 presents some of the main results. It shows the estimated number of people in rural areas of Malawi who cultivate their own gardens, the number who considered that they had been registered for SP2, and the number who say they received SP2. The first column of the table shows the actual numbers of people registered in the SPLU database.

The survey figures need to be treated with some caution, for the reasons already explained. The figures are necessarily indicative only, since they are based on a relatively small sample. This fact has been emphasized by presenting all figures as rounded to the nearest 10,000. Even with this level of rounding, the figures for individual ADDs are probably subject to fairly large margins of error. The tables do, however, reveal some interesting broad trends.

Some 2.9 million people appear in the SPLU database for SP2. The survey produced an estimate of 3.7 million people who cultivate their gardens. The numbers cultivating a garden are broadly in line with the SPLU registration figures for the three ADDs in the southern region (Machinga, Blantyre and Shire Valley), but in the northern and central regions the survey produces estimates of numbers cultivating which are well in excess of the number registered. The reasons for these apparent discrepancies are unclear. For the central region it seems possible that the difference may be due to the very high weighting factors used for some villages, which would lead to large sampling errors. In addition, the survey team found many more 'households' in some of these small villages than the number of 'farm families' listed in the SPLU register (see listing in Annex Table A.2). The estimates for the numbers considering themselves registered match up quite well for the three southern ADDs, but are still well out of line in the central region. On the basis of the survey results, some 2.4 million people would acknowledge having received SP2, compared with the 2.9 million who actually did receive it. One interesting point in Table 5.1 is that the number of people in the southern region who considered themselves registered for SP2 was actually higher than the number of people who said they cultivated their own garden.

Table 5.1 Main indicators provided by GTIS, by region and ADD

	Registered farm families SP2 (from SPLU)	People who cultivated own garden 99/00	Consider themselves registered	Say they received SP2
Malawi	2,890,039	3,690,000	3,680,000	2,390,000
Region				
Northern	278,689	470,000	400,000	290,000
Central	1,129,923	1,910,000	1,810,000	1,010,000
Southern	1,481,427	1,370,000	1,470,000	1,090,000
ADD				
Karonga	77,218	130,000	100,000	70,000
Mzuzu	201,471	340,000	300,000	220,000
Kasungu	360,878	730,000	740,000	430,000
Lilongwe	541,761	760,000	670,000	420,000
Salima	227,284	420,000	400,000	170,000
Machinga	565,079	510,000	560,000	430,000
Blantyre	758,757	710,000	770,000	540,000
Shire Valley	157,591	150,000	140,000	130,000

Females are very much involved in cultivating gardens. Although the table does not show the separate figures for females, it is relevant to note that females account for just over half (52%) of those cultivating their own gardens. The proportion was particularly high in Blantyre ADD, where two-thirds (68%) of those cultivating their own gardens were female. In terms of being registered for SP2 and receiving SP2, women seem to have fared reasonably well, constituting 50% of all registrations and 48% of those receiving SP2.

Table 5.2 provides a more detailed breakdown of the survey data, in an attempt to identify the link between the SPLU concept of 'farm family' and the census/survey concept of 'household'. Out of the 2.4 million people who, according to the survey, reported that they had received SP2, some 1.7 million were heads of household, and the remaining 0.7 million were other members of the household. Looking first at the head of household group, we see that the great majority of those who received SP2 did cultivate a garden, but about 10 percent of them (160,000) said they did not cultivate a garden. A much larger group of heads of household were the 820,000 who said they had a garden but who did not receive SP2. This situation occurred in all parts of the country, but was most marked in Salima, where the majority of heads of household with a garden appear not to have received SP2.

Almost all rural women who are heads of household cultivate their own garden. Only 50,000 did not. Of the 690,000 female heads of household with a garden, some 490,000 (or just over 70%) appear to have received SP2. Female heads of household with gardens in the southern region did particularly well, with 80% receiving SP2.

As indicated in Table 5.2, household members who were not heads of household were also often the recipients of SP2, but their chances of getting SP2 were much lower than for heads of household. Some 65 percent of household heads with a garden received SP2. Amongst the other household members with a garden, the comparable figure was less than 40 percent.

Children under 15 have been excluded from Table 5.2, since very few of them cultivate gardens and hardly any of them reported receiving SP2.

Table 5.2 Heads of household and other adult family members, by region and ADD and by whether they cultivated garden and received SP2

	Heads of household				**Other household members aged 15+**			
	With a garden		Without a garden		With a garden		Without a garden	
	Whether received SP2		Whether received SP2		Whether received SP2		Whether received SP2	
	Yes	No	Yes	No	Yes	No	Yes	No
Malawi	1,500,000	820,000	160,000	290,000	510,000	860,000	190,000	2,000,000
Region								
Northern	230,000	80,000	10,000	10,000	40,000	110,000	10,000	280,000
Central	730,000	490,000	30,000	70,000	190,000	460,000	60,000	1,080,000
Southern	540,000	250,000	120,000	200,000	280,000	280,000	130,000	640,000
ADD								
Karonga	60,000	20,000	*	*	10,000	40,000	*	100,000
Mzuzu	170,000	60,000	*	10,000	30,000	70,000	10,000	180,000
Kasungu	330,000	160,000	*	20,000	70,000	160,000	20,000	510,000
Lilongwe	280,000	150,000	20,000	20,000	100,000	210,000	10,000	300,000
Salima	120,000	170,000	*	20,000	20,000	90,000	20,000	270,000
Machinga	210,000	100,000	30,000	40,000	100,000	90,000	90,000	190,000
Blantyre	240,000	130,000	80,000	150,000	170,000	170,000	40,000	310,000
Shire Valley	100,000	20,000	10,000	10,000	10,000	20,000	*	140,000

5.2 Deletions

A contentious issue during SP2 was the matter of deletions, which we referred to earlier. When it was discovered that the number of registered farm families was apparently much too high, SPLU tried to get the field staff to prune the lists. The GTIS provides an opportunity to measure the effects of this pruning exercise. Each household member was asked whether their name had been deleted from the register. If it had been, the household as a whole was asked why they thought this had happened. The responses to this latter question have been entered verbatim onto the computer, and cannot easily be analysed as they stand, though some subjective impressions can be obtained by scanning the list of responses. The first question about deletions is, however, a simple yes/no question and yields useful information.

The estimated number of deletions are shown in Table 5.3. Children under 15 have been excluded from the table, since in general they do not cultivate their own garden . Some children were initially registered for SP2 and subsequently deleted. The GTIS gives an estimate of 100,000 children under 15 who were subsequently deleted. Table 5.3 indicates that more than a million adults were deleted from the register. A quarter of a million people acknowledged that they did not cultivate their own garden, and their deletion seems reasonable, but the deletion of the remaining three-quarters of a million seems less justified, since they reported that they did cultivate a garden. Some parts of the country seemed to report higher rates of deletion than did other parts. The central region suffered the highest rate of deletions, with more than a quarter of those cultivating a garden finding themselves deleted. The situation was particularly critical in Salima ADD where over 40 percent of adults cultivating a garden found themselves deleted. In general, females do not seem to have fared any worse than their male counterparts in terms of deletions. Females accounted for 52% of deletions, equivalent to their share in the cultivation of gardens.

Households gave a wide variety of reasons for why names had been deleted. Groups who appear to have been deleted include: children, the second and subsequent people in a household, people who are unmarried, civil servants and teachers, and people who were absent or sick at the time. In other instances it was felt the deletion was due to bias or corruption on the part of a local official. (The use of a simple coding frame would have enabled those analysing the data to carry out a more scientific analysis of the responses to this question, and would have saved the considerable time needed for data entry of the verbatim responses.)

Table 5.3 Estimated deletions from the SP2 register, by region and ADD, and by whether cultivating own garden, for rural population aged 15+

	Cultivating own garden		Not cultivating own garden	
	No. deleted	% Deleted	No. deleted	% Deleted
Malawi	810,000	22%	260,000	10%
Region				
Northern	80,000	18%	20,000	6%
Central	530,000	28%	140,000	11%
Southern	200,000	15%	110,000	10%
ADD				
Karonga	30,000	20%	10,000	6%
Mzuzu	60,000	18%	10,000	6%
Kasungu	160,000	22%	80,000	14%
Lilongwe	200,000	27%	20,000	6%
Salima	170,000	41%	40,000	11%
Machinga	90,000	18%	40,000	12%
Blantyre	100,000	15%	60,000	10%
Shire Valley	10,000	6%	*	1%

5.3 The influence of household size

The presentation of results has so far been in terms of individuals and households, but with no attempt made to link the individuals to other persons in the same household. Table 5.4 shows the distribution of households according to the number of household members, and according to the number of household members involved in various aspects of the starter pack scheme.

It is estimated that there are only 130,000 rural households where no one in the household cultivates a garden. Two-thirds of all households (1.7 million) have one household member with a garden, while the remainder have two or more with a garden. As one would expect, larger households are more likely than smaller households to have at least two members cultivating a garden. Table 5.4 shows that there are as many as 880,000 households in Malawi in which at least two household members were registered for SP2. The distribution of households by number in the household registered for SP2 has a similar pattern to that for the number cultivating a garden, but with a slight downward displacement in the number of household members involved. Again, larger households are much more likely than smaller households to have at least two members registered for SP2.

Some 830,000 households appear to have experienced a deletion from the register. For 600,000 of these households only one household member was involved, but the remainder had at least two members deleted. Finally, Table 5.4 shows the distribution of households receiving SP2. There were 340,000 households in which at least two members received SP2.

Table 5.4. Distribution of households by household size, and by number of persons in the household who were cultivating their own garden, (2) were registered for SP2, (3) were deleted from the register, and (4) received SP2.

Figures in thousands

	Size of household							
	1	2	3	4	5	6	7+	Total
Total households	200	420	580	510	380	300	370	2760
Number in the household...								
...cultivating their own garden								
0	10	20	30	20	20	10	10	130
1	190	290	360	300	220	150	160	1670
2	-	100	170	170	130	120	160	860
3+	-	-	10	20	20	20	40	110
...registered for SP2								
0	30	60	30	30	20	10	10	190
1	180	260	380	310	230	160	180	1690
2	-	100	150	150	110	100	130	740
3+	-	-	20	20	10	20	5 0	140
...deleted from the register								
0	160	310	420	340	260	190	240	1920
1	40	80	110	130	80	80	80	600
2	-	20	40	30	30	20	40	180
3+	-	-	10	10	10	10	20	50
...receiving SP2								
0	70	150	140	150	90	70	80	750
1	130	230	370	290	240	180	210	1670
2	-	40	60	60	40	40	60	300
3+	-	-	*	10	*	*	20	40

Table 5.5 shows the distribution of households according to the number cultivating a garden and the number registered for SP2. The 1.8 million households shown on the leading diagonal in the table represent those where there was a good match between the number cultivating and the number registered. We say 'good' rather than 'perfect', for while one person in the household may cultivate a garden and one person in the household may be registered, we cannot tell from this table that it is the same person (though it would usually be reasonable to assume that it was). All those households located below the leading diagonal (roughly 400,000) are households where the number registered for SP2 exceeded the number stating that they cultivated a garden.

Table 5.5 Distribution of rural households in Malawi by number in the household cultivating their own garden and number in the household registered for SP2

Figures in thousands

	Number in household cultivating a garden				
	0	1	2	3+	Total
Number in household registered for SP2					
0	40	110	40	*	190
1	70	1260	330	20	1690
2	10	250	450	30	740
3+	*	40	40	60	140
Total	130	1670	860	110	2760

The final table, Table 5.6, shows the distribution of households according to the number of household members registered for SP2 and the number receiving SP2. Again those below the leading diagonal would represent those households where the number of household members receiving SP2 exceeded the number registered for SP2. It is encouraging to note that there are very few households that fit into this category.

Table 5.6 Distribution of rural households in Malawi by number in the household registered for SP2 and number in the household receiving SP2

Figures in thousands

	Number in the household registered for SP2				
	0	1	2	3+	Total
Number in household receiving SP2					
0	180	350	190	30	750
1	10	1340	280	50	1670
2	-	*	280	20	300
3+	*	-	*	30	40
Total	190	1690	740	140	2760

5.4 Other aspects of registration and distribution

The GTIS also collected some information on who was involved in the registration process. Households were asked whether field assistants, village headmen or members of parliament were involved in SP2. Sometimes they did not know whether these people had been involved. Nearly always the village headmen were involved. In all, 92% of households reported the involvement of village headmen in SP2. For only 5% of households was the village headman not involved, and in a further 4% of cases the household respondent did not know if this person had been involved. Field assistants were also usually involved in the registration exercise. Some 78% of households mentioned their involvement, 9% said they had not been involved, and 4% did not know. From the perspective of the householders, members of parliament seem to have played little part in the registration exercise. Only 2% of households mentioned their involvement.

Households were also asked whether they thought the registration exercise for SP2 had been satisfactory, and if not, why not. In general, some 90% of households expressed satisfaction. According to the results of the GTIS, Salima was the ADD with the highest rate of dissatisfaction with the registration exercise. As many as 100,000 households (31% of the total in Salima) expressed dissatisfaction. A wide variety of reasons were given, but these have been entered verbatim onto the computer without coding. Some of the reasons appear to mirror those given in earlier responses to the reasons for deletion. In some cases people reported that they did not receive starter packs because of shortages at the distribution centre.

6. Conclusions and Recommendations

The conclusions and recommendations from Module 5 can be summarized under these headings:

Estimates of the rural population

As described above, the GTIS has yielded estimates of the size of the rural population (in terms of both population and number of households) that are very much higher than the counts obtained in the 1998 census. The reasons for these differences are unclear. The GTIS estimates are, however, based on a relatively small sample, and it would therefore not be reasonable to reject the census figures on the basis of these figures alone. A more detailed and extensive sample survey would be required to determine whether there has been an undercount in rural areas.

It had been hoped to match up the GTIS for the sample villages with the corresponding information from the 1998 census. This would have provided a valuable cross check, but unfortunately it has not been possible to obtain the data in the time required. It is recommended that the NSO provide this information as soon as possible.

In view of the importance of migration in the 1987 census, when there were many refugees from Mozambique in the country, any calculations of growth rates based on simple analysis of changes between censuses are likely to be unreliable as tools for planning purposes. Planners should take account of the effect of these refugees in the census figures when making population projections.

The Ground Truth Investigation Study

Besides providing estimates of the size of the rural population, the GTIS also yielded some useful information on the effects of the starter pack campaign, and part of the information has been reported here. Carrying out the fieldwork and data entry for this survey has provided useful experience for the Mzuzu team. For any future surveys of this nature, the following points are recommended:

- Sampling: Great care is required in selecting the sample (the Mzuzu team were not involved in this operation). For a small survey like GTIS, where resources are limited, it might be worth selecting the two villages at a fixed interval (i.e. one-quarter and three-quarters) of the way through the list of villages in the EPA (ordered by size) rather than starting at a random point. Although the resulting sample is not strictly random, this would be one way to avoid getting a very large village or a very small village in the sample, which might have implications for fieldwork loads and for grossing up at the analysis stage. Care should also be taken to ensure that the EPAs are ordered in a meaningful geographic order prior to selection, so that a good spread of EPAs is obtained.

- Fieldwork: All selected villages must be visited. This is particularly important in the case of inaccessible villages, since their very inaccessibility may mean that they have different characteristics from other villages. For instance, because they are inaccessible, they may not have received starter packs.

- Coding: Prior to data entry, coding frames should be designed for all open questions, so as to avoid having to enter answers verbatim, and so as to facilitate data analysis. The appropriate codes could be chosen on the basis of an initial coding exercise involving, say, 100 questionnaires which have responses to the question on them.

- Data cleaning: More attention needs to be paid to data cleaning, to ensure that all households have been entered onto the computer, and that there are no duplicates, etc.

SPLU database

The SP2 database of starter pack beneficiaries provides a valuable resource base which could be used by MoAI for planning purposes. The value of the database, and of MoAI statistics generally, would be enhanced if their geographic classification systems could be brought into line with the systems of classification used by other agencies, notably NSO.

Farm families

It is generally agreed that the term 'farm family' in its present form is unsatisfactory as a unit of analysis The NSO concept of household may not be directly applicable to the starter pack registration scheme, because it does not contain any reference to the agricultural activities of the household, but a small modification could be made to the definition to make it more useful for starter pack purposes. In this regard the definition used in the National Sample Survey of Agriculture has much to commend it (NSO, 1984). There, a 'farming household' is defined as follows:

> Farming household (NSO): "A household is considered to be a farming household when at least one member of the household is operating one garden or more, or when the household head is economically active mainly in agriculture."

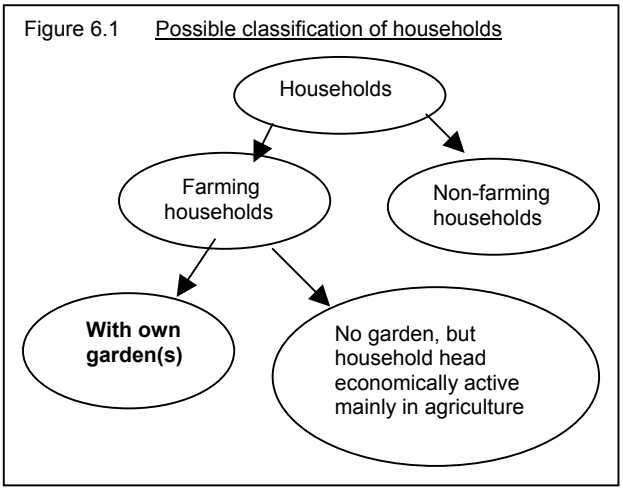
There are several interesting aspects to this definition. First, farming households as here defined form a subset of the larger group of all households, and statistics are therefore produced on a comparable basis. Second, the emphasis on the household means that it does not matter how many people in the household operate the garden, or whether there are several gardens being operated. All count as one farming household. The last part of the definition would presumably not be relevant to the starter pack scheme, since the scheme is concerned with helping those who operate their own gardens, rather than helping other people (such as those with only livestock) working in agriculture.

Figure 6.1 shows how this classification would work. Households are first identified. Three subgroups are then identified:

- those where at least one member of the household cultivates their own garden
- those without a garden, but where at least one household member is active in agriculture
- other households (i.e. non-farming)

Only the first group would count for starter pack purposes. They could be called 'smallholder households', to distinguish them from the larger group of 'farming households'. Identification of the other types of household would be helpful information, as it would enable a matching to be done with other available statistics.

Figure 6.1 Possible classification of households



It is possible, of course, that such stringent criteria as only one starter pack per smallholder household might not desirable. In that case one could identify those members of the household engaged in cultivating their own gardens, and give more than one starter pack to the smallholder household.

ANNEXES

A. Sampling and Grossing up Factors

The basic sample design

The initial planning for the survey was done in conjunction with a research team at Bunda College of Agriculture, supported by specialists from Chancellor College (in demography) and from the Centre for Social Research (in sampling).

Resources were available for a complete enumeration of about 60 villages. The aim of the survey would be to provide a reasonably reliable national estimate of the number of households in rural areas, and an estimate of potential starter pack beneficiaries. Estimates (though much less precise) would also be provided for each of the eight ADDs. The opportunity would also be taken to attempt a reconciliation of GTIS and SPLU data for each of the 60 villages, since this might help us to understand better the causes of the differences between the various estimates.

The SPLU list of farm families was used as the sampling frame for this survey. The EPAs were listed within their RDPs, which in turn were classified according to their ADD. Bearing in mind the spread of EPAs by ADD, and the number of farm families in each ADD, the team decided to allocate the sample to 30 EPAs, with two villages being visited in each selected EPA. The allocation of EPAs by ADD was as follows: Karonga 3, Mzuzu 3, Kasungu 4, Lilongwe 5, Salima 3, Machinga 4, Blantyre 5, and Shire Valley 3. The EPAs were selected by random sampling within each ADD, without using PPS.[6] Within each selected EPA, two villages were then selected at random.[7]

Grossing up the results to get national estimates

As part of the data processing work, it was essential to be able to attach a weight to each household record. This weight was applied to the household record and to the records of each individual within it, so that revised population estimates (and estimates on any other variable of interest) could be derived at the national and ADD level.

Karonga ADD is used as an example to show how this was done. There were two stages in the weighting process. First a grossing up factor was calculated, which could be applied to each record. In addition it was necessary to make adjustments to this grossing up factor, in order to make allowances for two kinds of non-response: non-response (if any) from individual households in those villages that were covered in the survey; and non-response arising from a failure to visit six of the selected villages.

---

[6] Although no stratification was made according to RDP, it had been planned originally that some implicit stratification would be obtained by listing the EPAs within each ADD according to their RDP, and then making a systematic selection of the required number of EPAs from across the list. However, in practice this was not done. The result is that the selected EPAs are sometimes not as dispersed within the ADD as one might wish. Mzuzu ADD is a good example of where this has happened, with all three selected EPAs touching each other.

[7] The original plan was to list the villages by size, largest to smallest, in terms of the number of SP beneficiaries. Two villages would then be selected systematically from the list, starting from a random point. In this way the two villages selected would consist of one larger village and one smaller village, since the first village was in the first half of the list, and the second village in the second half. This should have avoided the risk of selecting two very large villages or two very small villages, and so reduce the sampling error of resulting estimates. It would also have made the fieldwork easier, by ensuring that there was no risk of a field team spending a long time going to an EPA and then getting very few households. Unfortunately this procedure was not followed, with the result that sometimes both of the selected villages in an EPA are very small. Examples of this happening are villages 21 and 22, and villages 27 and 28, and the consequences can be seen in the very high weights needed for these villages, as indicated in the last column of Table A2.

The SP2 data provide the following information for Karonga and the selected EPAs and villages (Table A1). The GTIS listing information is also shown. (FF = Farm families, Hhlds = Households). Table A2 gives a detailed listing of the weights used in this survey, calculated in the way described here.

Table A1. Information used for grossing up data from selected villages in Karonga ADD

	SP2 FF	Selected villages	SP2 FF	GTIS Hhlds	SP2 FF	GTIS Hhlds
ADD Karonga	77,218					
EPA Chitipa North	7,305	11. Mkandi	78	74	129	128
		12. Mnyela	51	54		
EPA Kaporo North	9,697	13. Mwasulama	290	278	446	456
		14. Timoti	156	178		
EPA Karonga Centre	10,228	15. Malema 2	100	128	153	211
		16. Mwamatope	53	83		

The village information is first used to provide household estimates for the three selected EPAs. In using this ratio method, it is assumed that the ratio between farm families and households remains fairly steady across the EPA. Later, estimates of the number of farm families can easily be obtained by filtering on the answer to the appropriate survey question.

Estimated number of households in Chitipa North EPA = 128 x (7305 / 129)
Estimated number of households in Kaporo North EPA = 456 x (9697 / 446)
Estimated number of households in Karonga Centre EPA = 211 x (10228 / 153)

These revised estimates are now used to provide an estimate of the number of households in Karonga ADD:

$$\text{New estimate} = \frac{(128 \times 7305 / 129) + (456 \times 9697 / 446) + (211 \times 10228 / 153)}{7305 + 9697 + 10228} \times 77218$$

Adjustment (1):
To take account of non-responding households the household (and individual) records need to be inflated by an appropriate amount. In the case of Timoti village, for example, GTIS gave 178 households, but only 177 households were interviewed. The weighting factor is then made up of the following three elements, multiplied together:

(a) Non-response within village: 178 / 177
(b) Selection of two villages within EPA: 7305 / 446
(c) Selection of EPAs within ADD: 77218 / (7305 + 9697 + 10228)

Adjustment (2):
An adjustment has to be made for the six villages where no interviewing was done. It is considered that the characteristics of villages 5 and 6 are similar to villages 3 and 4. The missing data for 5 and 6 is then allowed for by adding in one more multiplier for villages 3 and 4: (23,965 + 46,777) / 23,965. Similarly the missing 55 and 56, and 59 and 60 are linked to villages 57 and 58 by using an additional first multiple for 57 and 58: (9,687 + 23,699 + 13,985) / 9,687. It is not ideal to have to link four of the non-response villages to the only two villages in that ADD (Shire Valley) for which there was a response. Fortunately the combined weight is not too large. A better alternative, if more time had been available, would have been to link villages 55 and 56 to villages 1 and 2, which they more closely resemble in terms of their characteristics. However, villages 1 and 2 are located in a different ADD (Blantyre), which would make the processing of the survey data rather more complicated. Simple inflation of the Blantyre figures would not be appropriate, since ADD is a key variable in the analysis.

Table A2. Weighting factors used for grossing up survey data

Village	ADD and Number of farm familes	SPLU size - EPA Farm Families		SP2 Village Hhlds listed	GTIS Hhlds interviewed	SPLU FF (2 Villages)	GTIS Hhlds	Weighting due to: Non-response	Villages to EPA	EPA to ADD	Villages missed	Total weight
1	Blantyre	31,569	293	400	392	646	697	1.020408	48.8684	5.8264		290.536
2	758,757	31,569	353	297	283			1.04947	48.8684	5.8264		298.811
3		23,965	253	172	170	331	212	1.011765	72.4018	5.8264	2.9519	1259.874
4		23,965	78	40	39			1.025641	72.4018	5.8264	2.9519	1277.153
5		46,777 *	92	*	*	276	*	*	169.4819	5.8264		0.000
6		46,777 *	184	*	*			*	169.4819	5.8264		0.000
7		12,460	160	125	123	225	189	1.01626	55.3778	5.8264		327.898
8		12,460	65	64	65			0.984615	55.3778	5.8264		317.688
9		15,457	255	280	277	473	432	1.01083	32.6786	5.8264		192.460
10		15,457	218	152	151			1.006623	32.6786	5.8264		191.659
11	Karonga	7,305	78	74	74	129	128	1	56.6279	2.8358		160.584
12	77,218	7,305	51	54	54			1	56.6279	2.8358		160.584
13		9,697	290	278	279	446	456	0.996416	21.7422	2.8358		61.435
14		9,697	156	178	177			1.00565	21.7422	2.8358		62.004
15		10,228	100	128	128	153	211	1	66.8497	2.8358		189.570
16		10,228	53	83	84			0.988095	66.8497	2.8358		187.313
17	Kasungu	9,858	46	40	40	79	83	1	124.7848	8.6754		1082.554
18	360,878	9,858	33	43	43			1	124.7848	8.6754		1082.554
19		10,614	49	46	46	111	95	1	95.6216	8.6754		829.553
20		10,614	62	49	48			1.020833	95.6216	8.6754		846.835
21		11,172	32	66	66	41	81	1	272.4878	8.6754		2363.932
22		11,172	9	15	15			1	272.4878	8.6754		2363.932
23		9,954	20	68	68	55	101	1	180.9818	8.6754		1570.084
24		9,954	35	33	33			1	180.9818	8.6754		1570.084
25	Lilongwe	12,242	99	117	117	140	154	1	87.4429	6.2610		547.483
26	541,761	12,242	41	37	37			1	87.4429	6.2610		547.483
27		24,926	15	30	30	63	62	1	395.6508	6.2610		2477.183
28		24,926	48	32	32			1	395.6508	6.2610		2477.183
29		24,341	104	47	47	186	130	1	130.8656	6.2610		819.354
30		24,341	82	83	83			1	130.8656	6.2610		819.354
31		12,652	125	81	79	149	111	1.025316	84.9128	6.2610		545.101
32		12,652	24	30	31			0.967742	84.9128	6.2610		514.492
33		12,368	28	27	26	49	46	1.038462	252.4082	6.2610		1641.118
34		12,368	21	19	19			1	252.4082	6.2610		1580.336
35	Machinga	3,690	304	257	249	354	285	1.032129	10.4237	7.8154		84.083
36	565,079	3,690	50	28	28			1	10.4237	7.8154		81.466
37		21,389	80	78	54	431	376	1.444444	49.6265	7.8154		560.231
38		21,389	351	298	292			1.020548	49.6265	7.8154		395.822
39		18,378	62	34	34	107	69	1	171.7570	7.8154		1342.355
40		18,378	45	35	33			1.060606	171.7570	7.8154		1423.710
41		28,846	143	68	68	193	105	1	149.4611	7.8154		1168.103
42		28,846	50	37	37			1	149.4611	7.8154		1168.103
43	Mzuzu	9,905	102	107	107	149	155	1	66.4765	8.5893		570.988
44	201,471	9,905	47	48	48			1	66.4765	8.5893		570.988
45		9,203	73	121	81	168	242	1.493827	54.7798	8.5893		702.877
46		9,203	95	121	121			1	54.7798	8.5893		470.521
47		4,348	77	102	98	155	193	1.040816	28.0516	8.5893		250.779
48		4,348	78	91	89			1.022472	28.0516	8.5893		246.359
49	Salima	11,166	127	117	112	375	397	1.044643	29.7760	4.8260		150.113
50	227,284	11,166	248	280	279			1.003584	29.7760	4.8260		144.213
51		15,993	104	278	242	247	481	1.14876	64.7490	4.8260		358.961
52		15,993	143	203	188			1.079787	64.7490	4.8260		337.408
53		19,937	122	126	121	291	355	1.041322	68.5120	4.8260		344.300
54		19,937	169	229	224			1.022321	68.5120	4.8260		338.017
55	Shire Valley	23,699 *	209	*	*	525	*	*	45.1410	3.3267		0.000
56	157,591	23,699 *	316	*	*			*	45.1410	3.3267		0.000
57		9,687	335	302	296	549	480	1.02027	17.6448	3.3267	4.8902	292.870
58		9,687	214	178	177			1.00565	17.6448	3.3267	4.8902	288.673
59		13,985 *	45	*		80	*	*	174.8125	3.3267		0.000
60		13,985 *	35	*				*	174.8125	3.3267		0.000

B. Questionnaire used in the GTIS

EVALUATION OF THE 1999/2000 STARTER PACK REGISTRATION PROGRAMME
MODULE 5

GROUND TRUTH INVESTIGATION

Household Number	**SP2**					Village Number		

Checked by		Date:	
Entered by		Date:	

Basic information

ADD _____ EPA _____
District _____ Section name _____
RDP _____ Village Name _____
Enumerator name _____ Date _____

Basic information about the respondent

Inheritance tradition — Patrilineal □1 Matrilineal □2

Type of house — Makeshift shelter □1 Round house □2 Square house □3

Walls — Mud □1 Unburnt brick □2 Burnt brick □3

Roof — Grass □1 Iron □2

The registration process

Were the following people involved in the registration for starter pack for the 1999/2000 season?

Field Assistant	Yes □1	No □2	Don't know □3
Village Headman	Yes □1	No □2	Don't know □3
Member of Parliament	Yes □1	No □2	Don't know □3
Other, specify			
Other, specify			

Do you think that the registration exercise for the starter pack this season was satisfactory? Yes □1 No □2

If No, Explain _____

Household Number	SP2				
Checked by		Date:			

VillageNumber		
Entered by		Date

Household membership

If a household has more than 10 members use a new form for those that didn't fit on this page. Make sure you complete the Household Number and Village name.

Name of HH Member	Sex 1=Male 2=Female	Relation to Head	Marital Status	Age	Main Occupation	Ask whether he/she cultivated own garden 98/99 season 1=Yes 2=No	Ask whether he/she cultivated own garden 99/2000 season 1=Yes 2=No	Ask whether he/she registered for starter pack 98/99 1=Yes 2=No	Ask whether he/she registered for starter pack 99/2000 1=Yes 2=No	Did he/she receive a starter pack in the 98/99 season 1=Yes 2=No	Did he/she receive a starter pack in the 99/2000 season 1=Yes 2=No	Was his/her name deleted from the register this year 1=Yes 2=No

Relation to head		
	1=Head	6=Grand parent
	2=Spouse	7=Brother/sister
	3=Son/daughter	8=Orphan related
	4=In-law	9=Orphan unrelated
	5=Grand child	10=Other

Marital Status	
	1=Married with spouse living in household
	2=Married with spouse living elsewhere
	3=Single (never married)
	4=Divorced/Separated
	5=Widow/Widower

If your name was deleted from the register, why do you think that happened?

_____

_____

C. Documents Consulted

C. Longley, J. Coulter and R. Thompson, *Malawi Rural Livelihoods Starter Pack Scheme 1998-9: Evaluation Report,* Overseas Development Institute, Aug 1999.

M.N.S. Msowoya, *Draft Report on Module 5: Ground Truth Investigation Study,* Mzuzu ADD, Jun 2000.

MoAI, *An Outline of the Starter Pack Initiative,* Jul (?) 1998 (briefing leaflet prepared for MoAI field staff)

NSO, *Report of Seminar on the Analysis of the 1977 Malawi Population Census and its Utilization in Development Planning,*(held at Kwacha International Conference Centre, 29 Oct – 2 Nov 1984), Mar 1986.

NSO, *National Sample Survey of Agriculture 1980/81, Vol. 1,* 1984

NSO, *Malawi Population and Housing Census 1987,* 1991.
Vol. I: Summary of Final Results; Vol. II: Population Characteristics; Vol. III: Economic Characteristics; Vol. IV: Migration; (Vol. V: Housing Characteristics), Vol. VI: Methodological Report; Vol. VII: Analytical Report

NSO, *1998 Population and Housing Census: Report of Preliminary Results,* Dec 1998.

SPLU, *Final report: Implementation of Starter Pack Scheme 1999,* Feb 2000.

Peter Wingfield Digby, *Statistical approaches to assessing data on rural households: Report on a visit to Lilongwe*, 29 January – 6 February 2000, DFID Contract no. MAL/03/2000

Annex 1 : A Guidance Note on Weighting in Estimation and on Interpretation of Results

The Module 5 report involves weighting each sample result differently according to the size of the village sampled and according to the population size of the area which that village serves to represent. Much greater emphasis is given to this theme in module 5 than elsewhere in the set of SP2 reports. Why is this?

As the Module 5 report by Peter Wingfield Digby explains, the estimate of a population total is usually more accurate - less biased - if that weighting is carefully undertaken.

If the weights attaching to different units in the sample - here village populations - are very different, that means that unusual features or errors in a small unit get amplified more in the scaling up to the overall total. They contribute more to the statistical variability - the sampling error - in the final answer.

We must also consider non-sampling error, which is due to other practical problems and issues. These tend to be rather important in difficult settings, such as rural Malawi, and where the pre-existing information for use in planning is limited, as was the case for much of SP2 evaluation work. In general, non-sampling errors are due to factors such as the following list, but the presence of these items here is for illustration and *not* to indicate that they are all very important in module 5:-

- frame definition error (e.g. if villages were missing from the SPLU register)
- fieldwork omissions (e.g. if unmotivated enumerators omitted some inaccessible households from a census count in a village, or did not register everyone in every household)
- fieldwork errors (e.g. if an enumerator coded some responses wrongly and the errors were not trapped through good supervision procedures)
- inaccurate or untruthful replies from respondents (e.g. if a household literally eats from the same pot *no-one* may be *able* to estimate how much a particular child ate at a given meal or e.g. if a few individuals stole a large batch of packs, they are unlikely to confess to this)
- definitional problems (e.g. those of household and farm family discussed in module 5)
- data entry errors (e.g. heavy-handedly entering ages as 220, 115 when they should be 20 and 15)

The estimation of population total size in a village, the main issue in Peter Wingfield Digby's main report, is likely to be pretty accurate in SP2 Module 5 - dwellings were carefully mapped and conscientiously enumerated.

The division of the villagers into "households" may have led to a larger number of smaller units than would be found by other processes, because of the distorting effects of issuing packs on the "farm family" basis. This is not of much significance to the grossing-up process whose input of primary data from module 5 is based on total population rather than households. The assumption is made that the module 5 villages were treated no differently in the course of SP2 registration than other villages nearby. Since their selection was unconnected with any information about their particular registration experience this is a natural assumption.

The particular variable concerned in this analysis, numbers of people, is quite unlikely to be misreported. People may not recall what they ate ten days ago, but they generally remember who is "normally" in their household!

When readers consider other data items in SP2 evaluation modules, they will often find that the figures presented represent a simple average across the individuals, households, villages or other units reported. In many cases this data concerns relatively difficult figures to elicit with an assurance of accuracy, e.g. 50 Kg bags of maize stored by someone who does not use 50 Kg bags, or e.g. “willingness to pay” (for a pack) values as cited by individuals. Is it wrong not to have weighted all of these figures?

Broadly speaking, where individual entries or blocks of data may be in error, very unequal weights are undesirable since there is a chance that bad data - from an unusual small village, say - may be given greatly increased prominence through receiving high weights, and may have undue influence on the results. This is much more likely to affect a value like “willingness to pay” than one such as household size. If there are substantial doubts, the analyst’s natural tendency is to “damp down” the effects of unequal weights, by some process which makes them “less unequal”. The most extreme case of such damping is to weight all responses equally. Then in a simple average, no individual has a bigger weight than any other. The results from a “damped-down” estimate of any sort will be less than perfectly efficient - some information is being reduced in “value”, but they are not likely to be so badly biased by non-sampling error.

A critical numerical value is one whose size materially affects what is *done* as a result of the finding. For example, the size of Malawi’s rural population could be critical if a donor gave credence to Module 5’s 54-village study, and related the size of a safety net provision to the estimate. If results are for a critical numerical value it is important to look carefully at the effects of weighting e.g. the figures in the module 5 main report have been recalculated in several ways, which produce very nearly the same overall population estimate.

If the results are *not* critical numerical values, but rather parts of a broader picture, their general scale and relativities are more important than their exact size. Often such results contribute to deciding between two (or a few) decision options, or to a qualitative message. If this is the case, we are concerned with the magnitude, rather than the exact numerical value, of a figure. What is important may be as crude as high/low or yes/no results. Usually such results are little affected by the selection of weighting scheme. The broad general interpretation of results from SP2 evaluation is in most cases unlikely to be seriously prejudiced by a lack of weights.

Much more important for the user of these reports and of specific information in them, is to interpret *carefully* the exact definition of a number cited! For example, are these numbers of individuals or of households? For example, are these numbers percentages of the whole sample or percentages of a restricted subset of the sample? In a number of instances, the numbers presented differ from place to place in the set of results, because rather similar-sounding results and interpretations are on different bases.

In almost all cases SP2 evaluation modules used random samples, so that biases due to personal preference or ease of access have been kept to a minimum. Reported qualitative findings from semi-structured discussions have been reported by small groups of writers, so are quite consistent across sites - indeed unusually so - but may contain some undercurrents of the authors’ preconceptions, rather less clearly identifiable than the assumptions underlying the more formal studies.

Ian Wilson
August 2000

Module 5: Annex 2 - The Problem of the Farm Family Basis of Benefit Distribution

The following is a personal note from the author to amplify a verbal comment made at the SP2 Evaluation Results Presentation of 24 August 2000.

The work done on a qualitative footing for Module 5 illustrates two difficulties that are compounded, and muddled together, in the results. Firstly, the process of defining who the recipients should be was tackled in an 'obvious' way, but not one that stood any chance of being acclaimed as reasonable.

Secondly, corrupt, unreasonable, unfair and wrong actions were taken by individuals along the SP registration and delivery chain. Of course the second difficulty is predictable and indeed inevitable. The qualitative study done under Module 5 in 54 villages has led to a set of narrative summaries written under fairly standardised headings by members of the fieldwork team. In an illuminating way these provide a series of instances of opportunistic thefts by individuals, and more extensive and systematic misappropriation and cheating by people with higher levels of access to the inputs. These are included in the CD in the form of scans which can be viewed, for 52 of the 54 villages covered by Module 5. No narratives were preserved from fieldwork in villages 51 and 53.

Of course, too, there are some cases where good has come out of evil e.g. the ousting of very corrupt village headmen as was reported in Village 1 in Thyolo RDP. Perhaps the optimistic might even claim that Starter Pack has contributed to the strengthening of popular democracy in civil society in some cases like this.

The concern of this note is with the first issue above, that of the farm family or household definition. Quite reasonably, Wingfield-Digby (2000) takes it as given that some form of household definition is inevitable, and discusses the disparity between the relatively conventional one used by NSO, and that used by SPLU. This discussion is not challenged here and is not repeated below: it is very clearly stated by Peter Wingfield-Digby and is worth reading.

The question here is: should we be determining who qualifies for Starter Pack using a farm family or household-based definition in the first place? There are several discussion points in this context, and these have to take full account of circumstances.

The development of good governance depends on general acceptance that the *institutions* of the state and society operate reasonably, even though everyone is familiar with instances of corrupt practice by *individuals*. There is very considerable evidence from the qualitative reporting of Module 5 that Starter Pack does *not* fit in with villagers' concept of the reasonable and to that extent contributes negatively to the perception of good governance.

There are frequently repeated comments to the effect that the qualifying criterion of household headship is unrealistic, unreasonable or unfair. Clearly in a single country where matrilineal & patrilineal, and polygamous & monogamous, traditions co-exist a household-based definition will be understood, interpreted and implemented in various incompatible ways. Clearly households are frequently not of idealised forms in a country where the social fabric is strained almost to breaking point by population pressure and the distorting effects of HIV-related mortality.

The household basis of NSO *population* statistics is not an issue in the same sense. That the people are counted at all, and according to what procedures, is a much more remote matter, whose consequences are little-understood even in highly-educated societies. This

issue of the basic unit matters to Starter Pack because registration has economically-important consequences which perfectly clearly link registration to the outcome of the exercise.

The rich qualitative evidence from module 5 sustains the assertion that "the people want" Starter Packs (and presumably other benefits of this general sort) for individuals, and at least in part this reflects a concern about "fairness" which makes it a governance issue.

Of course there are other messages not discussed herein. For instance, quite naturally "the people want" (they wish they got) other things, such as higher-value packs, in terms of quantity and quality, which were never on the cards. There are frequent messages concerning the aged, infirm and dependent, and a not unreasonable interpretation of SP as a welfare benefit that "should" be provided to them or their carers, though the narratives underscore the obvious point that the weakest are the most readily dispossessed of this form of benefit, not least because of its distribution system.

How would things be *if there were* an SP product delivered to qualifying individuals rather than households?

- A fixed total value of scheme would be divided into smaller units on a per-head basis than on a per-family basis.
- It appears that rural Malawians would readily understand the concept that an adult individual *who cultivated a garden* should qualify for an agricultural input.
- The interpretation (misinterpretation?) of SP as a "dependency culture" welfare handout - whether by recipients or observers amongst the donor community - would be reduced by stressing the italicised phrase in the qualifying condition.
- The process would not rely on the invasive and arguable conclusion (by one of a widely-varied and less than consistent range of decision-makers) as to whether a set of relationships currently constituted more or less than a family bond.
- The process of registration would be administratively easier to carry out and this would do something to offset the effect of dealing with more people individually.
- It would be clear that this was more nearly fair in terms of household size variation, e.g. in polygamous households.
- In terms of developing a more informed and involved civil society and encouraging public accountability of functionaries, it would be a far simpler definition to convey and to encourage the public to monitor - this is the first governance issue.
- The mischievous impact of certain political statements during the build-up to SP2 would not have occurred with the above alternative: there is substantial evidence from the qualitative reports from module 5 that social cohesion (a form of social capital) was damaged by the unscrupulous manipulation, or lack of understanding, of those who raised expectations and confused the issue of household-based distribution - this is the second governance issue.

Note that the general idea above is *not* restricted to universally distributed benefit. For example, it applies within regions if geographic targetting is attempted. It is more-or-less essential by definition if an individual "earns" a pack by labouring on public works or other workfare schemes.

I. M. Wilson
August 2000

Registration Information

1. PREAMBLE - DATA & ASSUMPTIONS

The data from Module 5 covered 54 villages where a census was carried out, so that the figures recorded should reflect the entire population of the set of villages. There is a certain amount of data missing from the individual records, as entered in the computer file, and in some cases data entry - completed under pressure of time - has included obvious errors such as values outside feasible ranges. These features mean there are tables below with slightly different totals, and there are minor degrees of uncertainty attached to some numbers in tables. For example only 25102 individuals' ages were recorded, but 25806 were aggregated from individual to household files, these latter recorded as comprising 6142 households with an average household size of 4.20 individuals.

The tables selected for inclusion in this description have as far as possible been chosen so that the *major messages* are not based on very small numbers of cases or very unusual values, but where such numbers are shown they should not be over-interpreted. While data have been screened for obviously wrong values that would clearly invalidate key conclusions, the analysis files undoubtedly still contained some errors.

This section *takes as given* the decisions by field staff as to what constitutes a farming family or household. The term 'household' is used below - as it was by the field team. It should be noted that since the module was concerned with Starter Pack registration amongst other issues, there is every likelihood that results relating to the composition of households reflect the same set of biases as may have operated in the SPLU's work. The average size of households may therefore be lower than by other methods of ascertainment, and the number of households correspondingly greater. Thus it may be that reported proportions of households *not* receiving starter packs will be larger than would have appeared if it had been possible to do a study which avoided these possible biases.

The record suggests that 5004 individuals received starter packs in these 6142 households, equivalent to about 815 per 1000 households. Since some households certainly received more than one pack, there were clearly more than [6142 - 5004] = 1138 households who did not receive [or acknowledge receiving] a pack at all as is explored below. The questioning did not ascertain whether an *individual* had received more than one pack so the figure of 5004 could be an underestimate, even if all respondents were wholly truthful.

For simplicity, the text below is worded as if:-

- all individuals were truthful in their responses and no individuals received multiple packs
- non-response can be ignored at the item level and at the questionnaire level
- fieldwork achieved exactly 100% identification of every household in every village

The broad thrust of conclusions will not be greatly affected by mild departures from these assumptions.

2. DISTRIBUTION OF PACKS TO HOUSEHOLDS

In broad outline the figures in Table 2.1 below indicate that many more households received NO pack than received more than one.

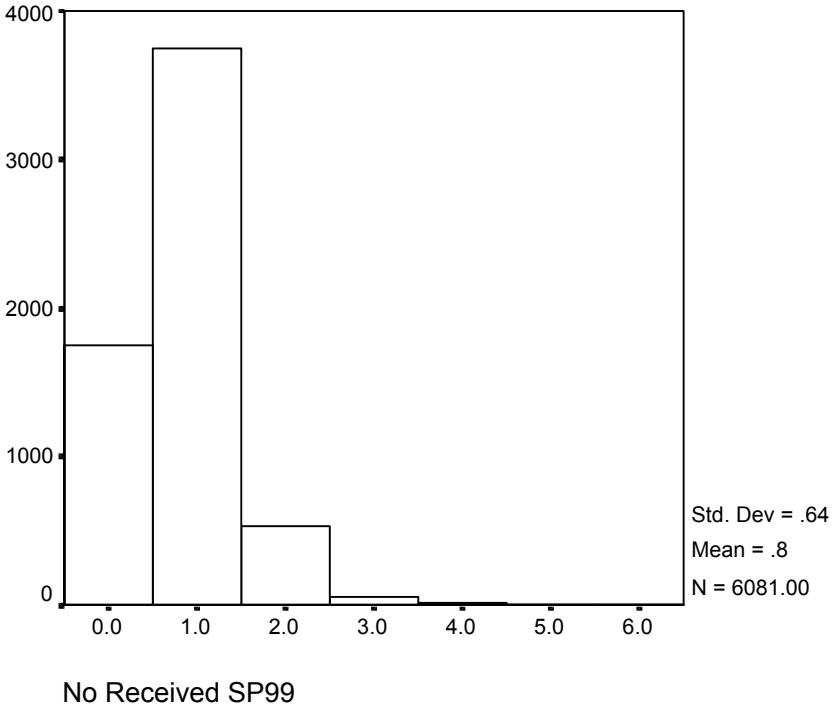


Table 2.1 below looks at the entire population recorded in the 54 villages, and at the numbers of persons recorded as receiving a pack. No attempt is made in this table to isolate “farming families” from other rural households.

		No Received SP99							
		.00	1.00	2.00	3.00	4.00	5.00	6.00	Total
FAMLYSIZ	1	162	311						473
	2	342	539	70					951
	3	323	736	102	6				1167
	4	302	653	94	12	3			1064
	5	231	553	84	5	1			874
	6	176	413	67	5	1			662
	7	97	260	46	8	2			413
	8	57	156	32	7		2	2	256
	9	33	69	17	3			1	123
	10	19	38	7	3				67
	11	4	8	5	2				19
	12	2	6						8
	13		2						2
	14			1					1
	20		1						1
Total		1748	3745	525	51	7	2	3	6081

Table 2.1

According to this table 525 families received one *extra* pack, 52 received 2 *extra* packs, etc. so the reported gross ***over***-distribution was 525 x 1 + 52 x 2 + 7 x 3 + 2 x 4 + 3 x 5 = 673 packs amongst 588 households (9.7% of all households)

The apparent gross ***under***-distribution was of 1748 packs to the households (28.5%) who received none. Therefore the ***net under***-distribution was 1075 packs (17.5% of the number of households). This is less than the 1138 reported earlier because of cases where some data was missing.

- *If the 25102 recorded inhabitants had been in households averaging 4.4 persons, the recorded national average from the recent NSO Census, there would have been about 5705 rather than 6142 households. This might have equated to a net under-distribution of 638 packs rather than 1075 i.e. about 11.2% of the 5705 households. The apparently large under-distribution <u>may</u> be <u>partially</u> explained by "over-registration of households" relative to the NSO definition.*

- *The qualitative information sheets collected by Module 5 teams record several instances of substantial numbers of packs being misappropriated at or near the point of delivery e.g. by village headmen and cronies.*

- *A rather rough coding of occupations of heads of households suggests that of the 673 excess packs amongst 588 households recorded above, 561(83%) went to 496 (84%) households where the head of household had farming recorded as his main occupation. Farming was the main occupation of some 90% of those for whom any occupation was identified.*

3. LOCAL FACTORS AFFECTING PACK DISTRIBUTION & SATISFACTION

Various factors affected registration, deletion and the receipt of packs and the informally collected narrative information often indicates reasons operative at local levels where individuals exerted influence, e.g. in one village in Lilongwe ADD in 1989-1999:-

> *"The GVH [group village headman] was very segregative. He refused to register those whom he said were rude or were of a higher living standard than him. Most women that refused to sleep with him were left out. He demanded a little something from those he hates (K.50.00). In some households he registered one person, or two or more depending on the amount they gave him."*
> *"Most of those that registered were not told about the distribution. The GVH just informed his few friends and relatives. He arranged some business with other people at the centre of distribution and these people answered to the names of those that did not come (due to ignorance of the exercise)."*
> *"..some people were registered [earlier by the FA] but their names were not called out during distribution. There were about 22 out of a hundred whose names were deleted. The deletion criterion was none other than hatred. The GVH is the one who did the job and created ghost names for his own benefit."*

Happily this GVH's "reign got terminated" in March 2000, after villagers' complaints to various authorities, but this does not seem to have been the outcome in all cases. There is important information in these village narratives which should be looked at those responsible for any future activity analogous to starter pack distribution.

The existence of local problems of this type can be seen in relation to overall satisfaction levels from the household level information collection of module 5.

Respondents were asked whether they thought the registration exercise for SP2 was satisfactory. Table 3.1 below looks at the relationship between family size, number of packs received by the household, and the response to the satisfaction question. It is of note that 1247 households (more than 20%) were recorded as being satisfied with the distribution arrangements though they had not received a pack. Those recorded as dissatisfied constituted about 13.6% of households. 56% of the dissatisfied (458/816) had not received a pack at all. Most of the remainder (40%) had received one pack.

				Family Size Group				
				Size 1	Size 2-3	Size 4-5	Size 6 or more	Total
Exercise OK	Yes	Grouped No Recd SP2	None	111	477	375	284	1247
			One	305	1164	1090	835	3394
			Two or Three		173	184	187	544
			Four +			4	8	12
		Total		416	1814	1653	1314	5197
	No	Grouped No Recd SP2	None	46	177	140	95	458
			One	2	104	108	113	327
			Two or Three		5	11	15	31
		Total		48	286	259	223	816

Table 3.1

The reasons for dissatisfaction *despite having received one or more packs* have moved (327 + 31) = 358 of the total sample of households (6%) to report such dissatisfaction. It is clear from the narrative information that in many cases the dissatisfaction is connected with a feeling of injustice engendered by local mismanagement of the registration process, and perhaps with deletion and distribution experience which had coloured their views. Table 3.2 below refers to the village from which the above quotations are reported. In this village about 17% have felt able to report dissatisfaction, the vast majority of these not having received a pack.

				Family Size Group				
				Size 1	Size 2-3	Size 4-5	Size 6 or more	Total
ExerciseSatisfactory	1	Grouped No Recd SP2	None	14	92	75	65	246
			One	5	24	21	6	56
			Two or Three		7	11	5	23
		Total		19	123	107	76	325
	2	Grouped No Recd SP2	None	6	28	17	7	58
			One		2	2	2	6
			Two or Three				2	2
		Total		6	30	19	11	66

Table 3.2

The reported levels of dissatifaction were consistently higher in Salima than other ADDs. Table 3.3 below summarises the information from Salima ADD.

				Family Size Group				
				Size 1	Size 2-3	Size 4-5	Size 6 or more	Total
ExerciseOK	1	Grouped No Recd SP2	None	25	89	91	61	266
			One	39	128	138	127	432
			Two or Three		20	16	16	52
		Total		64	237	245	204	750
	2	Grouped No Recd SP2	None	32	96	78	63	269
			One		32	39	38	109
			Two or Three		1	2	4	7
		Total		32	129	119	105	385

Table 3.3

This suggests that 47% ((266 + 265)/(750 + 385)) of the households in the Salima villages covered had not received a pack. That is a great deal higher than the 24% average reported over all the seven other ADDs. It provides a clear-cut indication of why dissatisfaction levels in Salima are high.

The high levels of non-receipt in Salima cannot sensibly be explained by differences in population composition, and clearly relate to administrative factors. the following tables indicate how different the deregistration/deletion process was in Salima than in other ADDs taken together.

		No Registered 99 in Salima ADD								
		.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	Total
No deleted 99 in Salima ADD	.00	69	421	69	6	1				566
	1.00	16	311	128	6	3				464
	2.00	3	8	98	5	2				116
	3.00		1		2	2				5
	4.00						1			1
	5.00					1	1	1		3
	7.00								1	1
Total		88	741	295	19	9	2	1	1	1156

Table 3.4

This Salima table has a certain number of people who are thought to have been deleted from the SPLU register, but who are also thought not to have been registered. This suggests a degree of misunderstanding and/or misrecording, but the *main message* is that over 35% of all households appear to have had all registered persons deleted (see dashed line). The equivalent figure for Table 3.5 (All Other ADDs) below is under 14%.

		No Registered 99 in ADDS other than Salima									
		.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	Total
No deleted 99 in ADDs other than Salima	.00	343	2786	538	42	7	5	2			3723
	1.00	13	399	377	32	9		1	1		832
	2.00	2	7	218	45	11	1				284
	3.00		1	4	29	11	1				46
	4.00					12	5	4			21
	5.00						3	5			8
	6.00							1	5		6
	7.00								1	1	2
	8.00			1						1	2
Total		358	3193	1138	148	50	15	13	7	2	4924

Table 3.5

Despite some uncertainties about the data accuracy and cleanness, this broad finding is in line with "common knowledge" about the deregistration process as it took place in Salima. It is quite clear that the combination of registration and deletion processes in Salima has worked to create a much higher level of popular discontent in Salima than elsewhere.

The above section of the report indicates one difficulty in analysing the whole data set. As well as the issues mentioned in section 1, we have to bear in mind that elements of the sample differ from others in *numerous* particular ways which may affect *certain* findings. A very full and detailed analysis might check every finding to see how far it was affected by known socio-political factors affecting villages, ADDs and other local areas. Given the limited time available, and the unreadable volume that such a comprehensive analysis would create, this has not been attempted here. Unless otherwise stated, results are concerned with the entire sample. Where possible, untoward effects of local circumstances have been noted.

4. DEMOGRAPHIC FACTORS AFFECTING PACK DISTRIBUTION

It is clear from narrative information that in a number of villages the registration and deregistration processes were implemented in differing ways, but that family composition, income sources and other factors were taken into account. The nature of family income was not addressed in module 5 except for a rough and incomplete account of 'main occupation'.

Table 4.1 illustrates that the average family size differed from region to region in the sample, by clear and substantial margins, being 10% bigger in the north than in the south.

			Sex of Head of H	
			Male	Female
			FAMLYSIZ	FAMLYSIZ
REGION	Nor	Mean	4.75	3.77
	Cen	Mean	4.54	3.67
	Sou	Mean	4.30	3.30

Table 4.1

The effect of restricting the dataset by removing Salima is included here to *illustrate* how the results may vary somewhat across the dataset. Table 4.1a below contains the same numbers for the North and South Regions, both of course unaffected. The average family size for MHH is barely affected in Central Region, but the figure for FHH is depressed a little.

			Sex of Head of H	
			Male	Female
			Mean	Mean
REGION	Nor	FAMLYSIZ	4.75	3.77
	Cen	FAMLYSIZ	4.55	3.46
	Sou	FAMLYSIZ	4.30	3.30

Table 4.1a

The *main finding* is that the average size of female-headed households (FHH) was lower, by approximately 1 in all regions. This finding is expressed in broad enough terms that it is not really affected by curious variations of the sample dataset . The *main finding* can be crudely construed as representing "the man missing " from FHH, though it conceals a great diversity of family compositions.

Table 4.1illustrates that the family basis of starter pack distribution, with distribution exactly according to 'the rules', means that one pack serves a rather different unit in each cell of the table. One pack per family in the southern region is a slightly greater measure of input per head - reasonable given the greater pressure on land in the south. One pack per FHH is a slightly greater measure of input per head - reasonable if FHHs are on average poorer. This appears to support the choice of a household basis for distribution, but there are other arguments against this approach.

The above argument relates to the *average* household, but there is clearly very great variation in household size and composition. In table 2.1, compare the 70 families of size 2 who apparently received 2 packs with the 69 families of size 9 who received one pack!

The smaller size of FHH is illustrated in Figure 4.1 below where there are *more* female headed than male headed households (MHH) of size 1. At family size 2 there are half as many FHH as MHH, but at family sizes 6 or 7 there are less than one-third as many FHH.

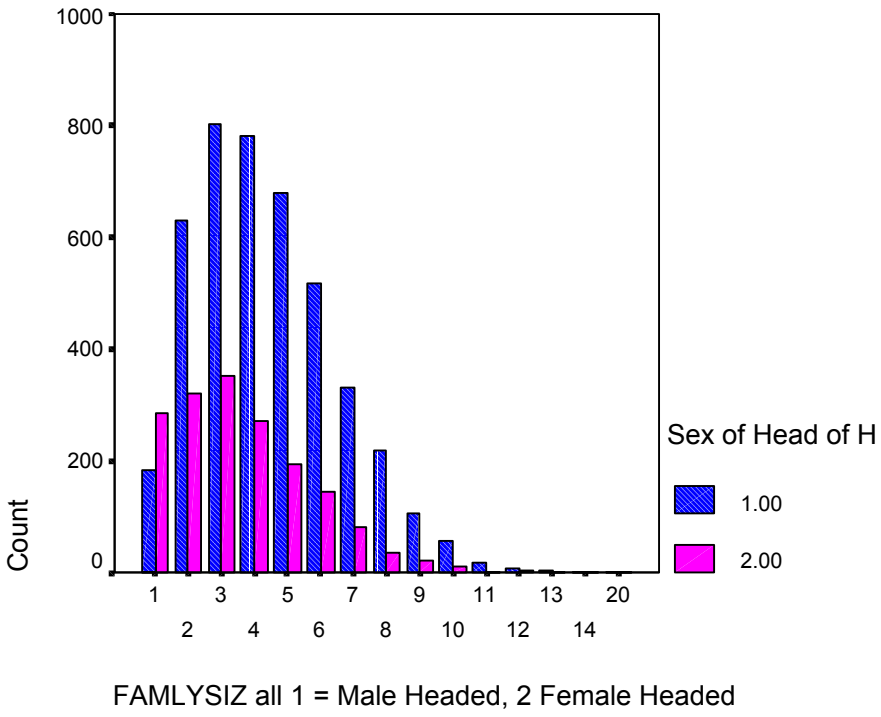


Figure 4.1

Combining the *distribution of family sizes* as above with the *regional variation* in means seen in table 4.1 produces a set of regional figures for family sizes for MHH and FHH as in Figure 4.2 below. The excess numbers of small families in FHH and the corresponding deficit of very large familes of FHH can be seen in every region.

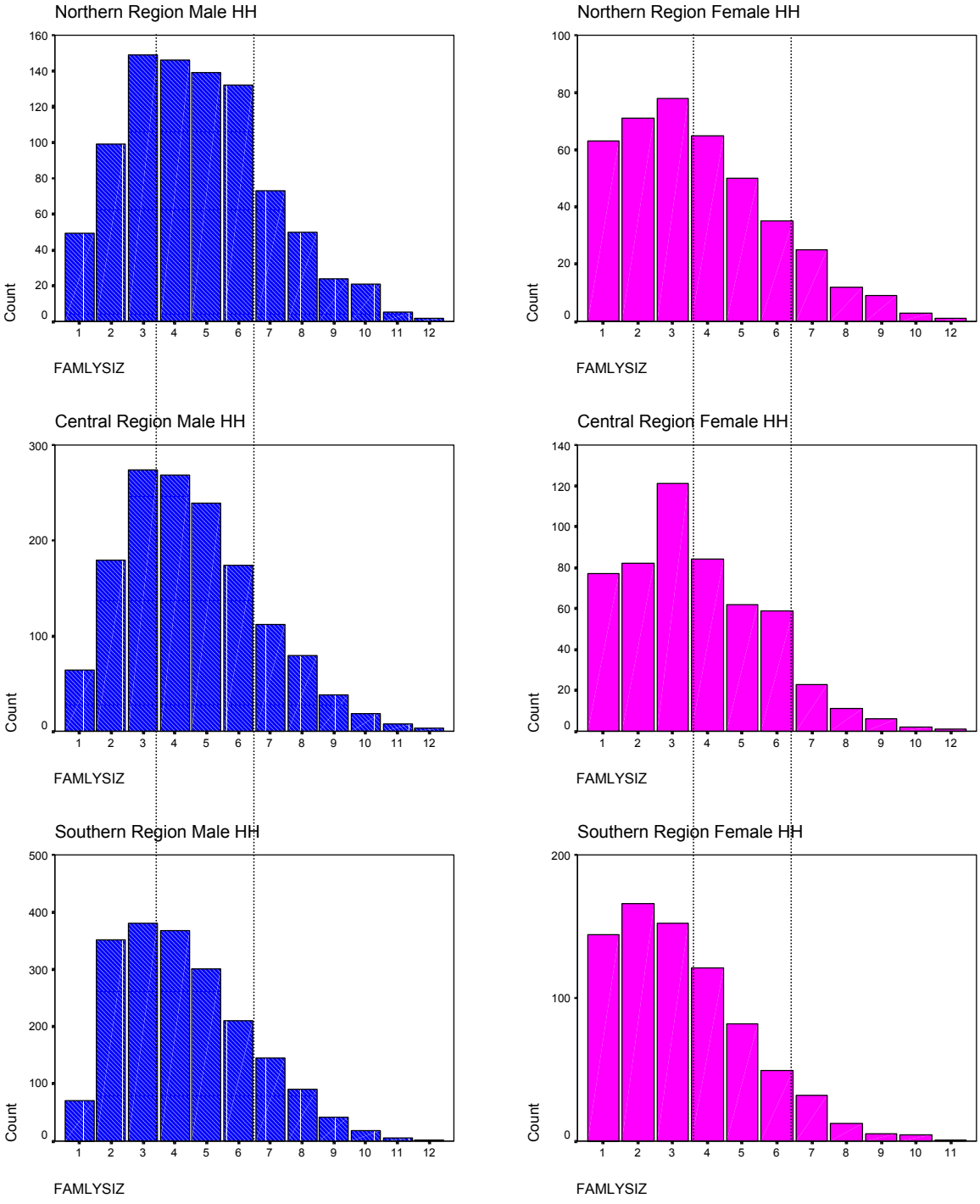

Northern Region Male HH
Count
FAMLYSIZ
Northern Region Female HH
Count
FAMLYSIZ
Central Region Male HH
Count
FAMLYSIZ
Central Region Female HH
Count
FAMLYSIZ
Southern Region Male HH
Count
FAMLYSIZ
Southern Region Female HH
Count
FAMLYSIZ

Figure 4.2

Families with no-one under 21are not infrequent, representing more than 15% of all families. Of course most of these are small in size, and they represent no more than 6.4% of the population. Such families are represented in Table 4.2 below.

		Sex of Head of Household		
		Male	Female	Total
FAMILY SIZE	1	171	268	439
	2	350	72	422
	3	49	10	59
	4	15	2	17
	5	3	1	4
	6	3	2	5
	8	2		2
Total		593	355	948

Table 4.2

				Family Size Group				
				Size 1	Size 2-3	Size 4-5	Size 6 or more	Total
ExerciseSatisfactory	1	Grouped No Recd SP2	None	14	92	75	65	246
			One	5	24	21	6	56
			Two or Three		7	11	5	23
		Total		19	123	107	76	325
	2	Grouped No Recd SP2	None	6	28	17	7	58
			One		2	2	2	6
			Two or Three				2	2
		Total		6	30	19	11	66

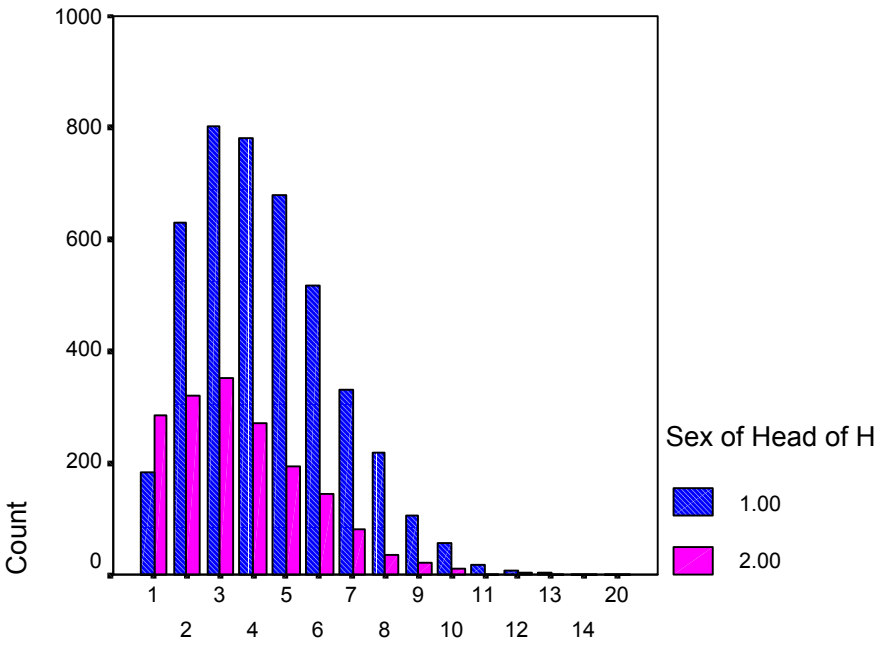
				Family Size Group				
				Size 1	Size 2-3	Size 4-5	Size 6 or more	Total
ExerciseOK	1	Grouped No Recd SP2	None	25	89	91	61	266
			One	39	128	138	127	432
			Two or Three		20	16	16	52
		Total		64	237	245	204	750
	2	Grouped No Recd SP2	None	32	96	78	63	269
			One		32	39	38	109
			Two or Three		1	2	4	7
		Total		32	129	119	105	385

		No Registered 99 in Salima ADD								
		.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	Total
No deleted 99 in Salima ADD	.00	69	421	69	6	1				566
	1.00	16	311	128	6	3				464
	2.00	3	8	98	5	2				116
	3.00		1		2	2				5
	4.00						1			1
	5.00					1	1	1		3
	7.00								1	1
Total		88	741	295	19	9	2	1	1	1156

		No Registered 99 in ADDS other than Salima									
		.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	Total
No deleted 99 in ADDs other than Salima	.00	343	2786	538	42	7	5	2			3723
	1.00	13	399	377	32	9		1	1		832
	2.00	2	7	218	45	11	1				284
	3.00		1	4	29	11	1				46
	4.00					12	5	4			21
	5.00						3	5			8
	6.00							1	5		6
	7.00								1	1	2
	8.00			1						1	2
Total		358	3193	1138	148	50	15	13	7	2	4924

			Sex of Head of H	
			Male	Female
			FAMLYSIZ	FAMLYSIZ
REGION	Nor	Mean	4.75	3.77
	Cen	Mean	4.54	3.67
	Sou	Mean	4.30	3.30

			Sex of Head of H	
			Male	Female
			Mean	Mean
REGION	Nor	FAMLYSIZ	4.75	3.77
	Cen	FAMLYSIZ	4.55	3.46
	Sou	FAMLYSIZ	4.30	3.30



FAMLYSIZ all 1 = Male Headed, 2 Female Headed

		Sex of Head of Household		
		Male	Female	Total
FAMILY SIZE	1	171	268	439
	2	350	72	422
	3	49	10	59
	4	15	2	17
	5	3	1	4
	6	3	2	5
	8	2		2
Total		593	355	948

Households with no-one under 21