KENYA
TRIP REPORT

Prepared by: Armando Levinson **Date Prepared: February 8, 2001**

Site Visited: Nairobi, Kenya **Submitted to: Joanne Jeffers, USAID**

Period Covered: January 21 to February 4, 2001

Trip Purpose and Summary

At the request of USAID/Kenya and the Kenyan Central Bureau of Statistics (CBS), Armando R. Levinson traveled to Nairobi, Kenya, to assist the CBS in the development of estimation procedures for the 2000 Post-Enumeration Survey (PES).

The CBS conducted the 1999 Kenya Population and Housing Census (CPH) in August 1999, with August 24-25, 1999 as the reference night. A PES is normally conducted approximately one month after all the census questionnaires are returned from the field and the results used to evaluate the census coverage and content errors. The 2000 Kenyan Post-Enumeration Survey took place during the period February 20-26, 2000, that is, almost six months after the census exercise. The matching exercise took place in August 2000 and the reconciliation visits have not yet taken place. The reconciliation visits are an integral part of the PES because the data obtained from this exercise permits the calculation of the census population from the "Census or E Sample" which, in turn, is used to calculate very important indicators such as the Erroneous Inclusion Rate, the Net Coverage Error and the Gross Error Rate per Unit Enumerated.

Upon arrival, Levinson met with David Nalo, CBS Director; J. Kekovole, Census Project Manager; F. Munene, CBS Deputy Director; P. W. Nyongesa, Chief of Data Processing Department; and Fred Otieno, Demographer to discuss the status of census and PES operations and the details of the work.

Levinson discussed the results of PES implementation and matching operation with Otieno. A summary of covered and matched households is presented in Annex 1. Page 4 of Annex 1 shows that 49.75 percent of households were matched in the urban areas and 73.39 percent of households were matched in rural areas (see page 13 of Annex 1). The national match rate was 64.30 percent. A copy of the PES questionnaire is shown in Annex 2. Here are some observations about the PES, followed by recommendations for a future PES.

1. The CBS intended to scan PES questionnaires and did not follow the questionnaire format recommended by Levinson and Weber (Trip Report of Levinson and Weber, August 9, 1999). The format used by CBS had the following disadvantages:

   a. No extra line was provided in the PES questionnaire to transcribe census information onto the PES questionnaire after the matching operation. Matched persons were identified in the census questionnaire with a line number (next to the person's name) that matched the line number of the same person in the PES questionnaire.

   b. The data from the matched census questionnaires was keyed directly from the census questionnaire after having located the correct person's information. This can cause errors in data entry. In addition, census questionnaires should not be used for this operation. They may get lost or mutilated in the process. Census questionnaires should be returned to their proper place after the matching operation and not be used for PES data entry.

   c. The back page of the PES questionnaire is mislabeled. It reads: "D: Information Regarding persons Enumerated in the Main Census but not in the PES." It should have read: "D: Information about Out-movers." The last column, Status of Census Enumeration, P-80A, should not have been placed there. If a person is identified as an out-mover, the person has been correctly enumerated and, therefore, there is no need to have this column in the out-mover section of the questionnaire.

   d. No areas were devoted to the "Information Regarding Persons Enumerated in the Main Census but not in the PES." These are the cases that have to be sent for reconciliation, at which time a person is determined to have been either "Correctly Enumerated" or "Erroneously Included."

   e. Levinson recommended that a sample of 25 percent of cases be sent back for reconciliation in order to obtain some information about those cases enumerated in the census but missed in the PES. Once this information is obtained, the CBS will be able to continue with the estimation process.

   f. If the CBS intends to document the results of the PES, the data users must be informed of all the problems encountered during all the phases of PES implementation and estimation.

   g. The PES is an integral part of a census of population and housing. Enough resources and staff should have been allocated during the planning stages to the PES. This is the most important lesson derived from the Kenyan 2000 PES. It is therefore highly recommended that for the next census, the CBS include the PES component as an integral part of the census project.

Levinson met with Jean Marc Hie, UNFPA Regional Advisor in Data Processing, and with Jason Ondora Osembe, UNFPA Regional Advisor in Survey Methods and Sampling, who were in country

assisting the CBS in areas such as census data processing (creation of a REDATAM data file) and PES implementation. Levinson informed Hie and Osembe of a future workshop in statistical methods and sampling methodology to be conducted in Nairobi. Hie and Osembe expressed interest in sending some participants from other English-speaking countries. The exact date for this workshop has not yet been determined. Levinson will inform Hie and Osembe of the date of the sampling workshop once it is determined.

Levinson also mentioned to Hie and Osembe about a workshop to be conducted in Senegal during the month of May in PES implementation and estimation procedures. Hie and Osembe expressed interest in conducting a similar workshop for English-speaking countries in Africa. They requested Levinson to send a proposal to their office in Addis Ababa.

Prior to his departure, Levinson met with the CBS Director, members of his staff, and the UNFPA delegation to discuss the work accomplished and future activities. Levinson was asked by USAID/Nairobi to review the project and integrate new activities, such as the establishment of a LAN (Local Area Network) and Internet access. The CBS has to obtain estimates from local providers and inform USAID/Nairobi and IPC of the results. Once these figures have been obtained and the budget analyzed, a new set of activities in support of the 1999 Kenyan Census of Population and Housing will be drafted and sent to the CBS and USAID/Nairobi for review and approval.

Attachments
cc:

T. Takona	USAID/Kenya
D. Vogel	USAID/Kenya
D. Nalo	CBS Director
F. Munene	CBS Deputy Director
J. Kekovole	Census Project Manager
F. Otieno	CBS Demographer
P. W. Nyongesa	Chief of Data Processing Department
S. L. Oyombe	Senior Economist Statistician
G. Sierra-Zorita	AF/EA, Desk Officer
J. M. Anderson	AF/EA, Director
J. Jeffers	USAID/G/PHN/POP/PE

ATTACHMENT A

Persons Contacted

USAID/Nairobi

Street Address:	ICIPE Complex Kasarani Nairobi, Kenya
Mailing Address:	Box 30261 Nairobi, Kenya
Telephone:	(254-2)-862-400; 862-402
Fax:	(254-2)-862-949

Timothy Takona, Project Manager, Health, Population, and Nutrition Officer
E-mail: ttakona@usaid.gov ([redacted])

Central Bureau of Statistics (CBS)

Street Address:	Herufi House (between Central Bank and Ministry of Finance) Nairobi, Kenya
Mailing Address:	P. O. Box 30266 Nairobi, Kenya
Telephone:	(254-2)-333-970
Fax:	(254-2)-333-030
E-Mail:	Not available

Census Operations Nyayo House (Near the InterContinental Hotel)

David Nalo, CBS Director (Phone: 340-929)
Francis M. Munene, Deputy Director
Dr. Kekovole, Census Project Manager
(Office 17th floor, Room 1720; phone: 22-43-85; 21-29-55; fax: 21-61-34)
P. W. Nyongesa, Chief of Data Processing Department
Fred Otieno, Demographer
Sammy L. Oyombe, Senior Economist Statistician, National Sample Survey and Evaluation Program (NASSEP), Division of Field Administration, NASSEP, and Cartography
Collins Opiyo, Demographer

United Nations Fund for Population Activities (UNFPA)

Mailing Address: P. O. Box 8714
Addis Ababa, Ethiopia

Telephone: 251-1-51-12-88, 51-31-34
Fax: 251-1-51-71-33
Telex: 21240

Jean Marc Hie, Regional Advisor in Data Processing (jm_hie@hotmail.com) (hie@unfpa.org)
Jason Ondora Osembe, Regional Advisor in Survey Methods Sampling (jonsembe@hotmail.com) (onsembe@unfpa.org)
Benson Morah, Director UNFPA, CST, Addis Ababa (morah@unfpa.org)

ATTACHMENT B

ESTIMATION PROCEDURES AND TABULATION FOR THE FEBRUARY 2000 KENYAN POST-ENUMERATION SURVEY

ARMANDO R. LEVINSON
U.S. CENSUS BUREAU

JANUARY 2001

1. Background

The report on "PRELIMINARY RECOMMENDATIONS ON THE SAMPLE DESIGN AND METHODOLOGY FOR THE POST-ENUMERATION SURVEY TO EVALUATE THE 1999 KENYAN POPULATION AND HOUSING CENSUS" (Levinson and Weber, August 1999) described the initial plans for the Post-Enumeration Survey (PES) to be conducted by the Central Bureau of Statistics (CBS) following the August 1999 Kenyan Census of Population and Housing (CPH).

The census was originally conducted in August of 1999, with August 24/25 as the census reference night. The PES exercise was conducted during the period February 20-26, 2000, almost six months after the census enumeration. The receipt of PES questionnaires in the main office took place during the period March 10-13, 2000. The matching operation started in August 2000 and out of the 529 original Enumeration Areas (EAs), 81 have yet to be re-matched because either too many census questionnaires could not be matched to PES questionnaires, or vice versa. For those EAs that were matched, summary statistics per census block were prepared to determine the number of omissions and match and non-match rates. Annex 1 presents the results of the matching operation for each covered EA.

The purpose of the reconciliation visits, an integral part of the PES, is to determine the match status and the census enumeration status of (a) cases determined to be "possible matches" and (b) cases enumerated in the census but not in the PES. This operation is vital to the success of the PES because the data obtained from this exercise permits the calculation of the census population from the "Census or E Sample" which, in turn, is used to calculate very important indicators such as the Erroneous Inclusion Rate, the Net Coverage Error and the Gross Error Rate per Unit Enumerated.

The purpose of this report is to provide the estimation procedures to carry out the estimation of characteristics and finalize the tabulation plan.

2. Distribution of PES Sample EAs

Table 1
Distribution of the Number of Households by Province and Urbanity Based on 1989 CPH Final Results and PES Sample Size in Each Domain of Estimation

Province	Urban			Rural			Total		
	HHs	%	Sample EAs	HHs	%	Sample EAs	HHs	%	Sample EAs
Nairobi	382,863	39.88	60	0	0	0	382,863	8.94	60
Central	72,335	7.54	10	591,027	17.79	68	663,362	15.49	78
Coast	159,202	16.58	22	201,680	6.07	28	360,882	8.43	50
Eastern	56,665	5.90	10	621,060	18.69	70	677,725	15.83	80
Nyanza	82,238	8.57	10	618,678	18.62	70	700,916	16.36	80
Rift Valley	172,367	17.96	20	834,894	25.14	74	1,007,261	23.51	94
Western	34,271	3.57	10	454,890	13.69	48	489,890	11.44	58
Total*	**959,941**	**100.00**	**142**	**3,322,229**	**100.00**	**358**	**4,282,170**	**100.0**	**500**

* Above total excludes the northeastern province.

Table 1 presents the *Distribution of the Number of Households by Province and Urbanity Based on 1989 CPH Final Results and PES Sample Size in Each Domain of Estimation.* This table was prepared during Levinson's visit to Kenya for the period July 21 to August 5, 1999. The new distribution of households and PES sample size in each domain of estimation is presented in Table 2 below.

Table 2
Distribution of the Number of Households by Province and Urbanity Based on 1999 CPH Final Results and Final PES Sample Size in Each Domain of Estimation

Province	Urban			Rural			Total		
	HHs	%	Sample EAs	HHs	%	Sample EAs	HHs	%	Sample EAs
Nairobi	649,426	33.30	60	0	0.00	0	649,426	10.19	60
Central	183,423	9.41	10	741,122	16.73	68	924,545	14.51	78
Coast	262,565	13.46	27	264,862	5.99	29	527,427	8.28	56
Eastern	144,168	7.39	21	813,480	18.40	70	957,648	15.03	91
North Eastern	33,424	1.71	10	114,582	2.63	0	148,006	2.33	10
Nyanza	197,678	10.14	10	770,336	17.42	70	968,014	15.19	80
Rift Valley	322,338	16.53	24	1,172,643	26.52	72	1,494,981	23.46	96
Western	157,210	8.06	10	544,113	12.31	48	701,323	11.01	58
Total	**1,950,232**	**100.00**	**142**	**4,421,138**	**100.00**	**358**	**6,371,370**	**100.0**	**529**

3. Estimation Procedures

3.1. Definitions

Census errors can arise from several sources such as imperfect data collection and processing procedures or poor operational control resulting in the loss of documents or erroneous coding and keying. Census designers also contribute to errors by producing faulty instruments, instructions, training materials, and procedures. Therefore, errors are intrinsic to a large data collection endeavor such as a census and they will occur in spite of efforts to avoid them.

In order to better understand the different categories of errors, it is important to design an analytical framework for them and divide them as follows:

a. Coverage versus content
b. Net versus gross
c. Sampling versus nonsampling

3.1.1. *Coverage Error*

It is the error in the count of persons or housing units resulting from cases that were "missed" during census enumeration or counted erroneously either through duplication or erroneous inclusion. It arises in census enumerations due to such factors as defective field operations, carelessness on the part of enumerators, misunderstanding or lack of cooperation on the part of respondents, or simply because census forms are lost or destroyed during the census processing operation. Within this category of error, it is important to distinguish three types:

a. ***Omissions*** result from:

i. Entire housing units, households, or persons having no established place of residence (nomads, for example) having been missed by census enumerators. Households can be missed because all of its members were at another place of residence at the time of the census enumeration, were temporarily absent during the hours of census enumeration (working, at school, etc.), or were in transit either within or outside of the country during the enumeration period.

ii. From one or more persons within enumerated housing units or households being missed. In the case where an entire housing unit is missed, it follows that all households and persons residing within the housing unit will also be missed by the census. Omissions of persons within enumerated households can occur for a variety of reasons, including deliberate or inadvertent omission of household members or the application of an incorrect definition of the household by the census enumerator. For instance, very young children and adults aged 15-29 tend to be over-represented among persons missed in census enumerations. Very young children are sometimes omitted because of respondent recall error, while persons aged 15-29 are in the age range in which the rate of residential mobility is high and they tend to be omitted due to uncertainty on the part of respondents and enumerators as to where they should be enumerated.

There are two primary causes of omission of housing units:

1. Failure to include part of the land area of the country in creating enumerator assignments. This can be caused by such factors as imprecise boundaries of geographic or census administrative units, faulty maps, or simply by coverage errors made by field staff in the pre-census listing operation.

2. Enumerator canvassing error within assigned areas. This can result from such factors as imprecise definition of enumerator assignments, faulty maps, or simple oversight on the part of the enumerators. These errors tend to occur more frequently in sparsely settled rural areas where villages are separated by large distances and in densely settled urban areas characterized by multi-unit structures which are arranged in a complex fashion.

b. ***Duplications*** occur when housing units, households, or persons are enumerated more than once in a census. Frequently, duplication is caused by the overlapping of enumerator assignments due to errors made during pre-census listing or enumeration area delineation, or because of the inability of enumerators to identify on the ground the proper boundaries of a particular enumerator assignment. Persons who are mobile residentially or have more than one residence are especially prone to being enumerated more than once.

As a result of duplications, it is conceivable that a census count of the total population may be larger than the actual population. It is more frequently the case, however, that the number of omissions exceeds the number of duplications resulting in a net census undercount.

c. ***Erroneous Inclusions*** include housing units, households, and persons that were enumerated in the census and either should not have been or were enumerated in the wrong place. Examples of erroneous inclusions are persons who died before the census, who were born after census day, or aliens (if the methodology used defines them to be out-of-scope for census purposes, as is the case in *de jure* censuses, where only persons who "usually" reside in the country are counted). It is not uncommon for persons to be enumerated in the wrong geographic area; this would result in overenumeration for that area and undercoverage for the area where they should have been enumerated.

3.1.2. *Content Error*

It is the error in the recorded characteristics of those persons that were enumerated in the census. It arises from erroneous or inconsistent reporting of characteristics by respondents, from failure on the part of enumerators to obtain or record accurately the required information, or simply from errors introduced in the clerical and processing operation. As an example, we can mention a person who reported sex as male in the census and as female in the PES.

3.1.3. *Net Versus Gross Error*

Gross error refers to the total number of errors made in the census. ***Net error*** refers to the total effect of these errors on the resultant statistics. Gross errors affect the nonsampling variances, while net errors affect the nonsampling biases.

As an example, ***gross census coverage error*** would consist of the total of all persons omitted plus all the duplicates plus all erroneous enumerations. In measuring ***net census coverage error***, one must take into account the fact that one of these types of error results in an underestimate of total population (omissions) while the other two types (duplications and erroneous inclusions) result in overestimates. Thus, net census coverage error would be measured by the excess or deficit of errors resulting in population underestimates over those errors resulting from population overestimates. A ***net census undercount*** exists when the number of omissions exceeds the sum of the number of duplications and the number of persons erroneously enumerated, while a ***net census overcount*** exists when the number of duplications plus erroneous enumerations exceeds the number of omissions.

3.1.4. *Sampling versus Nonsampling Error*

Sampling error arises because information is not collected from the entire target population, but rather only from some portion of it.

Nonsampling error is defined as a residual category consisting of all other errors which are not the result of the data having been collected from only a sample. These include errors made by respondents, enumerators, supervisors, office clerical staff, data entry operators, etc.

The total error is defined to be the sum of the above-mentioned two errors. It is usually the case that the larger component of the total error are the nonsampling errors. The latter are very difficult and expensive to measure and must therefore be prevented or controlled.

3.2 Procedure for Evaluation of Census Coverage

There are three procedures for evaluating census coverage in a post-enumeration survey. These three procedures, known as A, B, and C, respectively, are discussed in Chapter A-13 of the POPSTAN document, and in Chapter 2 of the manual titled "Evaluating Censuses of Population and Housing," both published by the U.S. Census Bureau.

Procedure A attempts to reconstruct the households as they existed at the time of the census; the respondent must identify all persons in the sample households as of the census reference date. These persons are then matched to the census questionnaires. Based on this information, the number and percentage of matched non-movers and out-movers are estimated. In procedure A, the matching of non-movers and out-movers is relatively easy and inexpensive because the search is limited to the sample areas and their adjacent areas. The weakness of procedure A lies in the fact that it is difficult and very expensive to locate movers given that they are no longer at the sample address. Therefore, there is a strong possibility of underestimation of the number of out-movers.

Procedure B tries to identify all persons currently in the sample household, that is, as of the date of the PES. Often, the respondent is asked to provide the addresses where these persons were living on "Census Day." The persons are later matched to the corresponding census records and, based on this information, the number and percentage of matched non-movers and in-movers are estimated. Even though procedure B provides a better estimate of the number of movers than procedure A, the difficulties and costs of matching associated with it are greater because it involves searching for in-movers in the area where they were enumerated during the census.

Procedure C tries to identify all current members of the sample household as of the PES date and, in addition, any other person present as of the reference date of the census. However, only the residents as of the date of the census, that is, the non-movers and the out-movers, are matched to the census records. The estimates for the number of non-movers and movers are based on procedure B; the matching rates are estimated based on procedure A.

In short, procedure C is a combination of procedures A and B. It takes advantage of the features of each to reduce matching difficulties and, at the same time, improve the estimation of movers. For this

reason, the recommended procedure for the Kenyan PES is procedure C. This procedure involves identifying all the residents at the sample household as of the date of the PES, as well as those present at the census reference date. ***However, only the residents at the time of the census, that is, the non-movers and the out-movers, are matched to the census records.***

4. Dual System Estimation (DSE) to Estimate Coverage Errors

4.1. Rationale for the Use of the DSE

To develop estimates of coverage combining the results of the two surveys, it is necessary to use a statistical model. The model indicates how errors in the census occur as a stochastic process and how an appropriate estimator of census undercoverage may be derived. This section presents a simple derivation of the dual system estimator, followed by a modification of the estimator to allow for erroneous enumerations in the census and measurement problems in the PES.

The model for coverage evaluation was originally developed for use in biometric studies to estimate the size of closed populations. A closed population in the biometric sense is one in which the composition of the population remains relatively unchanged over the time the study is being conducted: there are no births or deaths, and there is no immigration or emigration. The earliest use of the technique, known in biometrics as capture/recapture estimation, was to estimate the size of wildlife populations. An example is given to indicate how the model was originally derived.

Consider a lake with no inlets or outlets. We want to estimate the number of fish in the lake. Assume that the fish are uniformly and randomly distributed throughout the lake. We take a net and "capture" as many fish as possible in one catch. The fish are counted and the total denoted by N_1. The fish are tagged (for identification purposes) and they are released back into the lake. The tags do not injure or affect the fish in any way. After sufficient time has passed for the fish to redistribute themselves randomly throughout the lake (but not so long a period that one might encounter births or deaths of any fish), another net is cast into the lake and a second catch is taken. These fish are also counted and their total is denoted by N_2. A separate tally is also made of fish tagged during the first catch who were "recaptured" in the second catch, and this is denoted by M (for matched). We now have three totals and the captures can be distributed as in Figure 1 below.

Figure 1

Distribution of Captures in a Capture/Recapture Study

		Second Capture		
		Total	Caught	Not Caught
First Capture	Total	N_T	N_2	
	Caught	N_1	M	
	Not Caught			

where:

N_1 = the number of fish caught in the first attempt

N_2 = the number of fish caught in the second attempt

M = the number of fish caught in both attempts

N_T = the total number of fish in the lake

The objective is to estimate N_T, the number of fish in the lake. However, the number of fish ***not caught*** in either attempt remains unknown at this point. To model the capture process, the capture of a fish is conceptualized as a stochastic event, or more specifically, as a Bernoulli event (an experiment having only two possible outcomes, usually denoted *success* and *failure*, with the properties that the probability of occurrence of each outcome is the same in each trial and the occurrence of one excludes the occurrence of the other in any given trial). This implies that the counts of fish caught, M, N_1, and N_2, are random variables (the sums of Bernoulli outcomes). If the probability that a fish will be caught in the first attempt is "a" for each fish, and "b" in the second attempt, and the two captures are independent of each other, the expected values of the random variables M, N_1, and N_2 may be obtained as follows:

(4.1) $E(M) = abN_T$ (by independence of events)

(4.2) $E(N_1) = aN_T$

(4.3) $E(N_2) = bN_T$

An estimator may be obtained by substituting the observed data values for the expected values of M, N_1, and N_2, and substituting equations (4.2) and (4.3) into equation (4.1). Thus,

(4.4) $$M = abN_T = \frac{aN_T\, bN_T}{N_T} = \frac{N_1 N_2}{N_T}$$

Rearranging the terms in equation (4.4), we obtain:

(4.5) $$N_T = \frac{N_1 N_2}{M}$$

4.2. The Model Applied to Censuses Under Perfect Conditions

In a census evaluation application, the theory is exactly the same as the capture/recapture study described above, except that human populations require slightly different assumptions. Suppose that

the first capture is the census and the second capture is the PES. Let N_c (replacing N_1) be the census count, and N_p (replacing N_2) be the weighted sample total from the PES. N_p represents the total number of people (or units) to be found if the entire sampling frame used for the PES sample was contacted. It is critical, of course, that the sampling frame be independent of the census. For estimation purposes, it is best to have as complete a frame as possible so as to minimize the variance of the estimate and the number of persons (units) not found in either data source. The random variable M represents the number of people (units) captured in the PES who were enumerated in (are matched against) the census. The estimate is made in exactly the same way as in equation (4.5).
Note that there are a number of assumptions that have to be satisfied in applying the model to human populations. The model assumes that matching of persons and units is done perfectly between the two sources, that there are no multiple observations of individuals in either source, and that the two sources are independent. This latter assumption must be built into the design of the study by keeping the PES as independent as possible from the census. Therefore, the frame for the PES sample must be independent from the census. Perhaps the best choice for a sampling frame would be a current updated listing of the previous census or an area sample based on a recent listing of areas. Other requirements for keeping the two sources independent are using interviewers for the PES who are different from those used for the census and processing the PES separately from the processing of the census to avoid cross-contamination. ***The PES sample is used to estimate census misses. The sample drawn from the census, the E-sample, is used to estimate the population erroneously enumerated in the census.*** These are the two surveys mentioned at the beginning of section 4.1, page 5.

4.3. The Model Applied to Censuses Under Less than Perfect Conditions

There are several types of errors that can affect the estimation procedure described above:

a. The census may contain duplicate or multiple enumerations.

b. Housing units listed in the census may be correctly enumerated but allocated to the wrong geographic area.

c. Members of a housing unit may be enumerated in the wrong location or may not be in scope for the census (that is, should not have been counted at all).

d. Members of a housing unit may be incompletely enumerated so that there is insufficient identifying information for an individual.

Any of these factors can introduce bias into the estimation of the total population size. To measure or counter the effect of each of these factors, a second sample is used to provide correcting factors to the estimation process. This second, or E-sample, is drawn directly from the census for the same areas and using the same stratification as the sample described in the previous section.

A follow-up enumeration is conducted at each of the households in the E-sample and several pieces of information are collected. When the sample housing unit is initially contacted, the geographic location is checked and compared to the geographic location recorded for that housing unit in the census. If the locations differ, the case is reviewed (usually by a third party) and a determination is made as to whether the original geographic coding in the census was correct or incorrect. The reason for checking

geographic coding is that the housing units incorrectly located in the census are unlikely to be found when the PES match is conducted, unless unlimited resources are available to conduct the search. Thus, although the housing unit is correctly enumerated in the census, none of the occupants can be found or matched. As a result, the census total, N_c, will be correct, but the match total, M, will be too low. The method to correct the match total is presented in the next section.

In the enumeration undertaken for this sample, the interviewer ascertains for each occupant of the sample housing unit whether the occupant listed in the census actually exists (there are sometimes cases which are fabricated in the census by over zealous census enumerators) and whether that individual was correctly enumerated at that address in the census. A count of the number of persons who should <u>not</u> have been enumerated at that address is obtained for each household. In the case of census fabrications, these "non-persons" should be subtracted from the census totals since they don't exist. In the case of incorrect enumeration, if the PES interview was conducted correctly following the census enumeration procedures, the PES interview will either not list the individual (out-of-scope) or have them listed at another address, and in neither case will there be a match. In this case, the census totals, N_c, are incorrect and there should be no matches.

There are two further operations to be conducted in the PES processing office. For the area of search specified for the match portion of the PES operation, the sample cases drawn from the census should be compared to all other census returns in the same area. This comparison should be done for each person listed on the sample unit to determine whether there are any duplicate enumerations in the same area. Duplication in the census will cause the census total, N_c, to be overstated, though the number of matches made from the PES sample, M, will be correct.

Finally, there will be cases in the census which have insufficient information for matching to the PES sample. Although tallied into the census (probably as a result of imputing much of the missing information) so that the census total, N_c, is correct, the match cannot be completed and the match total, M, will be understated.

The following components of census erroneous enumeration would be estimated from the E-sample:

G = the number of persons incorrectly located geographically in the census;

E = the number of persons incorrectly enumerated in the census (fabricated or not in scope);

D = the number of duplicate enumerations in the census; and

I = the number of persons who are enumerated in the census but have insufficient information for matching.

Based on these estimates of census erroneous enumeration, the dual system estimate of total population could then be adjusted as follows:

$$N_T = \frac{N_p (N_c - G - E - D - I)}{M} \tag{4.6}$$

where N_p is the estimate of the total population from the PES sample (equivalent to N_2 in section 4.1 above).

The ***net undercount rate***, R_N, can be calculated as the total census count relative to the estimate of the total population size, or

$$R_N = \frac{N_c}{N_T} = \frac{\frac{M}{N_p}}{(N_c - G - E - D - I) / N_c} \tag{4.7}$$

4.4. Actual Dual System Estimator to Estimate Coverage Errors

The main purpose of the PES coverage error analysis is to estimate the ***census omission rate***, ***erroneous enumeration rate*** and the ***net coverage error rate*** for households and persons. The PES analysis of coverage errors will be based on the dual system estimator. This estimation involves the matching of records from two sources which are assumed to be independent (in this case, the census and the PES). For this purpose, we have two different samples:

> **P-sample** (population sample) consists of the PES sample EAs drawn from the same target population, but independently from the census, for the purpose of estimating **census omissions** when compared to census records.
>
> **E-sample** (enumeration sample) is drawn from the cases already enumerated in the census, but selected for independent re-enumeration for the purpose of estimating **census erroneous inclusions** when compared to original census records. This estimate of erroneous inclusions provides a correction in the dual-system of the true population. In practice, the E-sample is allowed to overlap completely with the P-sample to reduce costs and improve the precision of the estimates. The E-sample then consists of the same segments selected for the PES. A two-way match is conducted between the P-sample and the E-sample to identify both the omissions and the erroneous inclusions. The match also produces the estimates of the "matched population" component required in the dual system estimation.

A two-way matching is carried out between the PES and census questionnaires.

The dual system estimation for the population is illustrated here. The true population can be divided into the following components, based on whether a person was included or missed in the census and the PES.

True Population (T)		Census	
		Included	Missed
PES	Included	a	b
	Missed	c	d

Following the matching between the PES questionnaires for the P-sample and the census records from the E-sample, the matched population, *M*, is used to estimate the component *a*. The total number of **census correctly enumerated persons**, T_c', represents the components (*a*+*c*), and the total population estimated from the PES, T_p, represents the components (*a*+*b*). The purpose of the dual system estimator is to estimate the true population based on the results of the matching. Assuming that the census and the PES are independent, the probability of being matched (that is, included in both the census and the PES) can be expressed as follows:

$$P(Matched) = \frac{M}{T} = P(Included\ in\ Census)\ x\ P(Included\ in\ PES) = \frac{Matched\ Population}{True\ Population}$$

(1) $$\frac{M}{T} = \frac{T_c'}{T} x \frac{T_p}{T}$$

It can be seen from this expression that T can be estimated from the other components as follows:

(2) $$\hat{T} = \frac{T_c' x T_p}{M} = \frac{T_c'}{\left(\frac{M}{T_p}\right)} = \frac{T_c'}{R_M} = \frac{(a+b)(a+c)}{a}$$

where R_M is the ***census coverage rate*** (or match rate) estimated from the PES data and $\hat{T}$ is the dual system estimator of the true population.

The ***census omission rate*** (gross coverage error rate) is estimated by ($1-R_M$). Note that the dual system estimation does not assume that the coverage for the PES is better than the coverage for the census. The important assumption is the independence between the census and the PES. It is difficult to achieve complete independence between the two enumerations, since the sampling frame for the PES is based on the maps of census EAs, and persons who are difficult to enumerate in the census may have a higher probability of being missed in the PES. When there is such a lack of independence, the dual system estimator has a tendency to underestimate the census omission rate.

The ***total census correctly enumerated population*** (T_c') is estimated by subtracting an estimate of the

census erroneously enumerated population from the PES data (based on the E-sample), $\hat{E}_c$, from the official census figure for the total population. This can be expressed as follows:

(3) $$T_c' = T_c - \hat{E}_c$$

where T_c is the ***official census total population***

The dual system estimator of the total population ($\hat{T}_A$) is used to estimate the true population, *T*, using the following formula:

(4) $$\hat{T}_A = \frac{T_c - \hat{E}_c}{R_M} = \frac{T_c'}{R_M}$$

The PES results will probably not be used to adjust the official results from the census; they will only be used to evaluate the census coverage errors. The most important measure of coverage error is the ***net census undercount rate*** ($\hat{R}_U$), calculated using the following formula:

(5) $$\hat{R}_U = \frac{\hat{T}_A - T_C}{\hat{T}_A}$$

Based on the procedure C methodology, the PES estimate of the total population, $\hat{T}_p$, is equal to the weighted estimate of non-movers plus in-movers enumerated by the PES. The matched population *M* is estimated as follows:

(6) $$M = (MATCHED\ NON-MOVERS) + \left[IN-MOVERS \frac{MATCHED\ OUT-MOVERS}{TOTAL\ OUT-MOVERS} \right]$$

The expression in brackets is used to estimate the component of the matched population corresponding to movers. This is considered the most accurate (least biased) estimate for this component.

The dual system estimates will be calculated in two phases. The preliminary estimate of $\hat{T}_A$ will be based on the PES E-sample estimate of the total population enumerated in the census, $\hat{T}_C''$, as follows:

(7) $$\hat{T}_A' = \frac{\hat{T}_c'' - \hat{E}_c}{R_M}$$

where $\hat{T}_c''$ = weighted estimate of matched population + census correctly enumerated population (missed by PES) + census erroneously enumerated population

= PES survey estimate of the total population enumerated in census

Once the official census population (T_c) is available, this estimate of $\hat{T}_A'$ would then be ratio-adjusted as follows:

(8) $$\hat{T}_A = \hat{T}_A' x \frac{T_c}{\hat{T}_c''}$$

The final estimate $\hat{T}_A$ will then be used to calculate the final ***net census undercount rate*** $\left(\hat{R}_U\right)$ in equation (5) above.

The dual system estimation can be improved by post-stratifying the PES population into homogeneous subgroups. The PES coverage tables will be disaggregated by sex and age groups, since the census coverage tends to vary by these subpopulations. It is recommended to use the following eight age groups:

0 to 4 years
5 to 9 years
10 to 19 years
20 to 29 years
30 to 39 years
40 to 49 years
50 to 59 years
60 years and above

The dual system estimator of the adjusted population will be calculated using the formula specified previously for the subpopulation in each of the 16 sex/age groups within each domain of estimation.

5. Content Error Analysis

The content error evaluation will be based on the data for ***matched non-movers***. The out-movers will be excluded from this analysis since it is more difficult to obtain reliable information for the persons who have left the sample EAs after the census. However, since the out-movers generally represent less than 5 percent of the total population, this would not have a major effect on the results.

The census responses for person characteristics will be transcribed in the shaded cells of the PES questionnaire for matched non-movers. The content error tabulations will consist of cross-tabulating the PES responses for each item with the corresponding census responses. Lotus spreadsheets can be used to enter these results for calculating summary statistics, such as, the index of inconsistency for each category, the aggregate index of inconsistency, and the corresponding 95 percent confidence intervals, as well as the net difference rates and the rate of agreement. The formulas for the index of inconsistency and the other summary statistics can be found in the manual "Evaluating Censuses of Population and Housing." Copies of the spreadsheets have been left at the CBS for future use.

Content error is only estimated for matched persons and for selected variables in order to measure the inconsistency between census answers and PES answers to the same questions. This inconsistency is measured by means of four indicators: the net difference rate, the index of inconsistency (simple and aggregate), the gross difference rate, and the rate of agreement. These four indicators are presented below.

The variables included in the PES questionnaire for which content analysis is warranted are:

1. Name of Usual Household Member
2. Relationship to Head of Household
3. Sex
4. Age
5. Marital Status
6. Religion

5.1 <u>Net Difference Rate (NDR)</u>

$$NDR = \frac{Y_{.i} - Y_{i.}}{n} \times 100 \qquad i = 1, \ldots C$$

where:

$Y_{.i}$ = unweighted census number of cases in the i-th category;
$Y_{i.}$ = unweighted PES number of cases in the i-th category;
n = unweighted number of matched cases; and
C = total number of response categories for characteristic Y.

The 95 percent confidence interval of the NDR for category i, where i ranges from 1 to C, is obtained as follows:

$$\frac{Y_{.i} - Y_{i.} \pm 2\sqrt{Y_{.i} + Y_{i.} - 2Y_{ii} + 1}}{n} \; x \; 100$$

There are some exceptions to this formula:

(1) If $(Y_{i.} - Y_{ii}) = 0$, widen the high ninety-five percent confidence limit by adding:

$$\left[\frac{2}{n} \; x \; 100\right]$$

(2) If $(Y_{.i} - Y_{ii}) = 0$, widen the low ninety-five percent confidence limit by subtracting:

$$\left[\frac{2}{n} \; x \; 100\right]$$

(3) If both (1) and (2) occur, the ninety-five percent confidence limits are estimated as:

$$\left[-\frac{4}{n} \; x \; 100\right] \; to \; \left[+\frac{4}{n} \; x 100\right]$$

5.2 <u>Index of Inconsistency</u>

The index of inconsistency is the relative number of cases for which the response varied between the census and the PES. It is calculated for each response category i using the following formula:

$$\hat{I} = \frac{(Y_{.i} + Y_{i.} - 2Y_{ii})}{\frac{1}{n}\left[Y_{.i}\,(n - Y_{i.}) + Y_{i.}\,(n - Y_{.i})\right]} \; x \; 100 \qquad for \; i = 1,\ldots,C$$

where:

Y_{ii} = number of cases where category i was given as a response in both the census and the PES.

The ninety-five percent confidence interval of the index of inconsistency for category i, where i = 1, ..., C, is given by the following:

(1) If

$$\left(\frac{Y_{.i} + Y_{i.} - 2Y_{ii}}{n}\right) \leq .10$$

the ninety-five percent confidence limits are:

$$\frac{(Y_{.i} + Y_{i.} - 2Y_{ii} + 2) \pm 2\sqrt{Y_{.i} + Y_{i.} - 2Y_{ii} + 1}}{Y_{.i}\left(1 - \frac{Y_{i.}}{n}\right) + Y_{i.}\left(1 - \frac{Y_{.i}}{n}\right)}$$

(2) If

$$\left(\frac{Y_{.i} + Y_{i.} - 2Y_{ii}}{n}\right) > .10$$

the ninety-five percent confidence limits are:

$$\frac{(Y_{.i} + Y_{i.} - 2Y_{ii} + 2) \pm 2\sqrt{\frac{1}{n}(Y_{.i} + Y_{i.} - 2Y_{ii}\,(n_{..} - Y_{.i} - Y_{i.} + 2Y_{ii}}}{Y_{.i}\left(1 - \frac{Y_{i.}}{n}\right) + Y_{i.}\left(1 - \frac{Y_{.i}}{n}\right)} \; x\ 100$$

To calculate the Aggregate Index of Inconsistency (that is, for all the response categories of the characteristic under study), use the following formula:

$$\hat{I}_{AG} = \frac{\left[n - \sum_{i}^{C} Y_{ii}\right]}{\left[n - \frac{1}{n}\sum_{i}^{C} Y_{.i}Y_{i.}\right]} \; x \; 100$$

(1) If

$$\left[\frac{n - \sum_{i=1}^{C} Y_{ii}}{n}\right] \leq .10$$

the ninety-five percent confidence limits are:

$$\frac{\left(n - \sum_{i=1}^{C} Y_{ii} + 2\right) \pm 2\sqrt{n - \sum_{i=1}^{C} Y_{ii} + 1}}{\left(n - \frac{1}{n}\sum_{i=1}^{C} Y_{.i}Y_{i.}\right)} \; x \; 100$$

(2) If

$$\left[\frac{n - \sum_{i=1}^{C} Y_{ii}}{n}\right] > .10$$

the ninety-five percent confidence limits are:

$$\frac{\left(n - \sum_{i=1}^{C} Y_{ii} + 2\right) \pm 2\sqrt{\frac{1}{n}\left(n - \sum_{i=1}^{C} Y_{ii}\right)\left(\sum_{i=1}^{C} Y_{ii}\right)}}{\left(n - \frac{1}{n}\sum_{i=1}^{C} Y_{.i}Y_{i.}\right)} \; x \; 100$$

5.3 Gross Difference Rate (also Off-Diagonal Proportion)

The gross difference rate (GDR) is calculated for the variable as a whole. It is the number of discrepancies between the census responses and the PES responses relative to the total number of persons matched. It is equivalent to the sum of all cells off the diagonal, for all categories, or the complement of the sum of the diagonal cells.

$$GDR = \frac{\left[n - \sum_{i}^{C} Y_{ii}\right]}{n}$$

5.4 Rate of Agreement

The rate of agreement is the complement of the gross difference rate. A low rate of agreement indicates a high degree of variability and vice versa.

$$Rate\ of\ Agreement = \frac{\sum_{i}^{C} Y_{ii}}{n}$$

The following table provides standards for the interpretation of the different content error measures:

Measure	Low	Moderate	High
Index of Inconsistency	< 20	20 - 50	> 50
Aggregate Index of Inconsistency	< 20	20 - 50	> 50
Absolute Value of NDR relative to the mean or the proportion (NDR/P)	<.01	.01-.05	>.05

6. Calculation of Sampling Weights

In order to expand the PES data to the national or domain levels, it is necessary to apply a weight (expansion factor) to the data from each survey questionnaire. The basic weight for a sample household would be equal to the inverse of its probability of selection. Since the EAs will be selected independently with probability proportional to size (PPS) within each domain of estimation, the probability of selection of an EA is given by:

$$p_{hi} = n_h \frac{M_{hi}}{M_h} = n_h \frac{M_{hi}}{\sum_{i=1}^{N_h} M_{hi}}$$

where:

n_h = number of sample EAs selected in the h-th first-level stratum.

M_{hi} = measure of size (total number of households from the census map work) for the i-th sample EA in the h-th first-level stratum (domain of estimation).

M_h = cumulated measure of size (total number of households from the census mapping work) for the h-th first-level stratum (domain of estimation).

The basic weight for a sample household is equal to the inverse of its probability of selection and is given by:

$$W_{hi} = \frac{M_h}{n_h x M_{hi}} = \frac{\sum_{i=1}^{N_h} M_{hi}}{n_h x M_{hi}}$$

where:

W_{hi} = basic expansion factor for all households in the i-th sample EA in the h-th first-level stratum or domain of estimation.

Note that a separate weight would have to be calculated for each sample EA. A Lotus spreadsheet can be developed for maintaining the sampling information for each sample EA and for calculating the weights.

This basic weight should be adjusted for non-interviews due to refusals, temporary absence (not-at-home), etc. This weight adjustment can be carried out at the EA level as follows:

$$W'_{hi} = W_{hi} x \frac{M'_{hi}}{M''_{hi}}$$

where:

W'_{hi} = adjusted weight for households in the i-th sample EA in the h-th stratum.

M'_{hi} = number of valid households enumerated by the PES in the i-th sample EA in the h-th stratum (excluding abandoned or vacant housing units). In other words, the number of households initially slated to be interviewed.

M''_{hi} = number of households **with completed PES interviews** in the i-th sample EA in the h-th stratum.

6.1 Considerations for Further Weight Adjustment

Sometimes there is a desire to further adjust the weights for the PES sample by ratios based on the census population by age-sex groups, as it is believed that this would reduce the sampling error of the PES population estimates for these groups. The problem with this approach is that it would violate the basic assumption in dual system estimation of independence between the census and PES estimates. This can be seen from the expected values of the following components of the dual system estimate. Recall that:

(4.1) $E(M) = abN_T$ (by independence of events)
(4.2) $E(N_1) = aN_T$
(4.3) $E(N_2) = bN_T$

where:

N_1 = population enumerated in the census

N_2 = estimate of population from PES sample

M = PES estimate of matched persons

N_T = true total population

a = probability of person being "captured" (enumerated) in the census

b = probability of person being "captured" in the PES

If the weights are ratio adjusted to make the estimate N_2 equal to N_1, then $E(N_2)$ would equal aN_T (that is, a=b); this would violate the assumption that a and b are independent.

7. PES Estimates

Some of the survey estimates will be in the form of totals, such as the weighted PES estimate of the matched population. The survey estimate of a total will be obtained as follows:

$$\hat{Y} = \sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{M''_{hi}} W'_{hi} x Y_{hij}$$

where:

$\hat{Y}$	=	weighted estimate of total for variable Y
L	=	number of strata (Colombo City; Other Urban; Estate; Other Rural)
Y_{hij}	=	value of variable Y for the j-th sample household (or person) in the i-th sample EA in the h-th stratum
n_h	=	the number of sample EAs in stratum h
W'_{hi}	=	adjusted weight for households (and persons) in the i-th sample EA of stratum h.

In the case of the survey estimate of the total number of persons with a particular characteristic, the variable Y may be defined as follows:

$$Y \begin{cases} = 1 \text{ if the person has the particular characteristic} \\ = 0 \text{ otherwise} \end{cases}$$

Other survey estimates may be in the form of ratios, defined as follows:

$$\hat{R} = \frac{\hat{Y}}{\hat{X}}$$

where $\hat{Y}$ and $\hat{X}$ are weighted total estimates, calculated as specified above.

The principal PES estimates are in the form of ratios such as (M/T_p), where the numerator and denominator are estimated from the same sample and have a high correlation which reduces the sampling error of the ratio estimate. Also, a post-stratification of the PES subpopulation by age-sex groups can be used to improve the estimation. The age groups that are recommended are:

0 to 4 years
5 to 9 years
10 to 19 years
20 to 29 years
30 to 39 years
40 to 49 years
50 to 59 years
60 years and above

Thus, dual system estimates of coverage and total population would be computed for each of the 16 age-sex groups within each domain of estimation. These would be combined to obtain estimates at higher levels of aggregation. The allocation of the sample to the different domains of estimation would give us enough sample of persons per post-stratification group, although the sample size would vary by age-sex group.

8. Estimation of Coverage Rates and Total Population

Since different groups of the population may have quite different rates of coverage, the estimation can be improved by post-stratifying the population in the PES sample into homogeneous groups. The PES sample design includes a stratification (Colombo City; other urban; estate; other rural) which is efficient for the estimation of coverage, since the coverage rates would vary by these geographic areas. It also appears ideal to post-stratify the sample by age-sex groups within each domain of estimation, since the coverage rates would also vary by age-sex groups.

Following the determination of the matching status of each person in the PES sample and the assignment of the appropriate weight to the record for each housing unit, the components of the dual system estimate of coverage for each age-sex category within geographic strata (domain of estimation) would be computed as follows:

$\hat{N}_{(NM)_a} = \sum W_i (NM)_a$ = PES estimate of non-movers in age-sex category *a* within stratum *h*

$\hat{N}_{(IM)_a} = \sum W_i (IM)_a$ = PES estimate of in-movers in age-sex category *a* within stratum *h*

$\hat{N}_{(OM)_a} = \sum W_i (OM)_a$ = PES estimate of out-movers in age-sex category *a* within stratum *h*

$\hat{N}_{(TP)_h} = \hat{N}_{(NM)_a} + \hat{N}_{(IM)_a}$ = PES estimate of total population in age-sex category *a* of stratum *h*

$\hat{M}_{(NM)_a} = \sum_i W_i (MNM)_a =$ estimate of matched non-movers (enumerated in both the PES

and the census) in age-sex category *a* within stratum *h*

$\hat{M}_{(OM)_a} = \sum_{i \in H} W_i (MOM)_a$ = Estimate of matched out-movers in age-sex category *a* within stratum *h*

$\hat{M}_{(IM)_a} = \hat{N}_{(IM)_a} \dfrac{\hat{M}_{(OM)_a}}{\hat{N}_{(OM)_a}}$ Estimate of in-movers for age-sex category *a* within stratum *h*

$\hat{M}_{(TP)_a} = \hat{M}_{(NM)_a} + \hat{M}_{(IM)_a}$ = estimate of total matched number of persons enumerated in both the PES and the census for age-sex category *a* within stratum *h*

W_i' = final weight for the i-th sample housing unit

$(NM)_a$ = number of non-movers in age-sex category *a* in the i-th sample housing unit

$(IM)_a$ = number of in-movers in age-sex category *a* in the i-th sample housing unit

$(OM)_a$ = number of out-movers in age-sex category *a* in the i-th sample housing unit

$(MNM)_a$ = number of <u>matched</u> non-movers in age-sex category *a* in the i-th sample housing unit

$(MOM)_a$ = number of <u>matched</u> out-movers in age-sex category *a* in the i-th sample housing unit

The PES estimate of the coverage rate for age-sex category *a* in stratum *h* would be given by:

$$\frac{\hat{M}_{(TP)_a}}{\hat{N}_{(TP)_a}}$$

The corresponding ***dual system estimate of the total population in this group*** $\left(\hat{N}_{(DTP)_A}\right)$ would be:

$$\hat{N}_{(DTP)_a} = N_{C_A} \frac{\hat{N}_{(TP)_a}}{\hat{M}_{(TP)_a}}$$

where:

N_{C_a} = population enumerated in the census for age-sex category *a* in stratum *h*

The estimate of total population for stratum *h* $\left(\hat{N}_{T_h}\right)$ would be:

$$\hat{N}_{T_h} = \sum_a \hat{N}_{(DTP)_a}$$

In the same way, the stratum totals would be summed to obtain the national total:

$$\hat{N}_T = \sum_h \hat{N}_{T_h}$$

The E-sample will be used to estimate the census erroneous enumeration. Consequently, the dual system estimate of total population for age-sex category *a* in stratum *h* would be adjusted as follows:

$$\hat{N}'_{T_a} = \left(\hat{N}_{C_a} - \hat{N}_{E_a}\right)\frac{\hat{N}_{(TP)_a}}{\hat{M}_{(TP)_a}}$$

where:

$\hat{N}_{E_a} = \sum_i W_i\left(G_{ai} + E_{ai} + D_{ai} + I_{ai}\right) =$ E-sample estimate of the total number of persons erroneously enumerated in the census for age-sex category *a* in stratum *h*. G_{ai}, E_{ai}, D_{ai}, and I_{ai} are the components of the erroneous census enumeration count for age-sex category *a* in the i-th housing unit in the E-sample.

The estimate $\hat{N}_{E_a}$ can be improved through ratio estimation, as follows:

$$\hat{N}'_{E_a} = N_{C_a}\left(\frac{\hat{N}_{E_a}}{\hat{N}_{C_a}}\right)$$

where:

$\hat{N}_{C_a} = \sum_i W_i N_{ai}$ = E-sample estimate of population enumerated in the census

9. Calculation of Variances for the PES Estimates

In analyzing the PES results, it is important to calculate the precision of the sample estimates. The standard error, or square root of the variance, is used to measure the sampling error, although it may also include a small part of the nonsampling error. The variance estimator should take into account the different aspects of the sample design, such as the stratification and clustering. The CBS staff was given a demonstration on how to use CENVAR, the variance calculation software which is part of the Integrated Microcomputer Processing System (IMPS) developed by the International Programs Center, Population Division, U.S. Census Bureau. This software, which uses the ultimate cluster variance estimator, can be utilized to calculate variances for survey estimates based on a stratified multistage sample design such as that for the PES.

The following formulas are used for calculating the ultimate cluster variance estimates for totals and ratios:

(1) Variance for the survey estimate of a total:

(9.1)
$$VAR(\hat{Y}) = \sum_{h=1}^{L} \left[\frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h} \right)^2 \right]$$

where:

n_h = number of sample EAs in stratum h

L = number of strata (domains of estimation)

$\hat{Y}_{hi} = \sum_{j=1}^{M_{hi}} W_{hij} Y_{hij}$ = weighted total for variable Y in the i-th sample EA in stratum h

M_{hi} = number of sample housing units in the i-th sample EA in stratum h

W_{hij} = final weight for the j-th sample housing unit in the i-th sample EA of stratum h

$\hat{Y}_h = \sum_{i=1}^{n_h} \hat{Y}_{hi}$ = Total for Y in stratum h

$\hat{Y} = \sum_h \hat{Y}_h =$ Estimate of a total for the entire population

(2) Variance for the survey estimate of a ratio:

$$(9.2) \quad VAR(\hat{R}) = VAR\left(\frac{\hat{Y}}{\hat{X}}\right) = \frac{1}{\hat{X}^2}\left[VAR(\hat{Y}) + \hat{R}^2 x VAR(\hat{X}) - 2\hat{R}Cov(\hat{X}, \hat{y})\right]$$

where:

$$(9.3) \quad COV(\hat{Y}, \hat{X}) = \sum_{h=1}^{L}\left[\frac{n_h}{(n_h - 1)}\sum_{i=1}^{n_h}\left(\hat{X}_{hi} - \frac{\hat{X}_h}{n_h}\right)\left(\hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h}\right)\right]$$

$Var(\hat{Y})$ and $Var(\hat{X})$ are variances of survey estimates of totals, calculated using the formula specified previously in (9.1).

Annex 1 presents the tabulation plan and the list of steps involved in developing the variance specifications for CENVAR. It will be necessary to set up a special file for CENVAR with a record for each PES sample housing unit, with the stratum number, EA number, weight, classification variables and other variables corresponding to the numerator and denominator of each ratio for which the variance will be calculated. Given that some PES estimates will be made for age-sex groups, it would be necessary to develop a CENVAR file for PES sample persons, with age-sex groups and matching status as classification variables.

In order for the estimates from the PES to correspond to the estimation requirements of CENVAR, sometimes it will be necessary to calculate the variance of a "proxy" estimate as an approximation to the variance of the actual estimate. Since the estimate $\hat{M}$ includes a component for movers based on ratio estimate (using procedure C), the PES coverage estimate $\hat{M} / \hat{N}_2$ and the corresponding estimate $\hat{N}_T$ would not be simple ratio estimates. To calculate the actual variance of these estimates would involve a very complicated estimation procedure or the use of a variance estimator based on replications (such as random half-samples or jackknife estimators). However, since the component of $\hat{M}$ related to movers should be relatively small, it is possible to define a simpler "proxy" estimate for $\hat{M}$ which would be very close. This would involve using the number of matched out-movers to estimate the component of $\hat{M}$ related to movers, which corresponds to the estimation based on procedure A. In this case, the estimate $\hat{M}_{(TP)}$ defined above would be changed for purposes of variance computation to the following:

$$\hat{M}'_{(TP)_a} = \hat{M}_{(NM)} + \hat{M}_{(OM)_a} = \sum_i W_i[(MNM)_a + (MOM)_a]$$

(Each term was defined in section 8 above)

In order to be consistent with the estimate $\hat{M}'_{(TP)_a}$ based on procedure A, the estimate $\hat{N}_{(TP)_a}$ should also be changed to:

$$\hat{N}'_{(TP)_a} = \hat{N}_{(NM)_a} + \hat{N}_{(OM)_a} = \sum_{i \varepsilon H} W_i[(NM)_a + (OM)_a]$$

(The definition of each term is provided in section 8 above)

The coefficient of variation of $(\hat{M}' / \hat{N}'_2)$ (using the more general terms) could then be calculated as the approximate C.V. of $(\hat{M} / \hat{N}_2)$. This approach will also make it possible to compare the coverage estimates based on procedure C with those based on procedure A. Past experience indicates that procedure A tends to underestimate the number of movers, since it is more difficult to enumerate out-movers than in-movers.

In the same way, the following proxy estimate can be used for purposes of calculating the variance of $\hat{N}_T$, the dual system estimate of total population:

$$\hat{N}'_T = \frac{N_1 N'_2}{\hat{M}'}$$

The C.V. of $\hat{N}'_T$ can be expressed as follows:

$$C.V.(\hat{N}'_T) = \frac{\sqrt{V(\hat{N}'_T)}}{\hat{N}'_T} = \frac{\sqrt{N_1^2 V\left(\frac{\hat{N}'_2}{\hat{M}'}\right)}}{N_1 \frac{\hat{N}'_2}{\hat{M}'}} = C.V.\left(\frac{\hat{N}'_2}{\hat{M}'}\right)$$

Therefore, the C.V. of the survey ratio estimate $(\hat{N}'_2 / \hat{M}')$ can be calculated to obtain the approximate C.V. of $\hat{N}_T$.

If an E-sample is used to estimate the census erroneous enumeration $\hat{N}_E$, the census total population (N_1) would be adjusted by subtracting $\hat{N}_E$. The corresponding dual system estimate of the total population estimate would be:

$$\hat{N}_{T(E)} = \frac{\left(N_1 - \hat{N}_E\right)\hat{N}_2}{\hat{M}}$$

For the purpose of calculating the variance of this estimate, the term $\left(N_1 - \hat{N}_E\right)$ can be treated as a constant; this assumption should not have much effect on the variance estimate, since $\hat{N}_E$ should be relatively small. Therefore, the C.V. of $\hat{N}_{T(E)}$ would also be approximated by C.V. $\left(\hat{N}_2' / \hat{M}'\right)$. However, the variance of $\hat{N}_E$ could also be calculated with CENVAR, using a separate E-sample file.

The CENVAR variance will not take into account the post-stratification by age-sex groups within each geographic area involved in the estimation procedures, since these groups are not statistically independent. Therefore, when dual system estimates of total population are aggregated across age-sex groups within a geographic area, it would be necessary to calculate the C.V.'s of "proxy" estimates which ignore the age-sex post-stratification; such proxy estimates should be fairly close to the corresponding survey estimates. For example, the estimate of total population for stratum h, presented before in section 8, would be:

$$\hat{N}_{T_h} = \sum_a \hat{N}_{(DTP)_a} = \sum_a \left(N_{C_a} \frac{\hat{N}_{(TP)_a}}{\hat{M}_{(TP)_a}} \right)$$

The definition of these terms are presented in section 8 above.

The following "proxy" estimate would be used for the purpose of calculating the approximate C.V. of $\hat{N}_{T_h}$:

$$\hat{N}_{T_h} = \frac{N_{C_h} \hat{N}_h}{\hat{M}_h}$$

where:

$$N_{Ch} = \sum_a N_{C_a}$$

$$\hat{N}_h = \sum_a \hat{N}_{(TP)_a}$$

$$\hat{M}_h = \sum_a \hat{M}_{(TP)_a}$$

The C.V. of $\hat{N}_{T_h}$ would be estimated by calculating the C.V. of $\left(\hat{N}_h \,/\, \hat{M}_h\right)$ using CENVAR. As mentioned before, the estimates $\hat{N}_h$ and $\hat{M}_h$ would be based on the definitions corresponding to procedure A.

10. Instructions for the Reconciliation Visit

It is very important to carry out the reconciliation visit after the matching operation is finished in order to resolve the "possible match" cases and those enumerated in the census but not in the PES. Note that the information obtained from the reconciliation visits will be used to update PES information. The cases for which reconciliation visits are necessary are:

10.1 Introduction

1. Corresponding census questionnaire for a PES questionnaire does not exist

   In this case, search adjacent EAs to locate census questionnaire. PES households or persons matched with corresponding census records in adjacent EAs would be treated as matched for estimation purposes.

   a. Census questionnaire was not found

      i. Verify if all the household members were usual members at the time of the census (non-mover or out-movers). If all the household members were usual members at the time of the census, they must be declared to be not matched and omitted.

      ii. If the household members arrived after the census, they must be declared to be in-movers.

   b. Census questionnaire was found

      Proceed with the matching and determine census enumeration status and matching status.

2. Corresponding PES questionnaire for a census questionnaire does not exist

In this case, go back to the census enumerated household.

i. If the household cannot be found, ask neighbors to determine if that household existed at the time of the census. If it did exist at the time of the census, all persons in that household must be declared to be out-movers and correctly enumerated. If the neighbors provide the names of the members of the census household and the latter can be matched to the census questionnaire, these persons are considered to be matched.

ii. If the household cannot be found and neighbors cannot help determine the enumeration status, we must assume that it does not exist (maybe it was fabricated by the census enumerator). All the persons in that household must be declared to be erroneously enumerated (fabrications) and non-matched. These cases are, however, rare.

3. Census questionnaire with persons that do not appear on corresponding PES questionnaire

Many of these cases could apply to students who were enumerated at their parents' house but were absent at the time of the PES enumeration. We must determine whether these out-movers were usual members of the household at the time of the census enumeration.

4. Census questionnaire which does not contain people that were enumerated in the PES

This case may occur if the respondent made a mistake in considering a household member as a non-mover. For those people who were not enumerated in the census, we must determine whether the person arrived to the household after the census (in-mover) or whether the person was not a usual household member. There are cases in which the husband does not live with his wife because of work reasons. In these cases, we must determine if the husband is a usual member of another household. We may also find cases of children born after the census who must be classified as "born after" and not as non-movers.

5. Difficult to determine the enumeration status due to lack of information

In these cases, we must try our best to identify these people and discuss the situation with a supervisor.

10.2 Organization of Questionnaires

GROUP A: Case 1, Case 3, Case 4, and Case 5

By grouping cases 1, 3, 4 and 5, we have together all the cases for which we have questions concerning the PES.

Organize the questionnaires by household number. Mark on the household listing sheet the households that need a reconciliation visit.

***GROUP B*: Case 2**

This is the only case for which we do not have a PES questionnaire.

Organize the questionnaire by census household number. Mark on the census household listing sheet the households that need a reconciliation visit.

10.3 Examples

1. We have found 5 people in a PES household, of whom 4 are non-movers and 1 is an out-mover. All 5 persons were matched. However, there is also a person enumerated in the census who does not appear in the PES. In this case, we must determine if this person was a usual member of the household at the time of the census.

   If the person was a usual member and a non-mover, then the person is considered to be correctly enumerated and a non-mover.

   If the person was a usual member but has left since the census, the person is considered to be correctly enumerated and an out-mover.

   If the person was not a usual member of the household (visitor) at the time of census enumeration, the person was erroneously enumerated.

2. We have found 5 people in a PES household, of whom 4 are non-movers and 1 is an out-mover. The head of the household has not been matched (that is, he has not been enumerated in the census), but the other members have all been matched. There are also two other people who have been enumerated in the census but not in the PES.

   Regarding the head of the household, we must determine if he was a usual member at the time of the census. It could happen that the head of the household (in our case, the husband) does not usually live with his wife because of his work. We must determine if he has been enumerated in the census somewhere else (at his usual place of residence). If he is not a usual member of this household, we must declare so in the observation area of the questionnaire and cross him out. If he has not been enumerated in the census because he arrived after the census enumeration, we must write this in the observation area and register him as an in-mover instead of a non-mover.

   If the head of the household was a usual member at the time of the census enumeration, he must be declared as a non-match.

   Regarding the two other people, we must follow the guidelines outlined in the previous example. We must determine if they were correctly enumerated in the census.

3. We have found 7 people in a PES household and all of them are non-movers. There is a woman without relationship to head of the household who is 21 years old and who has not been matched; there is also an infant of age 00 who has not been matched.

Regarding the 21 year-old woman, we must determine if she is a usual member of the household. Did she arrive to the household after the census? Or maybe she was erroneously enumerated?

Regarding the infant, was he born after the census? If that is the case, he must be declared as a "born after". If that is not the case, was the infant a usual member of the household at the time of census enumeration? If so, he must be declared as a non-match. Otherwise, did he arrive after the census? If so, he must be declared as an in-mover.

4. In the case where there is a PES questionnaire but no corresponding census questionnaire, we must try to locate the census questionnaires in adjacent EAs. It may happen that there were boundary problems. If the census questionnaire is found, we have a match. Otherwise, we have a non-match.

5. In the case where we have a census questionnaire but no corresponding PES questionnaire, we must try to find the household in the field. If the household is found, we must verify that all the members were usual members of the household at the time of the census enumeration. If that is the case, the members are declared correctly enumerated and non-movers.

   If the household does not exist, we must try to find a neighbor that can tell us whether the members of the household were there at the time of census enumeration. If that is the case, the members are declared as correctly enumerated and out-movers.

   If the household does not exist (rare cases), we must write this in the observation area and declare all the members of the household as erroneously enumerated.

Annex 1

TABULATION PLAN FOR A POST-ENUMERATION SURVEY

A. Data Entry

The CENTRY module of IMPS should be used for the PES data entry. Based on the questionnaire design, there are three types of records per person. Record type 1 contains geographic identification and the characteristics of persons enumerated in the PES. Record type 2 contains geographic identification and the characteristics of out-movers and record type 3 contains geographic identification and the characteristics of persons enumerated in the census who were not found in the PES.

The CONCOR module of IMPS should be used to edit the PES data. Its purpose is to detect missing and invalid values, particularly inconsistencies among the PES enumeration status (also known as the moving status), the match status, and the census enumeration follow-up status from the reconciliation visits.

The QUICKTAB module of IMPS should be used during the pre-tabulation phase for reviewing frequencies per questionnaire item. This is very useful for contrasting actual frequencies with expected ones. Hence, this review would help identify certain invalid or blank entries, for example, cases without a final match status.

The CENTS module of IMPS should be used for the tabulation of the coverage measures (A Series). The program takes into account the sampling weights to calculate estimates of population parameters from the sample, based on the tabulation plan outlined below. The weights must be calculated prior to the tabulation of characteristics.

The CENVAR module of IMPS should be used to calculate the standard errors (E Series).

B. Convention for Table Numbering

1. Series
   - A = Coverage Analysis
   - B = Content Analysis
   - E = Standard Errors and Confidence Intervals for Estimates of Coverage Measures in Series A
   - I = Initial Tables for Calculation of Coverage Measures in Series A

2. Only Table Series A, B, and E are for publication; tables in Series I are only for input into Series A.

3. Tables in Series I, A, and Tables B1-B3 (of Series B) are calculated using the CENTS module of the Integrated Microcomputer Processing System (IMPS) software. Tables B4-B7 (of Series B) are calculated in spreadsheet software, using as input *unweighted* CENTS tables B1-B3.

4. The final tables are all weighted. While Tables B1-B3 are weighted for publication, Tables B4-B7 are based on the *unweighted* version of B1-B3.

5. All totals are rounded to whole numbers and all rates (and indices) are expressed as percentages rounded to one decimal place.

6. The sex/age group estimation cells should be limited to certain age/sex categories and should be formed taking into account: homogeneity with respect to omission rates, types of individuals for which separate estimates are desired (for example, infants or young males entering the labor force who might have high mobility), the number of observations in the resulting cell, and the need for the groups to coincide with the census categories. **All tables that cross-tabulate sex and age, should have the age groups as rows and the sex as columns.**

C. Range and Consistency Checks

It is necessary to detect and correct the inconsistencies listed below before tabulation can proceed. A quick review can be performed with the QUICKTAB module of IMPS by crossing the enumeration status with the match status. A more elaborate CONCOR program can be developed to identify the specific records involved.

NOTE: It is assumed at this stage that all the PES phases–enumeration, matching, reconciliation visits, data entry, and verification–have been completed with a reasonable degree of quality. The editing phase presumes only to correct a small number of errors remaining despite all the efforts made in the preceding phases. The correction of errors at this stage CANNOT substitute for the quality of the preceding operations.

PES MOVING STATUS		MATCH STATUS
❑ non-mover	←	❑ matched
❑ out-mover		❑ non-matched
❑ in-mover	→	❑ in-mover
❑ born-after		❑ born after

1. The moving status (also called PES enumeration status), the match status, and the follow-up status can never be blank. The technical staff must examine the questionnaires concerned to try to determine the correct response.

2. If the moving status indicates the person is an in-mover, then the only response possible in the match status is "in-mover."

3. If the moving status indicates the person is born after the census, then the only response possible in the match status is "born after the census."

4. If the moving status indicates the person is a non-mover or an out-mover, then the only responses possible in the match status are "matched" or "non-matched."

D. Table Titles by Series

I Series - Initial Tables (all weighted). Repeat for each domain of estimation.

I1 **a** = Estimated Number of Non-Movers, by Sex and Age Group (P-sample)
I2 **b** = Estimated Number of Out-Movers, by Sex and Age Group (P-sample)
I3 **c** = Estimated Number of In-Movers, by Sex and Age Group (P-sample)
I4 **d** = Estimated Number of Matched Non-Movers, by Sex and Age Group (P-sample)
I5 **e** = Estimated Number of Matched Out-Movers, by Sex and Age Group (P-sample)
I6 **f** = Estimated Number of Matched In-Movers, by Sex and Age Group (P-sample)

In procedure C, the total number of matched in-movers cannot be calculated directly given that no match is attempted for the in-movers in the sample. Under the hypothesis of a closed population, the "out-movers" and the "in-movers" constitute the same group: the "movers."

Therefore, an assumption can be made that, in the universe, the match rate for in-movers would be the same as that for out-movers (total number of matched out-movers divided by the total number of out-movers). Therefore, the total number of matched in-movers in the universe is estimated indirectly by multiplying the above ratio by total number of in-movers in the universe.

Similarly, its complement f', the total number of non-matched in-movers is estimated by applying the complement of the match rate for out-movers to the total number of in-movers. In symbols:

Total number of matched in-movers = $f = \frac{e}{b} c$

and total number of non-matched in-movers = $f' = \left(1 - \frac{e}{b}\right) c$

I7 **g** = Estimated Number of Erroneous Inclusions in the Census, by Sex and Age Group (E-sample)

I8 **h** = Estimated Number of Correctly Enumerated Persons Missed in the PES, by Sex and Age Group (E-sample)

I9 Match Rate for Out-Movers, by Sex and Age Group

The match rate for out-movers = (e/b)

The "matched" population is given by the total number of matched non-movers (d) plus the estimated total number of matched in-movers (f) in the universe.

MATCHED = MATCHED NON-MOVERS + ESTIMATED MATCHED IN-MOVERS
*= [d + f] = [d + (e/b * c)]*

A Series - Estimates of Coverage Measures. Repeat for each domain of estimation.

A1 E-sample Estimate of the Population Enumerated in the Census, by Sex and Age Group

This is the same as T_C' (page 10). This can also be obtained as follows:

Estimate of census population = $d + e + g + h$

A2 P-sample Estimate of the Total Population, by Sex and Age Group

This represents the sum of the total number of non-movers and in-movers in the population.

PES Estimate of population = $a + c$

A3 Census Omissions and Census Omission Rate, by Sex and Age Group

There are two ways of calculating the census omissions. The first involves subtracting the matched population from the PES estimate of the total population to obtain the total "non-matched" (omitted) population:

Census Omissions = PES Population - Matched = $[a + c] - [d + f] = [a + c] - [d + (e/b * c)]$

The second way of calculating census omissions is as follows:

Census Omissions = NON-MATCHED NON-MOVERS + NON-MATCHED IN-MOVERS
$= d' + f'' = d' + [((1 - e/b) * c]$

The census omission rate is the missed population relative to the PES estimate of total population.

$$\textit{Omission Rate} = \frac{Omissions}{PES\ Population}$$

The coverage rate is the complement of the omission rate:

$$\text{Coverage Rate} = \frac{Matched\ Population}{PES\ Population}$$

A4 Census Erroneous Inclusions and Census Erroneous Inclusion Rate, by Sex and Age Group

The estimated total number of census erroneous inclusions is given by item *g* above. This variable includes fabrications, duplications, and geographic misallocations, etc. As mentioned before, the purpose of the E-sample is to provide an estimate for this variable in order to permit a correction in the dual-system estimate of the true population.

Erroneous Inclusions = g

The census erroneous inclusion rate is equal to the total number of persons erroneously included in the census (g) relative to the E-sample estimate of the census population.

$$Erroneous\ Inclusion\ Rate = \frac{Erroneous\ Inclusions}{Census\ Population}$$

A5 Preliminary Dual-System Estimate of the True Population, by Sex and Age Group
See section 8 in the body of the document.

A6 Net Census Coverage Error and Net Census Coverage Error Rate, by Sex and Age Group

The net coverage error is the difference between what should have been counted (estimate of the true population obtained in A5) and what was counted (census population)

Net Coverage Error = True Population - Census Population

The net coverage error rate is the total net error relative to the dual system estimate of the true population. This measure constitutes the single most important indicator of the quality of the census coverage.

$$Net\ Error\ Rate = \frac{True\ Population - Census\ Population}{True\ Population}$$

A7 Gross Census Coverage Error and Gross Census Coverage Error Rate, by Sex and Age Group

The gross coverage error is the sum of the errors in the census enumeration, that is, the omissions and the erroneous inclusions. It is an indicator of the operational quality of the census enumeration.

Gross Coverage Error = Omissions + Erroneous Inclusions

The gross coverage error rate per unit enumerated is the absolute gross error relative to the census enumerated population:

$$Gross\ Error\ Rate = \frac{Omissions + Erroneous\ Inclusions}{Census\ Population}$$

A8 Adjustment Factors, by Sex and Age Group

The estimated census totals in table A5 are based on the E-sample. When the actual census figures for the household population become available, they can be used to obtain an improved estimate of the true population. This is accomplished by using a ratio estimator of total at each

estimation cell. The ratio estimator has a better precision because it takes advantage of the known census counts. See section 8 for the adjustment factors.

A9 Adjusted Population, by Sex and Age Group

Adjusted Population = Census Population + Omissions - Erroneous Inclusions

B Series - Content Error Estimates (CENTS, weighted)

B1 Sex as Reported in Census and As Reported in PES, for Matched Persons
B2 Age Group as Reported in Census and As Reported in PES, for Matched Persons
B3 Relationship to Head of Household As Reported in Census and As Reported in PES, for Matched Persons

For the following tables, use unweighted data only and spreadsheet software (Lotus, Excel)

B4 Net Difference Rate and Index of Inconsistency for Sex
B5 Net Difference Rate and Index of Inconsistency for Age Group
B6 Net Difference Rate and Index of Inconsistency for Sex and Age Group
B7 Net Difference Rate and Index of Inconsistency for Relationship to Head of Household

Example of Table B1.
Net Difference Rate and Indices of Inconsistency for Sex

Category	Number of Consistent Cases	Number of cases in census	Number of cases in PES	Net Difference Rate			Index of Inconsistency		
				Rate	95% C. I.		Index	95% C. I.	
					Lower	Upper		Lower	Upper
Male	610	625	617	.63	-0.13	1.39	3.47	2.27	5.30
Female	637	644	652	-.63	-1.39	0.13	3.47	2.27	5.30
Undetermined	0	0	0	NA	NA	NA	NA	NA	NA
TOTAL	1,247	1,269	1,269	-	-	-			
Aggregate Index							3.47	2.27	5.30
Gross Difference Rate		1.73%	(Proportion off-diagonal cases)						
Agreement Rate		98.27%							

1.73% = (# cases in PES - # of consistent cases) / (#cases in PES)
98.27% = (# of consistent cases / # of cases in PES)

E Series - Standard Errors and Confidence Intervals for Estimates of Coverage Measures (CENVAR)

Standard Errors and Confidence Intervals for PES Estimates of:

E0	Summary of Estimates of Coverage Measures at the Domain of Estimation Level
E1	E-Sample Estimate of the Population Enumerated in the Census, by Sex and Age Group
E2	P-Sample Estimate of the PES Population, by Sex and Age Group
E3A	Census Omissions, by Sex and Age Group
E3B	Census Omission Rate, by Sex and Age Group
E4A	Census Erroneous Inclusions, by Sex and Age Group
E4B	Census Erroneous Inclusion Rate, by Sex and Age Group
E5	Preliminary Dual-System Estimate of the True Population, by Sex and Age Group
E6A	Net Census Coverage Error, by Sex and Age Group
E6B	Net Census Coverage Error, by Sex and Age Group
E7A	Gross Census Coverage Error, by Sex and Age Group
E7B	Gross Census Coverage Error Rate, by Sex and Age Group
E8	Final Dual-System Estimate of the True Population, by Sex and Age Group

E. PES Data Dictionary Specifications (CENVAR)

Variable Name	Data Type	Position	Length	Value and Value Name
STRATUM	N	1	1	1 = Nairobi 2 = Central 3 = Coast 4 = Eastern 5 = North-Eastern 6 = Nyanza 7 = Rift Valley 8 = Western
EA	N	2-3	2	
WEIGHT	N	4-10	7 (4 dec.)	
RELATIONSHIP-P-SAMPLE	N	11	1	1 = Head 2 = Spouse 3 = Son 4 = Daughter 5 = Brother/Sister 6 = Father/Mother 7 = Other relative 8 = Non-relative 9 = NS/DK
RELATIONSHIP-E-SAMPLE	N	12	1	1 = Head 2 = Spouse 3 = Son 4 = Daughter 5 = Brother/Sister 6 = Father/Mother 7 = Other relative 8 = Non-relative 9 = NS/DK
SEX-P-SAMPLE	N	13	1	1 = Male 2 = Female
SEX-E-SAMPLE	N	14	1	1 = Male 2 = Female
AGE-P-SAMPLE	N	15-16	2	1 = 0 to 4 years 2 = 5 to 9 years 3 = 10 to 19 years 4 = 20 to 29 years 5 = 30 to 39 years 6 = 40 to 49 years 7 = 50 to 59 years 8 = 60 years and above

AGE-E-SAMPLE	N	17-18	2	1 = 0 to 4 years 2 = 5 to 9 years 3 = 10 to 19 years 4 = 20 to 29 years 5 = 30 to 39 years 6 = 40 to 49 years 7 = 50 to 59 years 8 = 60 years and above
NON-MOVER	N	23	1	1 = Non-Mover 0 = Otherwise
OUT-MOVER	N	24	1	1 = Out-Mover 0 = Otherwise
IN-MOVER	N	25	1	1 = In-Mover 0 = Otherwise
MATCHED-NM	N	26	1	1 = Matched-NM 0 = Otherwise
MATCHED-OM	N	27	1	1 = Matched-OM 0 = Otherwise
MATCHED-IM (Obtain the match rate for out-movers by sex and age group, which is expressed as a percentage, and apply that rate divided by 100 to a person who is an in-mover)	N	28-31	4 (3 dec.)	0000:1000
TOTAL MATCHED= $\hat{M}$ = (M-NM) + (M-IM)	N	32-35	4 (3 dec.)	1000:2000
ERRONEOUSLY INCLUDED	N	36	1	1 = Erron. Included 0 = Otherwise
CORRECTLY ENUMERATED (enumerated in the census but missed in the PES)	N	37	1	1 = Corr. Enumerated 0 = Otherwise
CENSUS POPULATION = (M-NM) + (M-OM) + (ERR INC)+ (CORR. ENUM.)	N	38	1	1 0
PES POPULATION = Non-Mover + In-Mover	N	39	1	1 0
OMISSIONS = PES-POP - (M-NM) - (M-IM) = PES-POP - (TOTAL MATCHED)	N	40-43	4 (3 dec.)	0000:1000

THE 1999 POPULATION AND HOUSING CENSUS

ANNEX I

SUMMARY OF COVERED AND MATCHED HOUSEHOLDS

URBAN

DISTRICT	EA NAME	EA CODE	CENSUS HHLDS	PES HHLDS	ATCHE HHDS	PES %	CENSUS %
NAIROBI	KIMATHI	101050401	155	152	121	79.61	78.06
NAIROBI	KALOLENI	101020102	214	217	151	69.59	70.56
NAIROBI	MABATINI	101010302	310	280	214	76.43	69.03
NAIROBI	HIGHRIDGE	101060303	176	155	125	80.65	71.02
NAIROBI	DANDORA PHASE II	101040601	34	47	21	44.68	61.76
NAIROBI	EASTLEIGH NORTH	101050102	224	200	136	68.00	60.71
NAIROBI	MUCHAGUCHA	101060403	172	165	103	62.42	59.88
NAIROBI	NJIKU 'A'	101070201	157	152	98	64.47	62.42
NAIROBI	MATHARE NORTH	101030503	253	246	146	59.35	57.71
NAIROBI	SAVANNAH PH.IV}	101040302	73	66	45	68.18	61.64
NAIROBI	PRECIOUS BLOOD	101070401	230	204	137	67.16	59.57
NAIROBI	GATWIKIRA	101080701	87	89	53	59.55	60.92
NAIROBI	RIRUTA SATELITE	101070501	165	151	100	66.23	60.61
NAIROBI	HIGH-RIDGE	101030202	153	140	89	63.57	58.17
NAIROBI	MARIGUINI	101020502	107	135	57	42.22	53.27
NAIROBI	KARANJA ROAD	101080102	134	123	72	58.54	53.73
NAIROBI	MAI-INI	101070401	117	110	73	66.36	62.39
NAIROBI	HURUMA	101010402	388	354	222	62.71	57.22
NAIROBI	GOLF COURSE	101070602	120	119	71	59.66	59.17
NAIROBI	BURU BURU PH.V	101020202	130	78	63	80.77	48.46
NAIROBI	LORESHO SOUTH	101060201	70	75	42	56.00	60.00
NAIROBI	MUKINDURI	101080201	398	342	194	56.73	48.74
NAIROBI	GITHURAI	101030401	143	147	64	43.54	44.76
NAIROBI	KWA REUBEN	101040202	185	177	89	50.28	48.11
NAIROBI	MUTHAIGA	101040302	93	84	48	57.14	51.61
NAIROBI	KIA MAIKO	101010401	209	206	89	43.20	42.58
NAIROBI	EASTLEIGH NORTH	101050102	168	156	70	44.87	41.67
NAIROBI	MATHARE NORTH	101030503	225	213	91	42.72	40.44
NAIROBI	JUJA RD. EST.(SEC II)	101010201	293	216	119	55.09	40.61
NAIROBI	EASTLEIGH SOUTH	101050201	80	87	32	36.78	40.00
NAIROBI	KAYOLE	101040401	276	357	135	37.82	48.91
NAIROBI	KIBERA	101080602	407	384	244	63.54	59.95
NAIROBI	KINGSTONE	101020402	578	490	219	44.69	37.89
NAIROBI	KAYOLE	101040401	618	461	231	50.11	37.38
NAIROBI	KAYOLE	101040401	296	370	110	29.73	37.16
NAIROBI	KOMAROCK	101040402	164	235	74	31.49	45.12
NAIROBI	SILANGA	101080103	331	300	203	67.67	61.33
NAIROBI	HURUMA	101010402	268	210	110	52.38	41.04
NAIROBI	KEY WEST 'B'	101070502	495	277	188	67.87	37.98
NAIROBI	VILLAGE 3 'A'	101010301	655	544	226	41.54	34.50
NAIROBI	LIGHT INDUSTRIES	101030101	203	200	76	38.00	37.44
NAIROBI	CLIFF	101030501	122	81	56	69.14	45.90
NAIROBI	KAYOLE	101040401	114	109	44	40.37	38.60
NAIROBI	KILIMANI	101060501	153	93	57	61.29	37.25
NAIROBI	KARIOBANGI SOUTH	101040701	86	36	29	80.56	33.72
NAIROBI	RIVER SIDE	101030102	75	71	24	33.80	32.00
NAIROBI	MOUNTAIN VIEW	101060402	467	292	149	51.03	31.91

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
NAIROBI	LAINI SABA	101080601	441	421	164	38.95	37.19
NAIROBI	KAIYABA	101020401	326	241	194	80.50	59.51
NAIROBI	MUTIRITHIA PH. 2	101030701	166	155	43	27.74	25.90
NAIROBI	KWA MUTISYA	101010303	106	93	26	27.96	24.53
NAIROBI	LUNGA LUNGA	101020402	300	184	105	57.07	35.00
NAIROBI	THOME V	101030601	449	367	79	21.53	17.59
NAIROBI	MUKURU KWA NJENGA	101040201	239	119	57	47.90	23.85
NAIROBI	MUKURU KWA NJENGA	101040201	181	181	32	17.68	17.68
NAIROBI	SINAI	101020402	719	579	197	34.02	27.40
NAIROBI	MUKURU KWA NJENGA	101040201	798	532	271	50.94	33.96
NAIROBI	INNERCORE	101040301	1424	1364	614	45.01	43.12
NAIROBI	GITHEMBE	101070502	506	423	244	57.68	48.22
NAIROBI	MADARAKA ESTATE	101080501006	NOT MATCHING				
			16226	**14055**	**7136**	**50.77**	**43.98**
KIAMBU	HOSPITAL WARD	201010302	111	105	81	77.14	72.97
KIAMBU	MUUTI 'B'}	201010903	823	616	446	72.40	54.19
KIAMBU	GACHIE }	201010904	988	852	277	32.51	28.04
KIAMBU	KINGEERO }	201050501	594	507	308	60.75	51.85
			2516	**2080**	**1112**	**53.46**	**44.20**
KIRINYAGA	KERUGOYA TOWN	202020101	65	65	57	87.69	87.69
MURANG'A	KABWE	203010402	543	566	246	43.46	45.30
NYERI	EQUATER }	205020303	432	270	144	53.33	33.33
NYERI	KAMUYU	205070110	121	108	86	79.63	71.07
			553	**378**	**230**	**60.85**	**41.59**
THIKA	KIANDUTU	206050203	300	258	137	53.10	45.67
THIKA	MWANAMUKIA "D"	206060201	197	88	58	65.91	29.44
			497	**346**	**195**	**56.36**	**39.24**
KILIFI	MTWAPA CENTRE	301030102	122	96	64	66.67	52.46
KILIFI	KWA MWANGO	301010103	161	207	119	57.49	73.91
KILIFI	KALOLENI T.C.	301071101	317	243	125	51.44	39.43
			600	**546**	**308**	**56.41**	**51.33**
KWALE	VINGUJINI	302040201	219	294	126	42.86	57.53
MOMBASA	KING'ORANI }	304010103	110	114	91	79.82	82.73
MOMBASA	KIPEVU ESTATE	304040201	84	87	65	74.71	77.38
MOMBASA	MAJENKO MAPYA	304020101	142	185	96	51.89	67.61
MOMBASA	TUA TUGAWE	304020102	175	161	105	65.22	60.00
MOMBASA	MAGOSI	304040301	110	119	58	48.74	52.73
MOMBASA	ZIWA LA NGOMBE	304020202	189	175	87	49.71	46.03
MOMBASA	MIGOMBANI	304030101	169	139	90	64.75	53.25
MOMBASA	MISHOMORONI	304020103	196	127	62	48.82	31.63
MOMBASA	MISHOMORONI	304020102	2441	2195	585	26.65	23.97
MOMBASA	MIKOROSHONI	304020303	2092	2342	892	38.09	42.64
MOMBASA	GRANDI	304030201	889	833	467	56.06	52.53
MOMBASA	KWA CHARO/KWA MWA	304040401	1757	1571	530	33.74	30.17
MOMBASA	BANGLADESH	304040402	2068	1900	404	21.26	19.54
MOMBASA	CHAMUNYU	304040501	772	726	514	70.80	66.58

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
			11194	**10674**	**4046**	**37.91**	**36.14**
TAITA TAVET	BURA NDOGO "B"	305010301	111	149	66	44.30	59.46
TAITA TAVET	MAHOO	305010302	171	179	158	88.27	92.40
TAITA TAVET	BOMANI	305030704	80	98	61	62.24	76.25
			362	**426**	**285**	**66.90**	**78.73**
MALINDI	KISUMU NDOGO	307030102	141	135	83	61.48	58.87
EMBU	MBUNA/DALLAS/MUSLI	401010201	121	155	67	43.23	55.37
EMBU	KANDURI `B'	401050204	159	156	124	79.49	77.99
			280	**311**	**191**	**61.41**	**68.21**
KITUI	MOSQUITO	403030402	229	236	128	54.24	55.90
MACHAKOS	EASTLEIGH	405010102	102	84	54	64.29	52.94
MACHAKOS	KONGOWEA	405070101	104	83	59	71.08	56.73
MACHAKOS	GITHUNGURI/KATANI	#########	201	200	41	20.50	20.40
			407	**367**	**154**	**41.96**	**37.84**
MERU CENTR	KAAGA	408040102	788	760	419	55.13	53.17
MERU CENTR	GITORO	408050101	107	76	39	51.32	36.45
MERU CENTR	KIRINGA "B"	408050301	400	416	238	57.21	59.50
			1295	**1252**	**696**	**55.59**	**53.75**
MWINGI	KARIBANI	410010106	125	126	62	49.21	49.60
GUCHA(S. KIS	BOKIMAI TABAKA }	601040303	92	92	79	85.87	85.87
KISII CENTRA	ISANA	603010501	288	263	113	42.97	39.24
KISUMU	LOWER KAROMBO "D"	604030102	98	89	70	78.65	71.43
KISUMU	KODEYO ALARA	604030501	111	105	90	85.71	81.08
KISUMU	KAROMBO II/SANGA II	604030703	528	413	358	86.68	67.80
KISUMU	KANYAKWAR II	604031101	110	80	51	63.75	46.36
			847	**687**	**569**	**82.82**	**67.18**
MIGORI	SOWETO	606010101	99	174	75	43.10	75.76
N.KISII(NYAM	MOSOSA	607010204	187	186	173	93.01	92.51
N.KISII(NYAM	ITERESI	607050301	116	139	76	54.68	65.52
SIAYA	ASIRO URBAN }	609020403	111	118	63	53.39	56.76
KEIYO	KAPSIO 'C'	703020202	98	91	74	81.32	75.51
KAJIADO	MAGADI/RAILWAY QTR	704040103	285	288	155	53.82	54.39
KERICHO	KERICHO TOWN (KAMPI	705040102	230	222	139	62.61	60.43
KERICHO	KERICHO TOWN (INDUS	705040102	83	66	46	69.70	55.42

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
KOIBATEK	MAJI MAZURI 3	706010301	139	140	109	77.86	78.42
LAIKIPIA	THINGITHU PHASE II	707010102	106	101	64	63.37	60.38
MARAKWET	KAPSOWAR "A"\KAPSOW	708040201	427	320	95	29.69	22.25
NAKURU	SIGALAME(OSERIAN "D"	709140801	485	472	269	56.99	55.46
NAKURU	KALOLENI `C'	709160302	89	80	76	95.00	85.39
NAKURU	RONDA	709160401	20	83	48	57.83	240.00
NAKURU	MWICIRINGIRI	709140802	321	308	239	77.60	74.45
NAKURU	PANGANI	709160302	141	124	90	72.58	63.83
NAKURU	HIGHLANDS	709100302025					
NAKURU	KIPSIGIS/TUGEN	709160401	194	181	82	45.30	42.27
			1250	**1248**	**804**	**64.42**	**64.32**
TRANS NZOIA	UPPER LESSONS	714030102	545	510	331	64.90	60.73
TRANS NZOIA	TUWAN KALOLENI	714030203	125	105	57	54.29	45.60
			670	**615**	**388**	**63.09**	**57.91**
UASIN GISHU	MARURA	716020701	295	284	198	69.72	67.12
UASIN GISHU	KORIO/TURUNDI MASHA	716040602	233	132	87	65.91	37.34
UASIN GISHU	HURUMA	716060502	138	139	82	58.99	59.42
			666	**555**	**367**	**66.13**	**55.11**
BURET	CHEBILAT TOWN	718050501	306	348	103	29.60	33.66
BUNGOMA	SIO 'A' }	801040402	304	287	186	64.81	61.18
BUNGOMA	SINOKO 'B'	801060302	55	83	51	61.45	92.73
			359	**370**	**237**	**64.05**	**66.02**
BUSIA	BUSIDIBU	802030501	101	119	96	80.67	95.05
MT. ELGON	PERETTEN	803010202	47	55	40	72.73	85.11
KAKAMEGA	SHIPOALO	804070202	145	139	74	53.24	51.03
LUGARI	TOWNSHIP	805020302	83	91	76	83.52	91.57
TESO	OPOPONG	806020201	61	58	34	58.62	55.74
TESO	AIR STRIP	806040602	76	77	36	46.75	47.37
			137	**135**	**70**	**51.85**	**51.09**
VIHIGA	MULWANYI	807050401	115	104	89	85.58	77.39
BUTERE/MUM	ESHITUKHUMI	808020203	215	189	181	95.77	84.19
TOTAL			**80313**	**72757**	**36198**	**49.75**	**45.07**
RURAL							
KIAMBU	KARUIBANGI	201040702	123	124	118	95.16	95.93
KIAMBU	MBAU-INI	201040202	97	98	91	92.86	93.81
KIAMBU	GATHANGARI	201020303	284	287	246	85.71	86.62
KIAMBU	KAHUHA }	201020203	182	180	157	87.22	86.26

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
KIAMBU	NYAGA	201020106	148	148	126	85.14	85.14
KIAMBU	MBOCA	201050603	131	121	110	90.91	83.97
KIAMBU	CENTRAL	201010502	448	484	347	71.69	77.46
KIAMBU	RUKUBI	201050104	136	117	108	92.31	79.41
KIAMBU	KAMITI/BRIDGE	201010303	136	119	97	81.51	71.32
KIAMBU	RIRONI "A"	201030201	198	178	128	71.91	64.65
KIAMBU	KIGUMO	201020405	470	465	303	65.16	64.47
KIAMBU	KARIA/KAWAIDA	201011002	1155	1142	542	47.46	46.93
KIAMBU	GATHANGA }	201010801	529	487	371	76.18	70.13
KIAMBU	GICHAGI "B" }	201020103	187	180	8	4.44	4.28
KIAMBU	KARIRANA FACTORY	201030603	544	372	304	81.72	55.88
			4768	**4502**	**3056**	**67.88**	**64.09**
KIRINYAGA	KAGUMO	202030503	197	224	186	83.04	94.42
KIRINYAGA	MURUBARA	202040101	236	242	219	90.50	92.80
KIRINYAGA	KIMURI	202040404	221	221	205	92.76	92.76
KIRINYAGA	RURII	202040601	230	233	212	90.99	92.17
KIRINYAGA	KIMANDI	202030203	250	261	228	87.36	91.20
KIRINYAGA	KARIKO }	202010302	99	94	89	94.68	89.90
KIRINYAGA	KIRUGARA 'A'	202040301	265	248	229	92.34	86.42
KIRINYAGA	GATHIRUTHIRU	202020402	167	174	134	77.01	80.24
KIRINYAGA	GIAKIRURI	202010601	356	351	186	52.99	52.25
			2021	**2048**	**1688**	**82.42**	**83.52**
MURANG'A	KIREGE	203030404	91	95	85	89.47	93.41
MURANG'A	KIRIMAHIGA	203040501	150	151	140	92.72	93.33
MURANG'A	TIKURA/KIANDUNDU	203040301	75	75	68	90.67	90.67
MURANG'A	NGUNYI	203030102	158	158	143	90.51	90.51
MURANG'A	ITHIKI	203020202	160	149	133	89.26	83.13
MURANG'A	KAMBUGI	203010202	220	232	175	75.43	79.55
MURANG'A	NDUTUMI I	203020303	254	248	105	42.34	41.34
			1108	**1108**	**849**	**76.62**	**76.62**
NYANDARUA	OFFICE-AREA	204060201	149	171	128	74.85	85.91
NYANDARUA	RITAYA NORTH	204060402	124	140	110	78.57	88.71
NYANDARUA	MACHENGO	204040203	139	141	106	75.18	76.26
NYANDARUA	GATHANGARI PLAIN	204040101	234	219	180	82.19	76.92
NYANDARUA	KANYUNGU	204050202	134	214	78	36.45	58.21
NYANDARUA	KIRITI	204010402	228	155	126	81.29	55.26
NYANDARUA	GITHIMA	204020103	204	200	99	49.50	48.53
NYANDARUA	KIRIMA	204030101	239	235	81	34.47	33.89
			1451	**1475**	**908**	**61.56**	**62.58**
NYERI	KIRURUMI/NYARUGUM	205060601	126	123	120	97.56	95.24
NYERI	NGIMAINI/KAIGUCHU }	205040402	128	131	119	90.84	92.97
NYERI	MARIKAINI	205060101	99	103	91	88.35	91.92
NYERI	KAGOCHI	205030401	173	191	151	79.06	87.28
NYERI	LUSOI CENTRE	205020502	308	287	242	84.32	78.57
NYERI	RUKANDU/MARUA	205030604	90	87	82	94.25	91.11
NYERI	KANDOROTHI/KIGONGO	205050203	150	136	113	83.09	75.33
NYERI	GATEMU	205010504	274	279	208	74.55	75.91
NYERI	GAKURUA	205020402	355	336	245	72.92	69.01
NYERI	PURA	205010301	328	202	190	94.06	57.93
			2031	**1875**	**1561**	**83.25**	**76.86**

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
THIKA	GATONO	206040101	114	116	109	93.97	95.61
THIKA	KIMATA "B"	206010403	99	101	92	91.09	92.93
THIKA	MUTUYUINI/ITURAMIRO	206020306	128	128	104	81.25	81.25
THIKA	KAGONDO EAGAD "A"	206050202	104	95	76	80.00	73.08
THIKA	GATUKUYU "B"	206040401	201	185	144	77.84	71.64
THIKA	BTL	206060203	197	195	115	58.97	58.38
THIKA	TOLL	206060202	268	285	114	40.00	42.54
THIKA	GACHORO I	206060102	1725	1628	640	39.31	37.10
THIKA	MAKONGENI 'A'	206030201	139	134	41	30.60	29.50
THIKA	BENVAR	206060103	38	70	2	2.86	5.26
			3013	**2937**	**1437**	**48.93**	**47.69**
MARAGUA	ITHANJI }	207010301	108	118	108	91.53	100.00
MARAGUA	KIHARO	207020105	208	211	201	95.26	96.63
MARAGUA	ITAAGA 'D'	207030101	113	111	106	95.50	93.81
MARAGUA	GATHUKIINI	207010103	98	115	89	77.39	90.82
MARAGUA	GATAIYU	207020502	479	486	433	89.09	90.40
MARAGUA	WANGAI }	207020402	93	88	84	95.45	90.32
MARAGUA	KIBERA	207010302	386	407	289	71.01	74.87
MARAGUA	NGATIA	207040302	267	297	166	55.89	62.17
MARAGUA	NDUINI/KORURU	207030302	199	245	123	50.20	61.81
			1951	**2078**	**1599**	**76.95**	**81.96**
KILIFI	MAUNGE/MIFUDUNI	301050401	80	77	76	98.70	95.00
KILIFI	NGULUWENI	301071001	71	70	64	91.43	90.14
KILIFI	KASIDI `A'	301070101	138	141	119	84.40	86.23
KILIFI	MWEMBE TAYARI	301070704	160	137	93	67.88	58.13
KILIFI	KOKOTONI	301070303	505	507	367	72.39	72.67
KILIFI	MTWAPA VILLAGE	301030102	47	24	17	70.83	36.17
KILIFI	MUTWAPA CENTRE	301030103	122	96	64	66.67	52.46
KILIFI	MWAKUHENGA	301030301	172	235	107	45.53	62.21
			1295	**1287**	**907**	**70.47**	**70.04**
KWALE	KILOLE	302040101	142	174	142	81.61	100.00
KWALE	NGULUKU	302040401	131	142	119	83.80	90.84
KWALE	MKONJE	302040601	159	164	130	79.27	81.76
KWALE	BAHAKANDA	302050102	213	206	170	82.52	79.81
KWALE	BANGA	302020401	115	88	102	#####	88.70
KWALE	BOFU	302020603	207	185	156	84.32	75.36
KWALE	LOWER MACKNNON	302010601	235	184	158	85.87	67.23
KWALE	KIGATO	302030401	339	335	210	62.69	61.95
			1541	**1478**	**1187**	**80.31**	**77.03**
TAITA TAVET	MARUMBENYI	305040302	90	87	81	93.10	90.00
TAITA TAVET	CAMP NO.2	305020502	162	158	130	82.28	80.25
TAITA TAVET	ITANGE	305030401	182	179	150	83.80	82.42
TAITA TAVET	BENGI	305020102	100	104	80	76.92	80.00
TAITA TAVET	MADARASANI	305010202	264	272	229	84.19	86.74
TAITA TAVET	BUGUTA WEST	305050401	172	189	146	77.25	84.88
			970	**989**	**816**	**82.51**	**84.12**
MALINDI	PISHI MINENGA	307030702	67	67	54	80.60	80.60
MALINDI	RABUKO	307030501	140	162	112	69.14	80.00

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
MALINDI	MWANGANI	307020102	113	122	75	61.48	66.37
MALINDI	THOROVENI }	307010103	84	53	51	96.23	60.71
MALINDI	SITA	307030202	309	298	164	55.03	53.07
MALINDI	GONGONI T.C.	307020203	711	702	306	43.59	43.04
MALINDI	TIMBONI	307030302	464	520	206	39.62	44.40
			1888	**1924**	**968**	**50.31**	**51.27**
EMBU	KIANJUKI `B'	401030102	83	82	72	87.80	86.75
EMBU	KIAMBOGO	401040102	130	144	119	82.64	91.54
EMBU	KAMUGERE	401040301	99	88	80	90.91	80.81
EMBU	MUGUI	401050103	215	185	160	86.49	74.42
			527	**499**	**431**	**86.37**	**81.78**
KITUI	NGUNZIU	403010103	73	75	40	53.33	54.79
KITUI	KITULA	403021301	105	106	102	96.23	97.14
KITUI	KANGULI	403050301	61	64	54	84.38	88.52
KITUI	WAANI	403070801	99	109	93	85.32	93.94
KITUI	IIANI	403080404	70	75	59	78.67	84.29
			408	**429**	**348**	**81.12**	**85.29**
MAKUENI	ITHIU	404030203	133	144	107	74.31	80.45
MAKUENI	MWAANI	404060302	301	301	261	86.71	86.71
MAKUENI	MUAANI "B"	404080103	90	102	87	85.29	96.67
MAKUENI	KATHONZWENI MKT.	404120104	332	406	204	50.25	61.45
MAKUENI	MUKONONI	404150107	110	109	92	84.40	83.64
MAKUENI	IVIANI	404160301	141	135	125	92.59	88.65
			1107	**1197**	**876**	**73.18**	**79.13**
MACHAKOS	KAVILINGUNI	405030402	99	103	97	94.17	97.98
MACHAKOS	KWANZASU	405010303	49	51	48	94.12	97.96
MACHAKOS	KIWANZA }	405110801	105	115	102	88.70	97.14
MACHAKOS	KAKUYUNI }	405020302	75	82	72	87.80	96.00
MACHAKOS	KITULA	405080101	116	127	110	86.61	94.83
MACHAKOS	KATISAA	405050604	159	166	152	91.57	95.60
MACHAKOS	CITY	405050303	81	85	73	85.88	90.12
MACHAKOS	KAKUNGU "B" }	405100303	106	130	70	53.85	66.04
MACHAKOS	MWISYANI "C"	405060601	178	102	98	96.08	55.06
MACHAKOS	MANYATTA	405010502	555	551	223	40.47	40.18
MACHAKOS	MLOLONGO	405070202	951	1047	288	27.51	30.28
			2474	**2559**	**1333**	**52.09**	**53.88**
MBEERE	NGORA	407010201	95	88	82	93.18	86.32
MBEERE	ISHIARA `A'	407020101	220	238	140	58.82	63.64
MBEERE	MARIARI `B'	407030102	47	45	38	84.44	80.85
MBEERE	RUROGWE `A'	407030203	101	51	50	98.04	49.50
MBEERE	GITARU	407030303	221	105	73	69.52	33.03
MBEERE	RUGAKORI `D'	407030402	347	301	215	71.43	61.96
MBEERE	KAMWELI "A"	407040102	98	97	76	78.35	77.55
MBEERE	UTITHINI	407040202	86	56	54	96.43	62.79
			1215	**981**	**728**	**74.21**	**59.92**
MERU CENTR	KOOTHU II	408040502	64	110	96	87.27	150.00
MERU CENTR	KITHUNGURI	408080301	145	148	149	#####	102.76
MERU CENTR	KIURONE	408090401	110	110	105	95.45	95.45

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
MERU CENTR	NKUMBO	408100103	177	168	162	96.43	91.53
MERU CENTR	GACURU	408070302	76	70	67	95.71	88.16
MERU CENTR	GIUTI	408090301	583	577	508	88.04	87.14
MERU CENTR	MWANIKA	408020602	216	194	169	87.11	78.24
MERU CENTR	NYAAGI	408090101	218	234	166	70.94	76.15
MERU CENTR	KIUNU	408100503	67	52	47	90.38	70.15
MERU CENTR	KABUBUNGI	408010302	747	668	428	64.07	57.30
MERU CENTR	LEWADOWNS	408010404	339	164	109	66.46	32.15
			2742	**2495**	**2006**	**80.40**	**73.16**
MWINGI	KISOLE	410010501	125	127	114	89.76	91.20
MWINGI	NGOMBENI	410030302	56	58	46	79.31	82.14
MWINGI	NGIENI }	410040401	54	54	51	94.44	94.44
MWINGI	KAMUTUKYA	410050303	49	52	47	90.38	95.92
			284	**291**	**258**	**88.66**	**90.85**
MERU NORTH	LAINI	411120302	91	90	87	96.67	95.60
MERU NORTH	MBARANGA	411010302	132	139	121	87.05	91.67
MERU NORTH	MUKWEONE	411040402	70	66	64	96.97	91.43
MERU NORTH	KABITAI	411070102	125	121	114	94.21	91.20
MERU NORTH	KALIATI	411110303	234	226	208	92.04	88.89
MERU NORTH	KINICHA	411090102	171	196	140	71.43	81.87
MERU NORTH	THANANGA	411100301	436	420	348	82.86	79.82
MERU NORTH	ANTUBANGAI	411080101	412	269	219	81.41	53.16
MERU NORTH	NYANKINE	411140101	216	206	198	96.12	91.67
			1887	**1733**	**1499**	**86.50**	**79.44**
THARAKA	CHANGA/NDANGARE/KI	412010101	119	121	97	80.17	81.51
THARAKA	KIRIANI/KAMARANDI	412010302	47	49	44	89.80	93.62
THARAKA	KATHITHINI/KAUJU/NJIR	412020502	118	107	94	87.85	79.66
THARAKA	MUKURUTI/GASERENI/K	412030302	116	117	101	86.32	87.07
THARAKA	KATITHINI	412030601	445	403	201	49.88	45.17
			845	**797**	**537**	**67.38**	**63.55**
NITHI (MERU	MURAGA	413010103	101	94	85	90.43	84.16
NITHI (MERU	NJURI	413020201	355	149	123	82.55	34.65
NITHI (MERU	KIRINGA	413020303	67	71	65	91.55	97.01
NITHI (MERU	KERERENI/KARUMBI	413030301	96	100	87	87.00	90.63
NITHI (MERU	KONYO	413030501	122	121	116	95.87	95.08
NITHI (MERU	K.K MWENDWA	413040501	157	150	144	96.00	91.72
NITHI (MERU	NTOORO	413050201	92	91	87	95.60	94.57
			990	**776**	**707**	**91.11**	**71.41**
GUCHA(S. KIS	NYABIOTO I	601060201	78	81	78	96.30	100.00
GUCHA(S. KIS	RIAYUBU/OMOSOCHO II	601030302	77	77	75	97.40	97.40
GUCHA(S. KIS	EMONGA	601040203	69	70	67	95.71	97.10
GUCHA(S. KIS	IRONGO	601070203	98	105	91	86.67	92.86
GUCHA(S. KIS	OBUOYO	601060102	183	194	168	86.60	91.80
GUCHA(S. KIS	EGETONGO	601020303	113	143	110	76.92	97.35
GUCHA(S. KIS	NYABIOBO/NYANGO	601040302	88	150	75	50.00	85.23
GUCHA(S. KIS	NYABIGONKORU	601010304	118	92	84	91.30	71.19
GUCHA(S. KIS	IBENCHO	601010302	150	133	110	82.71	73.33
GUCHA(S. KIS	GETARE/NYAMIOBO	601010202	173	167	140	83.83	80.92
GUCHA(S. KIS	BOKEGUSA I & II	601020201	124	85	80	94.12	64.52

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
GUCHA(S. KIS	BOMOOKU	601050201	211	212	135	63.68	63.98
			1482	**1509**	**1213**	**80.38**	**81.85**
HOMA BAY	PITH DIBUORO	602010201	164	170	140	82.35	85.37
HOMA BAY	ONYATA/KOMBOGO/KO	602030102	77	81	72	88.89	93.51
HOMA BAY	GINA,/GINA "B/OYIER	602030303	188	152	134	88.16	71.28
HOMA BAY	KACHIENG/KOTINA 'A'/K	602040202	116	109	97	88.99	83.62
HOMA BAY	KAWERE 'B'	602040801	87	85	80	94.12	91.95
			632	**597**	**523**	**87.60**	**82.75**
KISII CENTRA	NYABISABO	603010202	91	95	82	86.32	90.11
KISII CENTRA	NYABIOTO I & II	603010402	137	139	123	88.49	89.78
KISII CENTRA	MWAMARAGIA	603020302	78	81	74	91.36	94.87
KISII CENTRA	IBACHO MARKET	603030104	233	210	108	51.43	46.35
KISII CENTRA	NYAMAGESA/EKERUBO/	603030405	139	141	128	90.78	92.09
KISII CENTRA	MWAITEMBE	603050202	107	107	97	90.65	90.65
			785	**773**	**612**	**79.17**	**77.96**
KISUMU	ELA/KALUOCH	604010101	50	60	26	43.33	52.00
KISUMU	KAJULU EAST	604020102	137	147	121	82.31	88.32
			187	**207**	**147**	**71.01**	**78.61**
KURIA	KWIGOTO	605010102	126	155	111	71.61	88.10
KURIA	BORDER/BUTIAMA	605020101	84	85	81	95.29	96.43
KURIA	NYAMATABURO 'A'/'B'/'C	605030203	100	100	85	85.00	85.00
KURIA	GUKIHURU	605030601	78	96	63	65.63	80.77
KURIA	GURANI/KWIHO	605050102	107	109	86	78.90	80.37
			495	**545**	**426**	**78.17**	**86.06**
MIGORI	WASWETA	606010401	135	151	130	86.09	96.30
MIGORI	LWANDA OTANG/RAHA	606020402	150	179	145	81.01	96.67
MIGORI	GUGA/KINANDA	606050302	126	133	117	87.97	92.86
MIGORI	NYAMG'AO	606030304	129	140	119	85.00	92.25
MIGORI	TITO	606070401	113	121	104	85.95	92.04
MIGORI	NYAMBECHE	606080502	77	81	69	85.19	89.61
MIGORI	KISUMO/NYAKECH	606060401	102	104	82	78.85	80.39
MIGORI	DEDE TUJU 'A'B'	606040202	179	215	118	54.88	65.92
MIGORI	SORI TOWN	606060201	1258	1147	532	46.38	42.29
			2269	**2271**	**1416**	**62.35**	**62.41**
N.KISII(NYAM	NDURUMO	607010501	76	91	70	76.92	92.11
N.KISII(NYAM	BONYABAGA	607020203	88	90	77	85.56	87.50
N.KISII(NYAM	NYAISA/ESINYO	607030102	191	197	165	83.76	86.39
N.KISII(NYAM	NYARONGE	607040402	143	144	115	79.86	80.42
			498	**522**	**427**	**81.80**	**85.74**
RACHUONYO	GOD-NYANGO	608010202	64	61	53	86.89	82.81
RACHUONYO	KANTAMANGA/KODALO	608010901	477	461	367	79.61	76.94
RACHUONYO	APONDO	608011003	226	134	195	#####	86.28
RACHUONYO	KOYORE/KOGUTU/KODH	608020802	92	111	80	72.07	86.96
RACHUONYO	KANYANDIKO`A'	608030501	109	105	65	61.90	59.63
RACHUONYO	AHANGO	608040901	84	139	80	57.55	95.24
			1052	**1011**	**840**	**83.09**	**79.85**

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
SIAYA	LIGANO	609010302	78	93	70	75.27	89.74
SIAYA	ULAWE "C" }	609030103	73	68	65	95.59	89.04
SIAYA	SIT	609040101	43	40	39	97.50	90.70
SIAYA	SEGA .E	609040303	118	123	68	55.28	57.63
SIAYA	KAMALUNGA `A'2 }	609050201	104	103	87	84.47	83.65
SIAYA	KANYILAJI `A'	609060502	41	45	41	91.11	100.00
SIAYA	KAMJENI	609070402	99	103	92	89.32	92.93
			556	**575**	**462**	**80.35**	**83.09**
SUBA	KISUI EAST WAAGA	610010302	187	193	166	86.01	88.77
SUBA	ONGORO/GOGO	610010502	116	141	110	78.01	94.83
SUBA	MUKENDE/KIORAGA	610030202	76	72	62	86.11	81.58
SUBA	MASINA/MIRORE	610040302	121	122	112	91.80	92.56
SUBA	MIKUYU	610040502	121	138	105	76.09	86.78
			621	**666**	**555**	**83.33**	**89.37**
BONDO	ONGERE	611010102	64	64	64	#####	100.00
BONDO	NYAMWANGA/OCHUKA/	611010301	84	142	74	52.11	88.10
BONDO	KAKREMBA	611020203	103	102	82	80.39	79.61
BONDO	MIYARE "B" }	611030101	83	93	75	80.65	90.36
BONDO	KISWARO "B" }	611030302	93	96	53	55.21	56.99
BONDO	NYANGERA	611040502	168	149	112	75.17	66.67
			595	**646**	**460**	**71.21**	**77.31**
NYANDO	KATITO URBAN	612010402	125	78	50	64.10	40.00
NYANDO	JABER MASHAMBANI /M	612030302	51	46	37	80.43	72.55
NYANDO	AWEDO	612030403	49	49	42	85.71	85.71
			225	**173**	**129**	**74.57**	**57.33**
BARINGO	KAPSOWE	701010401	37	32	31	96.88	83.78
BARINGO	KAMIWINGO	701100202	81	82	76	92.68	93.83
BARINGO	KAPUNYANY/EMBOSITI	701140201	30	30	29	96.67	96.67
BARINGO	ILLIAGAT	#########	48	47	45	95.74	93.75
			196	**191**	**181**	**94.76**	**92.35**
BOMET	CHORWETAB-KIMOIK/KI	702010601	91	90	81	90.00	89.01
BOMET	MWOKYOT	702040402	116	114	110	96.49	94.83
			207	**204**	**191**	**93.63**	**92.27**
KEIYO	KIMWOGO	703011202	115	113	109	96.46	94.78
KEIYO	KAPCHEPNABEI	703020403	64	62	62	#####	96.88
KEIYO	CHEPYAM/CHEPSONGOL	703030101	82	76	67	88.16	81.71
KEIYO	KAMARINY 'B'	703040102	78	66	53	80.30	67.95
			339	**317**	**291**	**91.80**	**85.84**
KAJIADO	KAHUHO `B'}	704010702	225	201	94	46.77	41.78
KAJIADO	NANDONGORU	704041302	107	60	6	10.00	5.61
KAJIADO	OLTEPESI 'B'	704060601	187	204	134	65.69	71.66
			519	**465**	**234**	**50.32**	**45.09**
KERICHO	SARAMEK EST. -SALAM	705020201	70	85	50	58.82	71.43
KERICHO	KAPKIAM 'A'	705040503	56	65	53	81.54	94.64
KERICHO	CHEPKERIO	705070502	56	62	55	88.71	98.21
			182	**212**	**158**	**74.53**	**86.81**

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
KOIBATEK	CENTRE	706010302	168	164	120	73.17	71.43
KOIBATEK	BENONIN	706010601	83	92	64	69.57	77.11
KOIBATEK	SOYMINING 'B'	706020201	40	37	31	83.78	77.50
KOIBATEK	TUGUMOI	706070202	136	132	117	88.64	86.03
KOIBATEK	ESAGERI	706090302	127	124	102	82.26	80.31
			554	**549**	**434**	**79.05**	**78.34**
LAIKIPIA	MURANENE	707010501	163	209	131	62.68	80.37
LAIKIPIA	MAGADI 'A'& 'B'	707040102	252	186	144	77.42	57.14
LAIKIPIA	KINGUKA	707060501	138	129	111	86.05	80.43
			553	**524**	**386**	**73.66**	**69.80**
MARAKWET	KAPCHEROP/CHEPTERE	708010301	351	361	241	66.76	68.66
MARAKWET	KAPSIGORIA/KAPKAITA	708030201	267	246	206	83.74	77.15
MARAKWET	KACHEPKOK/KAMOGO	708060401	112	96	83	86.46	74.11
			730	**703**	**530**	**75.39**	**72.60**
NAKURU	SEKUTIET	709050201	190	185	100	54.05	52.63
NAKURU	MAJANI MINGI	709080501	544	406	389	95.81	71.51
NAKURU	WIYUMIRIRIE `A` (GITO	709090201	43	53	40	75.47	93.02
NAKURU	EGERTON UNIVERSITY	709110101	497	400	297	74.25	59.76
NAKURU	MUNIU	709140302	251	236	190	80.51	75.70
			1525	**1280**	**1016**	**79.38**	**66.62**
NANDI	KAPNYAZA	710041103	78	81	74	91.36	94.87
NANDI	KOBUSOI	710070703	100	120	29	24.17	29.00
NANDI	ABURWO }	710071104	69	69	55	79.71	79.71
			247	**270**	**158**	**58.52**	**63.97**
NAROK	KATAKALA	711010202	117	122	110	90.16	94.02
NAROK	OLOOLUA	711030101	59	31	28	90.32	47.46
NAROK	KISIRIRI 'A'	711050401	195	180	143	79.44	73.33
NAROK	OLOIRIEN	711070101	426	448	162	36.16	38.03
			797	**781**	**443**	**56.72**	**55.58**
TRANS MARA	KILETIEN/SOIMINGIN NO	713010204	180	181	168	92.82	93.33
TRANS MARA	KILUTOORI	713020702	135	141	124	87.94	91.85
TRANS MARA	NDAMAMA	713050301	216	221	203	91.86	93.98
			531	543	495	91.16	93.22
TRANS NZOIA	KAPSARA/KIRIITA	714020502	153	119	95	79.83	62.09
TRANS NZOIA	MUGANDA	714030501	359	389	237	60.93	66.02
TRANS NZOIA	MUTHANGARI	714050101	457	473	342	72.30	74.84
			969	**981**	**674**	**68.71**	**69.56**
UASIN GISHU	CHEPKOIYO	716010202	69	78	69	88.46	100.00
UASIN GISHU	CHELUGUI	716030901	78	79	75	94.94	96.15
UASIN GISHU	LENGUT	716010902	81	81	77	95.06	95.06
UASIN GISHU	SAMBUL 'A'/TULWET 'A'	716040201	106	133	99	74.44	93.40
UASIN GISHU	KAPTENDON/GROWER	716010402	124	118	114	96.61	91.94
UASIN GISHU	NGELEL	716030102	358	365	311	85.21	86.87
UASIN GISHU	SINENDET	716010602	224	208	194	93.27	86.61
UASIN GISHU	CHEPSAAS	716020602	95	87	82	94.25	86.32

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
UASIN GISHU	USWO II	716020303	99	93	80	86.02	80.81
UASIN GISHU	CHEMALAL 'B'	716060401	167	210	134	63.81	80.24
UASIN GISHU	SEIYO 'A'	716041002	97	104	72	69.23	74.23
UASIN GISHU	KAPBILAT	716040501	175	156	124	79.49	70.86
UASIN GISHU	KAPKONGONY/KAPSERT	716050401	130	121	87	71.90	66.92
UASIN GISHU	KAMPI MIWA	716060201	194	175	98	56.00	50.52
			1997	**2008**	**1616**	**80.48**	**80.92**
WEST POKOT	EMBOSIMATWA	717050903	84	83	80	96.39	95.24
WEST POKOT	SES	717040203	69	57	53	92.98	76.81
WEST POKOT	TIGHUM	717020301	58	66	45	68.18	77.59
WEST POKOT	PSUYWO	717050803	217	243	160	65.84	73.73
WEST POKOT	LILUNGA	717030201	60	70	43	61.43	71.67
WEST POKOT	KANGILIKWAN 'A'	717060202	108	124	56	45.16	51.85
WEST POKOT	TINOYO	717050305	61	59	49	83.05	80.33
WEST POKOT	NAKWAPUO	717010603	104	117	53	45.30	50.96
WEST POKOT	LOYAMURUK	717050702	113	78	50	64.10	44.25
WEST POKOT	MOLUOESECH	717010201	100	100	DON'T MATCH		#######
WEST POKOT	APIRAKWANG	717030403	221	103	85	82.52	38.46
			1195	**1100**	**674**	**61.27**	**56.40**
BURET	TENDWET	718010203	113	111	103	92.79	91.15
BURET	KAPKWEN CAMP	718030401	380	533	327	61.35	86.05
BURET	KAPSIRATET	718040102	79	106	70	66.04	88.61
			572	**750**	**500**	**66.67**	**87.41**
BUNGOMA	SIPALA 'B'}	801040104	31	32	29	90.63	93.55
BUNGOMA	LUKHOME	801080602	193	204	180	88.24	93.26
BUNGOMA	MISEMWA	801010202	128	123	117	95.12	91.41
BUNGOMA	NAKALILA "B"	801010602	144	138	130	94.20	90.28
BUNGOMA	NAKHWANA	801080401	86	83	77	92.77	89.53
BUNGOMA	NAMAKANDA	801050102	142	149	127	85.23	89.44
BUNGOMA	KHALALA	801090303	182	192	132	68.75	72.53
BUNGOMA	NANGILI 'A'(1)	801020201	150	149	62	41.61	41.33
BUNGOMA	MULMKHU	801070501	113	114	109	95.61	96.46
			1169	**1184**	**963**	**81.33**	**82.38**
BUSIA	EMALAYA	802020103	187	197	155	78.68	82.89
BUSIA	NAMISI	802030403	86	87	75	86.21	87.21
BUSIA	MADOLA	802040404	80	80	73	91.25	91.25
BUSIA	MUMBAKA 'B'/MIMBAKA	802050301	85	83	78	93.98	91.76
BUSIA	BUHALULU'B'/MADUA/SI	802060503	236	265	171	64.53	72.46
			674	**712**	**552**	**77.53**	**81.90**
MT. ELGON	NALONDO 'A'/NALONDO	803010102	29	29	29	#####	100.00
MT. ELGON	MALAKISI	803010402	86	85	81	95.29	94.19
MT. ELGON	CHEMRON	803030301	73	75	46	61.33	63.01
MT. ELGON	KIBUMET 'B'	803040102	65	68	61	89.71	93.85
MT. ELGON	TOYO-BEI/CHEPT.LETW	803040302	170	172	157	91.28	92.35
			423	**429**	**374**	**87.18**	**88.42**
KAKAMEGA	SENGETEI	804010302	164	171	157	91.81	95.73
KAKAMEGA	MASIAKALO }	804040304	82	83	78	93.98	95.12
KAKAMEGA	IKOLI }	804020503	79	76	71	93.42	89.87

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
KAKAMEGA	ESHIAMBOKO	804060301	151	151	135	89.40	89.40
KAKAMEGA	KAMATUNI	804010103	169	165	144	87.27	85.21
KAKAMEGA	SHIKANGANIA	804050102	183	179	155	86.59	84.70
KAKAMEGA	BUTWENE	804050502	250	223	128	57.40	51.20
KAKAMEGA	VIKOSHE }	804030202	116	120	52	43.33	44.83
			1194	**1168**	**920**	**78.77**	**77.05**
LUGARI	BONDENI	805010202	101	101	92	91.09	91.09
LUGARI	LUYEKHE }	805010303	101	104	62	59.62	61.39
LUGARI	MAKUTANO	805020102	180	182	166	91.21	92.22
LUGARI	MAKUTANO WEST	805020201	129	129	108	83.72	83.72
LUGARI	MUKHUYU	805020401	97	101	96	95.05	98.97
LUGARI	MAKHUKHUNI }	805030203	143	145	138	95.17	96.50
			751	**762**	**662**	**86.88**	**88.15**
TESO	ATAPA-OBURI `A' /OTEK	806010401	120	122	108	88.52	90.00
TESO	KAMUSENU/AKOBWAIT	806030603	211	205	198	96.59	93.84
TESO	ALABOROI	806040101	70	98	91	92.86	130.00
TESO	CHAKOL	806040402	147	139	131	94.24	89.12
			548	**564**	**528**	**93.62**	**96.35**
VIHIGA	EBUKHUNGU	807010304	118	109	97	88.99	82.20
VIHIGA	EMATIBINI }	807020302	128	125	107	85.60	83.59
VIHIGA	LUSOSI	807030602	98	101	92	91.09	93.88
VIHIGA	SHIKHAMBI	807040105	144	147	138	93.88	95.83
VIHIGA	KABINJARI }	807050202	89	104	86	82.69	96.63
VIHIGA	ISHAGALO	807060501	119	107	98	91.59	82.35
			696	**693**	**618**	**89.18**	**88.79**
BUTERE/MUM	EBUKANANACHI	808010203	203	218	175	80.28	86.21
BUTERE/MUM	EMALOMBA	808020104	99	105	89	84.76	89.90
BUTERE/MUM	BUYOKA	808030101	164	167	160	95.81	97.56
BUTERE/MUM	MUSHIKULU	808030404	107	106	101	95.28	94.39
BUTERE/MUM	KHWISERO	808040303	134	96	70	72.92	52.24
			707	**692**	**595**	**85.98**	**84.16**
TOTAL			**121669**	**119378**	**87609**	**73.39**	**72.01**
ASAL							
LAMU	KICHINJIONI	303010101	72	68	42	61.76	58.33
LAMU	MKUNGUNI	303010101	145	137	70	51.09	48.28
LAMU	MABATINI	303010201	89	99	72	72.73	80.90
LAMU	MPEKETONI TOWN	303030101	266	239	116	48.54	43.61
TANA RIVER	DOWN TOWN II	306060401	101	75	42	56.00	41.58
ISIOLO	OLA JAROLE	402020101	292	187	128	68.45	43.84
ISIOLO	KULA MAWE I	402020101	122	117	63	53.85	51.64
ISIOLO	MISSION/MWANGAZA I	402020201	72	76	43	56.58	59.72
ISIOLO	OLLA GANNA/MANYATT	402050302	108	109	87	79.82	80.56
MARSABIT	DARES-SALAAM	406010102	35	29	27	93.10	77.14
MARSABIT	SHAURI TAKO II	406010102	61	80	42	52.50	68.85
MARSABIT	TECHNICAL	406010202	177	167	115	68.86	64.97
MOYALE	DIDA	409010202	97	98	80	81.63	82.47
MOYALE	TULICHA	409010301	130	112	107	95.54	82.31
MOYALE	ALISO SAGE	409010402	72	77	67	87.01	93.06

DISTRICT	EA NAME	EA CODE	HHLDS	HHLDS	HHDS	%	%
MOYALE	SHAURI YAKO	409010503	102	100	76	76.00	74.51
GARISSA	GARISSA NDOGO 'A'	501010101	71	95	40	42.11	56.34
GARISSA	TOWN CENTRE	501010103	208	183	100	54.64	48.08
GARISSA	MASALANI	501010201002			DON'T MATCH		
GARISSA	AN-AAM	501010401	64	87	36	41.38	56.25
MANDERA	BULLA DANA	502060101	151	170	111	65.29	73.51
MANDERA	ADM III	502180101	89	85	18	21.18	20.22
MANDERA	HOSPITAL	502180203	29	29	29	#####	100.00
WAJIR	TOWN	503010101	21	21	21	#####	100.00
WAJIR	ABUDULLA BULLE ZON	503010402	66	87	32	36.78	48.48
WAJIR	BULLA JOGOO	503010402	792	747	324	43.37	40.91
SAMBURU	HURUMA	712020101	115	132	70	53.03	60.87
SAMBURU	BAZAAR ESTATE	712030205	270	269	133	49.44	49.26
SAMBURU	YOMO ESTATE	712030207	139	117	84	71.79	60.43
TURKANA	LODWAR YOUTH POLYT	715120103	105	112	73	65.18	69.52
TOTAL			**4061**	**3904**	**2248**	**57.58**	**55.36**
GRAND TOTAL			**206043**	**196039**	**126055**	**64.30**	**61.18**

ANNEX 2

Republic of Kenya
Census '99

Post Enumeration Survey, 2000

(STRICTLY CONFIDENTIAL)

Structure No.: S . - - - - - -
Serial No . 029727

ovince District Division Location Sub-Location E.A. Number Household No . Enumerations status

Household Totals: Males **Females** **Total**

Information Regarding All Persons									B:Education	C: Information Regarding females aged 12 years and above				For office use only	
							Whether Enumerated			Particulars of all Live Births					
Name (P-00)	Serial No .	Relationship (P-10)	Sex (P-11)	Age (P-12)	Religion (P-14)	Marital Status (P-15)	(P-16)	(P-17)	Education (P-23)	(P-40) (P-41)	(P-46)	(P-48) (P-49)		(P-60) Moving Status	(P-70) Matching Sta
hat are the names of l persons who spent st night in this ousehold? (Record at ast two names of l persons, young or d, starting with the ad of the ousehold).		What is ..... 's relationship to the head of the household? 1 Head 2 Spouse 3 Son 4 Daughter 5 Brother/ Sister 6 Father/Mother 7 Other relative 8 Non- relative 9 NS / DK	What is .....'s sex? Mark X in the appropriate box	How old is .....? (Record age in completed years using two digits. If under one year write "00")	What is.....'s religion? 1 Catholic 2 Protestant 3 Other Christian 4 Muslim 5Traditionalist 6 Other Religion 7 No Religion 8 NS / DK	What is.....'s marital status? 1 Never Married 2Monogamos 3 Polygamous 4 Widowed 5 Divorced 6 Separated 7 NS / DK.	Was...a usual member of this household on the census night, 24/25th August 1999? 1 Yes 2 No 3 Born after Census night 4 Visitor	Where was... enumerated? 1 In this household 2 Elsewhere 3 Not enumerated but residing in Kenya 4 Not enumerated not residing in Kenya 5 Born after census night 6 DK	What is the highest level reached by..? Write the appropriate code using the categories shown on the inside cover of the questionnaire	How many children Have you born alive?	How many children have you born alive who have died?	When was your last child born?		1 Non-Mover 2 In-Mover 3 Born after	1 Match 2 Possible Mi 3 Non-Match 4 In-Mover 5 Born after 6 Not Stated
			Male Female							Boys Girls	Boys Girls	Month Year			
	1														
	2														
	3														
	4														
	5														
	6														
	7														
	8														
	9														
	0														

Serial No . of continuation form

Outmovers.

D: Information Regarding Persons Enumerated in The Main Census But Not in the PES							For Office use only	
Name P-00A	Serial Number	Relationship P-10A	Sex P-11A		Age P-12A	Residence P-13A	Matching Status P-70A	Status of Census Enumeration P-80A
What are the names of all persons who were in this household during the main census but were not here last night? (Record at least two names of all persons young or old.		What is ..... 's relationship to the head of the household? 1 Head 2 Spouse 3 Son 4 Daughter 5 Brother/ Sister 6 Father/Mother 7 Other relative 8 Non- relative 9 NS / DK	What is .....'s sex? Mark X in the appropriate box		How old is.....? (Record age in completed years using two digits. If under one year write "00")	Was..... .a usual member of this household on the census night, 24 / 25th August 1999? 1 Yes 2 No 3 Visitor	1 Match 2 Possible Match 3 Non-Match 4 Not Stated	1 Correctly Enumerated 2 Erroneously Enumerated (Fabrication)
			Male	Female				
	1							
	2							
	3							
	4							
	5							
	6							
	7							
	8							
	9							
	0							