working paper series

EGYPT LABOR MARKET PANEL SURVEY 2006:
REPORT ON
METHODOLOGY AND DATA COLLECTION

Ghada Barsoum

Working Paper No. 0704

Egypt Labor Market Panel Survey 2006: Report on Methodology and Data Collection

Ghada Barsoum

Working Paper 0704

February 2007

Ghada Barsoum, ELMPS Project Coordinator
Email: gbarsoum@popcouncil.org

Executive Summary

The Egypt Labor Market Panel Survey 2006 (ELMPS06) is the first full-fledged panel study of its scale in Egypt. This panel follows a nationally representative sample of 4,816 households visited in 1998, households that split from that sample, plus a refresher sample of 2,500 households. Total number of households reached in 2006 is 8,349.

The ELMPS06 provides estimates of employment, unemployment and underemployment. The survey also collects information on job characteristics, mobility, and earnings. Collected data also covers issues of a household's socio-economic characteristics, demographic characteristics, family enterprises and women's status and work. Different questionnaire sections are listed in this report for easy reference. A separate community level questionnaire has been administered to collect data on access to services and work opportunities in sampled localities.

Reference Period:
The week proceeding the interview (short reference period) and the three months proceeding the interview (long reference period).

Data Collection Dates:
Data collection started in Cairo, Giza, Kaliobia, Sharkia, Gharbia and Menoufia on December 24, 2005. Data collection started in Alexandria, Behira, Kafr El Sheikh on December 31, 2005. Data collection in Ismailia, Suez, Port Said and Dumiat and Dakahlia started on January 14, 2006. Data collection for all Upper Egypt governorates (except Giza) started January 21, 2006.

Subjects Covered
- Household Assets and durable goods
- Education
- Employment
- Migration
- Women's unpaid work, women's status and fertility
- Household enterprises
- Remittances and non-work related income

Target Population

The ELMPS06 covers the civilian non-institutionalized population 6 years of age and above. Excluded from the survey's coverage are residents of the five frontier governorates of North Sinai, Matrouh, Red Sea, New Valley, and South Sinai. These represent an exclusion of less than 2% of the Egyptian population.

Sampling
The ELMPS06 uses a probability sample that is based on a stratified multi-stage design following the master sampling frames from the Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS). Households interviewed in 1998 and followed in this round were chosen from the CAPMAS 1995 master sample, which is described in details in Appendix I. The refresher sample of 2006 was selected from CAPMAS 2004/2005 master sample. The latter is a nationally representative cross-sectional sample also described in Appendix I.

The ELMPS 06 final sample consists of a total of 8,349 households distributed as follows: (i) 3,684 households from the original ELMS 98 survey, (ii) 2,167 new households that emerged from these households as a result of splits, and (iii) a refresher sample of 2,498 households. Of the 23,997 individuals interviewed in 1998, 22,987 were still alive or in the country in 2006 and 17,357 of those (75.5%) were successfully re-interviewed in 2006, forming a panel that can be used for longitudinal analysis. The 2006 sample contains an additional 19,743 new individuals. Of these 2,663 individuals joined the original 1998 households, 4,880 joined the split households, and 12,200 were part of the refresher sample of households.

Data Sources

The ELMPS06 was administered by field interviewers in all governorates. Data were collected using printed questionnaires carrying the names of CAPMAS and the Economic Research Forum (ERF). Proxy interviews were allowed for those less than 15 years old. Otherwise, researchers were instructed to get the data from the same respondent unless s/he could not be reached after three visits.

Instrument Design

The ELMPS06 uses two survey instruments: the household instrument and the community-level instrument. The household instrument is applied to all households sampled (both in 1998 and 2006) and split households. This instrument builds on the 1998 design with the addition of many new questions. A major addition in the panel questionnaire design is the addition of a new section that inquires about previous members of the household and their current location. In addition to this information, the new section also allows for the identification of the personal codes for all individuals in the 1998 data sets. Other new modules that were added collect information about siblings, remittances, fertility, and cost of marriage. The education section was significantly expanded and the work detection questions became more elaborate in terms of listing specific activities that can be considered as economic activities.

Another innovation of the ELMPS06 was the introduction of the community-level questionnaire. This questionnaire gathers data on the availability of services in the locality (defined as the village in rural areas and as the district in urban areas). It also gathers data on the availability of work opportunities, and prevalent economic activities and occupations.

Introduction

This report seeks to provide a full documentation of the Egypt Labor Market Panel Survey 2006 (ELMPS06). The report covers issues of data sources and methodology as well as a narrative description of the major activities undertaken as part of this study. These activities include the conclusion of negotiations between the Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS) and the Economic Research Forum (ERF); the preparatory activities for data collection including sampling and training; data collection; office review; and data entry.

The Egypt Labor Market Panel Survey (ELMPS 06) is a nationally representative household panel survey. It is the first full-fledged panel survey of its scale in Egypt. The first round of this panel study took place in 1998 covering a nationally representative sample of 4,816 households. A pilot round took place in 2004. This is the first full-scale round covering: households visited in 1998; households that split from the original sample as sons and daughters, among others, move out of the original household to form their own households; and a refresher sample of 2500 households.

Similar to the 1998 and 2004 rounds, survey activities are executed by the CAPMAS in close collaboration with a consultant team hired by the Economic Research Forum (ERF).

This survey utilizes two research instruments: a household-level questionnaire and a community-level questionnaire. The household questionnaire comprises three complementary questionnaires: the first inquires about the demographic characteristics of all household members, household assets and housing conditions (therefore dubbed as the household questionnaire); the second includes modules on the different education and work characteristics of individuals aged six years and more (the individual questionnaire); and the third inquires about on migrants from the household, household enterprises and non work-related sources of income. The community questionnaire includes modules on access to education, health services, utilities, and work opportunities in a locality. This questionnaire is applied on the district level for all sampled urban localities and on the village level for all rural localities.

Both questionnaires were administered by CAPMAS field staff. Data is reviewed in the field as each team had two or three reviewers depending on its size. Questionnaires are brought to the CAPMAS Cairo office for an office review stage, a data coding stage, and finally a data entry and verification stage.

The following sections of the report describe in detail these major project milestones.

Preparatory Activities for Data Collection
Contract with CAPMAS was duly signed in November 2005. Preparatory activities for data collection included: the printing of questionnaires; sample preparation; security clearance letters for field staff; purchase of gifts for respondents; purchase of cell phone and recharge cards for field supervisors; and the selection and training of field staff.

Questionnaire Design:
Questionnaire design was finalized by the ERF team prior to the conclusion of the agreement between CAPMAS and ERF.

The household-level research instruments comprise of three interrelated questionnaires for each household. The household questionnaire collects data on the different demographic characteristics of household members, household assets and access to services. This questionnaire also includes a module that tracks individuals who were part of the 1998

sample. The questionnaire allows space for 20 individuals as members of the household and for 10 splits.

The individual questionnaire includes modules on the education and work characteristics for individuals six years and up. The printed version allows space for only five individuals, but more than one individual questionnaire can be used for a household depending on its size.

The third questionnaire in the household-level research instrument is the "Migration, Family Enterprises and Non-wage Income" questionnaire, which includes the modules on migration, remittances, non-work-related income sources, and non-agricultural household enterprises.

Questionnaire Sections
Following are the different sections of the three questionnaires:

1. Household Questionnaire
Section 0.1 Basic Characteristics

Section 0.2 Basic Characteristics of members who lived in household in1998 but no longer live in household in 2006 and their new addresses for tracking purposes

Section 0.3 Housing, Service & Facilities

Section 0.4 Consumer Durable Goods

2. The Individual Questionnaire
Chapter 1 Parents' Characteristics

Section 1.1 Father's Characteristics

Section 1.2 Mother's Characteristics

Section 1.3 Data on Siblings

Chapter 2 Education

Chapter 3 Migration

Chapter 4 Employment and Unemployment

Section 4.1 Employment detection question during the past 7 days

Section 4.2 Unemployment

Section 4.3 Subsistence and Domestic Work

Section 4.4 Employment detection during the three months and characteristics of individual's first job

Chapter 5 Job Characteristics

Section 5.1 Primary Job in the past three months

Section 5.2 Secondary Job in the past three months

Chapter 6 Female Paid Employment

Chapter 7 Fertility, Women's Status, and Cost of Marriage

Section 7.1 Fertility

Section 7.3 Women's Status

Section 7.3 Cost of Marriage

Chapter 8 Job Mobility

Chapter 9 Earnings

Section 9.1 Earnings in Primary Job

Section 9.2 Earnings in Secondary Job

3. The Questionnaire on Migration, Remittances, Non-work Related Income and Household Enterprises

Chapter 10 Migration, Remittances and Non-work related income

Section 10.1 Migration (individuals currently working abroad)

Section 10.2 Remittances

Section 10.3 Non-work related income

Chapter 11 Household Enterprises

Section 11.1 Non-agricultural enterprises

Section 11.2 Non- agricultural Enterprises: Employment outside the household

Section 11.3 Agricultural Activities: Capital Equipment

Section 11.4 Agricultural Activities: Ownership of Livestock

Section 11.5 Agricultural Access to Credit

Additions and changes to the 1998 Questionnaire

New sections were added to the 1998 questionnaires and a number of questions were deleted because they did not produce useful results after the analysis. Following are the major changes to the 1998 questionnaire:

1. The panel design mandated a number of changes, including the addition of a new section, Section 0.2, which gathered information on the basic characteristics of members who lived in household in1998 but no longer live in household in 2006 and their new addresses to track them. The cover page also included a question regarding the type of the household (whether it is originally visited in 1998, is a split household, or from the new sample). Section 0.1, the household roster, also included an additional question (0105), which inquires about the individual's person number (**pn)** in the 1998 data set. Data collectors were able to get this information from the data sheets that were printed for each household containing basic demographic characteristics and a summary of her/his work and education characteristics.

2. Questions about land ownership and cultivation were added in section 0.3. Although they do not quite fit under housing and services, this was is best place to include them. Instructions during training were to write zero if no land was owned or rented by household.

3. The section on durable goods, section 0.4, now includes questions on whether the item was brought at the time of marriage and whether an item is bought to be used by a household member after she/he marries.

4. A short section on siblings (section 1.3) was added, which refers to total number of siblings, and whether or not they reside in the same household.

5. The section on education is expanded significantly. It now includes questions about the characteristics of secondary, preparatory, and primary schools, where relevant. Questions about repetitions and interruptions of schooling are included in order to gain better understanding of the number of years of schooling as opposed to grade

level achieved and age of exit/completion. The section also allows assigning a unique code for each school attended by the individual. These unique codes were received from the Ministry of Education and allow for analysis on school characteristics based on further data from the Ministry.

6. The migration section was moved earlier in the questionnaire so that it applies to all individuals whether they worked or not. In 1998, this section only applied to those who had previously worked. The section now applies to all those aged 15 and above. It also includes a new question about place of birth.

7. In the sections on work characteristics, we no longer have a reference week and a reference three months. We use instead the past seven days (counting back from day of first interview with individual) and past three months.

8. In the unemployment section (Section 4.2), we have added questions about the use of a land or cell phones in job search activities. We also separated the question on registering with a government agency from the job search question. Now all the activities listed under job search are limited by the past three months reference period.

9. We separated the questions on subsistence and domestic work in a new section. These questions now apply to all children aged 6-17 and all women aged 18-64, irrespective of employment status. The questions on domestic work are now much more detailed than before and ask about time spent on various domestic chores during past 7 days. If the same amount of time is spent everyday, then interviewers were instructed to multiply the daily times by seven. However, this is designed to allow for variations in schedules every day. One of the reasons this section now applies to the past week rather than a reference week was that that it might be difficult to get an accurate estimate due to recall problems. Only the last question of the section allows for the activity to be done concurrently with other activities (child care). Otherwise, interviewers were instructed that they are after the time spend exclusively on the activity in question

10. Questions about the "first job" were added into the section detecting employment in the forgoing three months. As in the job mobility section, to qualify as a job, the individual must have spent at least 6 months at the job. Thus, a job during summer vacation is considered a job in the current employment section, but does not qualify as a first job or as a job in the mobility section. Similarly, an individual could have worked in the reference 3 months (on summer jobs) but has no first job, because that job lasted less than six months. The same criterion of six months applies to the job mobility section to prevent listing of back and forth mobility between school and work for students who work only during the summer.

11. Chapter 7 contains two entirely new sections, one on fertility and one on the cost of marriage. Sections 7.1 and 7.2 apply to all ever-married women.

12. In the earnings chapter, the section on second main job was dropped. The instructions specify that if the main job has changed during the past three months, earnings should be collected in this section for all main jobs combined, not just the last one. This should not be a problem since very few people actually change their main job in a given three month period. The same applies for the earnings of the secondary job. If an individual has changed their secondary job during the past three months, earnings should be collected for all the secondary jobs combined in the three months period.

13. The household enterprises questionnaire now applies to all individuals, irrespective of whether they have enterprises or not. While these questions are better placed at the household questionnaire, we believe that having these questions early at the interview might scare people off. The household enterprise questionnaire comes towards the end of the interview and was therefore an ideal place for these questions. Household

enterprise questions are divided into four sections: one section on non-agricultural enterprises and three short sections on agricultural enterprises.

Sample Preparation:

The ELMPS06 sample consists of three types of households: households visited in 1998, households that split from the 1998 sample as sons and daughters (among others) form their own households, plus a refresher sample of 2,500 households.

Appendix I includes a detailed description of the ELMPS06 sample, including the description of the 1998 sample, the 2004 household identification process and the attrition of some households due to loss of identification data prior to data collection. Also included is a description of the master sample from which the refresher sample of 2,500 households was extracted in 2006.

Printing of Data Sheet for all 1998 Sample and Sample Lists:

Data sheets summarizing household and individual characteristics of the 1998 sample were printed for field staff to take with them to the field.   We printed a sheet for each household that contained:

- The household serial number in the 1998 data set

- The governorate name

- The district name

- The village/shiakha name

- A listing of all individuals and their characteristics including:

  o Gender
  o Relationship to head of household in 1998
  o Data of birth
  o Age in 1998
  o Marital status
  o Year of (first) marriage
  o Age at (first) marriage
  o Occupation
  o Employment status or reason for not working
  o Education status
  o School year if individual was in school in 1998
- Characteristics of the dwelling unit including ownership status, material used in ceilings, walls and floors, and number of rooms

Field staff used these data sheets for two objectives:
1. To verify that they are interviewing the right household from the 1998 sample, based on its characteristics
2. To copy the code of the household into the new questionnaire of the household for data analysis

We also printed sample lists for all the 1998 PSUs containing the code of the household in the 1998 data set and names of heads of households taken from the questionnaires of the 1998 data collection process (names were not input into the data set in 1998).   These lists, combined to CAPMAS 1998 sample lists which contained household addresses, were used by CAPMAS field staff to identify households.

We did not have a sample frame for split households, since these were to be identified during the data collection process.

The 2006 refresher sample had separate lists produced by CAPMAS and was, in the majority of cases, in different localities to avoid confusion with the 1998 panel lists.

Security Clearance Letters

CAPMAS issues letters to police offices in all governorates to explain the task of the team and to ask for their support in facilitating their task. In many instances, these letters are crucial for the facilitation of the data collection. Each researcher receives a copy of this letter in case it is required by the household members. These letters give legitimacy to the research team. The need for these letters was obvious in urban areas or newly urbanized areas where a higher rate of rejection is usually the case. In rural areas, the letter is often submitted to the mayor to facilitate the task of the team. Particularly in rural and recently urbanized areas, team supervisors also showed the letters to the *imams* of neighborhood mosques in order to announce in mosque microphones about the study. This procedure, as field staff noted, was pivotal in increasing response rates in such neighborhoods.

Printing of Questionnaires

CAPMAS proceeded with the printing of questionnaires even before the transfer of the first payment in order to save time. We ordered the following number of questionnaires to be printed: Cell-Phone Lines and Charge Cards
Field supervisors were provided with cell-phone lines and charge cards. Those who already owned a cell-phone line were provided with charge cards only, which were replenished at least once for most supervisors.

The objective of cell phones is twofold. First, field supervisors are instructed to inform colleagues in other governorates when there was a split household or a 1998 household that has moved to an exact address in that particular. Second, field supervisors can reach the ERF team whenever they have questions in the data collection process and the identification of households. As will be noted in the data collection section, a hotline was established to receive calls when addresses of 1998 sample households were not clear or when the identification data was confusing.

Purchase of Gifts

Our experience in 1998 and 2004 taught us that the provision of some compensation to household members for their time during the interview, primarily since the questionnaire is relatively long, can be very helpful. Specifically for panel studies, it is advisable in the literature to establish rapport with the household in order to be accepted in successive panel rounds. In Assaad, Barsoum and Dang (2005) we discuss the different options for gifts tried in the 2004 pilot panel round. We had to avoid edible giveaways in order to eschew problems related to food quality. In 2004, we experimented with alarm clocks but these were not very successful. They were particularly unsuccessful in poor rural areas as respondents did not see the need for alarm clocks. In upper scale urban areas, the alarm clocks were considered as too cheap a gift. We searched in the market for items that would be practical and usable and yet within our budget limitations. At the end, we decided to get two types of gifts: the first was a packet containing three soap bars and a plastic bag of detergent; the second was a porcelain or glass coffee cup. The first was used in rural and low-income urban areas and the second was used in upper-scale areas. Field staff noted that the soap was very welcomed in the field for its practical use.

Selection of Field Staff

Data collection personnel were selected from staff nominated by CAPMAS management, or from new graduates (also identified by CAPMAS staff) who had prior data collection experience. The task of the ERF team was to train these nominees and to choose from them. We were provided with an extra number of candidates to allow for the attrition of disqualified candidates.

All interviewers were females because females have easier access to all household members and can particularly reach female members who would not welcome a male interviewer at their homes. Most reviewers were males but there were a handful of female reviewers. All supervisors, except for the Beni Suef team and one of the two Alexandria teams, were males.

Training of Field Staff

To accommodate teams from all over the country, we held four training programs in four different governorates: Cairo, Alexandria, Ismailia, and Minia. We provide in this section details of each training program.

For the four training programs each day began with a review of the materials covered in the previous day, a review of the questionnaire sections, and role playing. The day program would end with an exam. In 1998, we used true/false type of questions for these exams, but in 2006, the trainees were asked to fill a questionnaire based on a case that was given to them. Questionnaires were reviewed each day by the research team and common mistakes were discussed on the following morning. Selection of field staff was based on their exam results throughout the week.

Phase I: Training in Cairo (December 17 – December 22, 2005)

This was the first training program. CAPMAS Cairo staff as well as those from Menoufia, Sharkia, and Gharbia received training during this first phase. Cairo staff covered the three governorates of Cairo, Kaliobia and Giza. Researchers and reviewers for Sharkia were also selected from CAPMAS Cairo staff. Only the Sharkia supervisor was from the Sharkia CAPMAS office. Menoufia was covered by a team from the Menoufia governorate joined by a reviewer from Cairo. Gharbia was covered by a team from Gharbia, except for the two reviewers who were selected from the Cairo CAPMAS office.

During this phase we trained 105 CAPMAS staff members. Only 75 CAPMAS staff members were selected. These were assigned for Cairo (4 teams), Kaliobia (1 team), Giza (1 team), Menoufia (1 team), Sharkia (1 team), Gharbia (1 team), and finally a team for quality control.

Phase II: Training in Alexandria (December 24 – December 27, 2005)

Phase II included two teams for Alexandria, one for Kafr El Sheikh and one for Behira. We followed the same training program. During this phase we trained 39 CAPMAS staff members and young graduates who were nominated by CAPAMS. Only 38 CAPMAS staff members and graduates were selected for Alexandria (2 teams), Kafr El Sheikh and Behira.

Phase III: Training in Ismailia (December 31 – January 3, 2006)

Team members for Port Said, Suez, Dakahlia and Dumiat came to Ismailia to receive training. Overall, we trained 36 CAPMAS staff members and young graduates who were nominated by CAPAMS in Ismailia. Only 26 CAPMAS staff members and graduates were selected to cover these governorates.

Phase IV: Training in Minia (January 14, 2006 – January 18, 2006)

Staff for Upper Egypt governorates (except Giza) received training in Minia during this final phase. Team members who came from all governorates were asked to stay in Minia. We had CAPMAS staff and graduates from Fayoum, Beni Suef, Kena and Aswan. Staff for Minia, Assuit and Sohag were nominated by CAPMAS management from Minia. Overall, we trained 74 CAPMAS staff members and young graduates who were nominated by CAPAMS in Minia. Only 69 CAPMAS staff members and graduates were selected to cover these governorates.

The Data Collection Process

Dates

Start dates for administering household questionnaires for each governorate followed the training schedule. Those trained in Cairo started data collection on December 24, 2005. Those trained in Alexandria started data collection on December 31, 2005. However, those trained in Ismailia started for two days and then had to wait for one week because of the Kurban Bairam. Data collection for Ismailia, Port Said, Suez, Dumiat and Dakahlia started in full gear on January 14, 2006. Data collection for all Upper Egypt governorates (except Giza) started January 21, 2006.

The quality control team started working on December 24, 05. During the first week, they focused on other teams who received training in Cairo. They then gradually expanded their work to include the rest of the governorates as all teams started to work.

Field Teams and Number of Researchers

We had the following number of teams and staff members for each governorate.

Table 1 Data Collection Teams

	# of Teams	# of researchers	# of reviewers	# of supervisors	Staff hired from
Cairo*	4	18	8	4	Cairo
Alexandria	2	12	4	2	Alexandria
Port Said	1	1	1		Port Said
Suez	1	3	1		Suez
Dumiat	1	4	1	1	Dumiat
Dakahlia	1	6	2	1	Dakahlia
Sharkia	1	5	2	1	Researchers and reviewers are hired from Cairo; Supervisor is hired from Sharkia
Kaliobia	1	6	2	1	Cairo
Kafr El Sheik**	1	6	2	1	Kafr El Sheik
Gharbia***	1	5	2	1	Researchers and supervisor were from Gharbia; the two reviewers were hired from the CAPMAS Cairo staff
Menoufia	1	3	2	1	Researchers and supervisor were from Menoufia; one of the two reviewers was from Cairo
Behira	1	6	2	1	Behira
Ismailia*****	1	3	1	1	Ismailia - supervisor was responsible for the Port Said and Suez teams.
Giza	1	6	2	1	Cairo

Table 2 Data Collection Teams - Continued

Beni Suef	1	7	2	1	Beni Suef
Fayoum	1	5	2	1	Fayoum
Minia	1	7	3	1	Minia
Assiut	1	7	3	1	Minia
Sohag	1	6	2	1	Minia
Kena	1	6	3	1	Kena
Aswan	1	6	2	1	Aswan
Luxor	1	4	2		Aswan
Quality Control Team	1		4		Cairo
Total	27	132	55	23	
Total field staff				**210**	

*Two researchers were fired during data collection from the Cairo teams, one for negligence and the other for absenteeism.
** Two researchers were fired during data collection from the Kafr El Sheik team for absenteeism.
**** One researcher was fired from the Gharbia team during data collection for negligence.
**** Two researchers were fired during data collection from the Ismailia team for negligence. The Suez team completed the work under the same supervisor.

Role of Field Staff

The division of labor within the data collection team is highly scripted following regular CAPMAS data collection procedures. Female researchers visit households escorted by supervisors and reviewers. Since most of the supervisors and reviewers were males, they did not usually stay with the researcher. Instead, they wait for the whole team to return at a meeting point (a café or a certain landmark). The researcher fills in the questionnaire during the interview. Researchers were instructed to visit the household at least three times in order to get data from the individual him/herself. For the panel sample, researchers also compare the data collected to the 1998 data. They also make sure they get the exact address for individuals who move out of the household and form new households. As bonus, researchers make the visits to the split households identified and this way they add to their production and final pay.

During training, researchers were instructed to stop the interviewing process if they find that they are not in the correct household as identified in the panel sample list. They were also instructed to report the problem to the supervisor who would contact the Cairo hotline using his cell phone to make sure s/he gets the right identification information.

Field reviewers, by definition, review the questionnaires during the data collection process and compare the 1998 data to that collected during the 2006 interview. Reviewers are also instructed to visit households with researchers when in doubt of the collected data.

Supervisors are responsible for the stock of blank questionnaire and for the collection of questionnaires after they are reviewed by field reviewers. They are also responsible for the distribution of the sample assigned to the team among researchers. During the ELMPS06 study, the supervisors had two added responsibilities:

1. To follow up on the completion of the household questionnaire section on split households, including informing teams in other governorates with addresses of splits to get their data. Question 214 in the household questionnaire, which inquires on whether a split household was reached, was the responsibility of the supervisor.
2. To arrange for the transportation of gifts to researchers and to keep an inventory of gifts. Supervisors contacted the ERF team in cases of shortage of gifts.

3. To arrange for the transportation of completed questionnaires to the CAPMAS office in Cairo.

Household Identification Hotline

A CAPMAS staff member based in Cairo and a research assistant were in charge of a hotline to respond to problems in the household identification procedures. Problems related to households with unavailable sheets or addresses were forwarded to this team. Particularly when there were clear discrepancies between data collected in the field and the data 1998 data sheets, the field supervisor would call the hotline to inquire on possible reasons for such discrepancy. The hotline team re-checked lists and addresses. The team assisted in the coordination of the follow up of split households across governorates

Cell phones were used extensively to call the project coordinator to check on certain questions in the questionnaires, and to ask about cases when they found discrepancies between 1998 data sheets and the actual situation of the household members.

Community Survey Questionnaires

An innovation over the ELMS 1998 was the inclusion of a community survey questionnaire. Community information is crucial in the analysis of household data for a number of reasons. First, it provides a natural link to policy variables such as social and physical infrastructure, including schools, family planning clinics, water and sewage networks, and markets. Second, it provides useful instruments for the estimation of simultaneous equation models. The interdependence of household decisions often makes it necessary to estimate such models and community variables describing the environment in which the household is operating.

The community questionnaires were administered to key informants in all villages in sampled rural areas and in all districts in sampled urban areas. The questionnaire gathered information on local availability of public services, local prices for key services and products, and the condition of local infrastructure including access to water, electricity, sewage, and education and health facilities.

Data collection for community survey questionnaires was undertaken by separate teams. The following table shows training and start dates for the different governorates.

Table 3 Start dates for data collection for community questionnaires

Governorates	Training Date	Data Collection Start Date
Cairo, Kaliobia, Giza, Sharkia, Gharbia, Dakahlia, Menoufia	December 17, 2005	December 24, 2005
Fayoum, Beni Suef, Minia, Assiut, Sohag, Kena, Luxor, Aswan	February 10, 2006	February 11, 2006
Port Said, Suez, Ismailia and Dumiat	February 13, 2006	February 14, 2006
Alexandria, Behira, Kafr El Sheik	May 2, 2006	May 3, 2006

Data collection for Cairo, Giza, Kaliobia, Menoufia, Sharkia and Gharbia started with the Cairo teams for the household questionnaires (December 24, 2005). Training for the team for Upper Egypt, constituting of one researcher for each governorate, was held for one day on Friday, February 10, 2006 in Minia. One researcher conducted the data collection for Port Said, Suez, Ismailia and Dumiat due to the limited sample size. He received training on February 13, 2006. Training for Alexandria, Behira and Kafr El Sheikh took place on May 2nd. The delay was due to staffing constraints as we had to use reviewers from the household

questionnaire data collection teams to undertake this task. For these three governorates, the data collection for the community questionnaire had to wait till the end of the data collection for the household questionnaire.

Table 4 Distribution of community questionnaires per governorate

Cairo	21
Alexandria	11
Port-said	4
Suez	2
Domiyat	6
Dakahliya	15
Sharkiya	17
Kaliobiya	12
Kafr-Elsheikh	16
Gharbiya	16
Menofiya	9
Behira	14
Ismailia	7
Giza	16
Bwni-Suef	11
Fayoum	13
Menia	19
Asyout	19
Sohag	17
Qena	13
Aswan	9
Luxor	2
Total	**269**

Teams for administering the community questionnaire were instructed to communicate with teams responsible for administering household questionnaires in order to collect community data on districts/villages of split households.

Important Data Collection Issues in the Panel Design

Definition of a Split Household

Split households are those households formed by or joined by an individual who was a member of the 1998 data collection round. In rural areas, we were faced by the situation that many of the sons who marry, continue to live in the same dwelling unit. During the training, we instructed researchers to inquire about eating arrangements to separate different households living within the same dwelling unit. However the data collection reality provided a number of complexities which we had to address. In some instances, the newly formed household would have a dwelling unit that is an apartment within the same rural house of the patriarchal family. Probed into the issue, particularly by quality control team members, we are faced with a situation where the informant notes "our stuff is in the apartment upstairs, but our meals and expenses are one". In this case, we considered them as living in one household with the patriarchal family.

What was noticeable in the field was that the definition of the situation varied significantly between different members in a household. In rural areas, fathers of married sons assume the role of a patriarch. Therefore, they usually say that they are providing for their sons' families, even if this is not the case. On the other hand, it was also noticeable that a son's wife would be the first to admit the financial independence of her nuclear family. In these situations, we instructed researchers that a separate household should have a separate kitchen to connote separate cooking arrangements.

The Household Identification Process

We gave training instructions to researchers to stop an interview if they notice clear discrepancy between the 1998 data sheets and the situation of a household they interview. However, quality control teams as well as office reviewers spotted a number of household cases that were clearly different from the ones interviewed in 1998. In some instances, we had to omit a number of questionnaires that did not match with the 1998 data set and consider them as households that we were not able to reach. In other instances, questionnaires were resent to the field with a quality control team member accompanied by the researcher/reviewer to investigate the case. This was the case in governorates where many discrepancies were found related to negligence in the data collection process.

De Jure and De Facto Household Members

This study followed the *de jure* rule for the inclusion of different household members as part of the household. The criterion was a presence in the household for at least six months, whether the member was actually present during the interview or not.

However, we realized that in order to capture newly formed households and splits that took place in the past six months, we had to relax this rule on the condition that the individual moves with the intention of staying permanently away from the household, primarily either for marriage for permanent work. Only in these cases we followed the *de facto* definition of household membership.

Issues Related to Data Collection Timing

One of the unfortunate events that coincided with the data collection process has been the spread of the avian influenza (bird flu) towards the end of this stage. The impact of this pandemic was primarily palpable in rural areas. Male reviewers and supervisors were primarily unwelcome for fear that they were from the Ministry of Health seeking to inspect houses checking for poultry. Female researchers were less threatening, and this alleviated the potential damage of this pandemic on the data collection process.

In Behira, a concurrent study by another organization, which we could not identify, weighed children and provided vaccination that was associated with cases of illness among children. Because of the problems caused by this study, our research team was received with great hostility in one particular village (El-Delengat). Villagers were extremely suspicious to the extent that they asked one researcher to use the soap provided as gift in front of them, for fear that it might be a source of danger. The team leader had to turn to the mayor of the village, who facilitated the task of the team within a couple of days.

Office Review

All questionnaires administered in different governorates were moved to CAPMAS head office in Cairo for office review and data processing.

The major tasks involved in the stage of office review are:

- To check for missing variables

- To check the logical sequence of collected data
- For panel and split households, to ensure that the data collection team visited the same household by comparing the 1998 data sheets to the newly filled questionnaire. This also entails checking that individual IDs are correctly copied from the 1998 data sheet to their proper places in the new questionnaires.
- To ensure that individual IDs are written on each page for data entry

Reviewers were given clear instructions to minimally change questionnaires and to refer to the research team should they find major changes requested into a certain questionnaire. In this case, we used to call the data collection team and ask them to re-visit the household.

Office review training took place during the data collection process on February 8, 2006. Six teams (each with 4 or 5 reviewers) started working a week later as questionnaires were gradually moved to Cairo. The office review process lasted till June 30, 2006.

Coding

Coding entails turning text describing occupations, economic activities, education certificates and schools into numeric values following CAPMAS coding booklets.

This stage began May 1st and was finalized June 30, 2006.

The following table illustrates the different coding books used in the dataset and the conversion process introduced to ensure the comparability of the 2006 data to that of 1998.

	Economic Activity		Occupation		Education Certificates	
	1998	**2006**	**1998**	**2006**	**1998**	**2006**
Classification Index Used	The Arab Unified Coding Book for Economic Activities (issued January 1986)	CAPMAS Economic Activities Codebook (January 1996)	The Arab Unified Coding Book for Occupations (January 1985)	CAPMAS Occupations Codebook (January 1996)	Education Certificates Book (January 1996)	Education Code book (January 2006)
Number of Digits	4-digit	4-digit	5-digit	6-digit	4-digit	6-digit
Conversion Process for data comparability	converted to 2006 on the level of 1, 2 and 3-digits	-	converted to 2006 on the level of 1-digits	-	-	converted to 88/98 on the level of 2-digits

We also introduced a coding system for all schools which were mentioned by respondents. The coding book was brought from the General Authority of Educational Buildings (GAEB). Schools are coded using a 10 digit unique code. Since the GAEB coding book did not include Azharite, nursing and mailing schools, we coded these schools as follows:

Azharite Schools: 7777777777

Nursing Schools: 8888888888

Mailing Schools: 8888888888

Other schools not listed in the code book: 999999999

Data Entry and Validation:

Data entry began on May 17, 2006 and was concluded on August 17, 2006. This stage took an extended period of time for two reasons:

- Delay of the provision of the computers to CAPMAS for reasons related to USAID customs and tax exemption.
- Many of CAPMAS staff members who usually take part in data entry were taken by the preparatory activities of the census.

The data validation process followed. CAPMAS staff had validation rules for data ranges, consistency, and skips. These were designed in close collaboration with the ERF team.

Appendix I: Description of Sample Selection

The ELMPS06 sample consists of three types of households:

1. Households visited in 1998
2. Split households
3. A refresher sample of 2,500 households

In this section we describe in details the sampling techniques for both the 1998 and 2006 samples. We also describe the attrition to the 1998 sample due to loss of some household identification data, which were kept by CAPMAS.

The Selection Process of the 2006 New Sample

This sample was selected from the CAPMAS 2005 Master Sample. This is a nationally representative two-stage self-weighted (to the extent possible) sample. Each governorate is allocated a number of PSUs in the master sample that is proportionate to its size and its urban/rural distribution.

The master sample was prepared through a two-stage process. First, shiakha's and villages are selected by probability-proportionate-to-size method from two different sampling frames (one urban and another rural). In the second stage, these selected primary sample units are divided into secondary sampling units of 700 households each. A total of 1200 sampling units is then randomly selected to constitute the final master sample of CAPMAS.

The ELMPS06 2006 new sample was proportionately selected from the CAPMAS master sample. Its urban/rural distribution was based on the following two formulas:

For urban locations:
$$\frac{100\, \text{PSUs} \times \text{Total Number of Urban PSUs}}{\text{Total Number of PSUs in the Master Sample}} = 46\, \text{PSUs in Urban Areas}$$

For rural locations:
$$\frac{1000\, \text{PSUs} \times \text{Total Number of Rural PSUs}}{\text{Total Number of PSUs in the Master Sample}} = 54\, \text{PSUs in Rural Areas}$$

Primary sampling units were then randomly selected from the CAPMAS master sample. Then within each PSU (containing 700 households in the master sample) we randomly selected 25 households.

The following table illustrates the distribution of PSUs in both the CAPMAS master sample and the ELMPS 06 refresher sample across governorates.

Table 5 Distribution of 2006 new sample across governorates

Governorate	CAPMAS master sample			ELMPS 06 sample
	Urban	**Rural**	**Total**	
Cairo	153	0	153	13
Alexandria	74	0	74	6
Port Said	11	0	11	1
Suez	9	0	9	1
Dumiat	6	14	20	2
Dakahlia	26	62	88	7
Sharkia	21	65	86	7
Kaliobia	28	39	67	5
Kafr El Sheikh	11	30	41	4
Gharbia	23	44	67	6
Menoufia	11	43	54	5
Behira	19	52	71	5
Ismailia	8	7	15	2
Giza	64	39	103	8
Beni Suef	9	26	35	3
Fayoum	9	29	38	3
Minia	14	53	67	5
Assiut	16	37	53	4
Sohag	14	47	61	5
Kena	10	35	45	4
Aswan	8	10	18	2
Luxor	4	4	8	2
Matrouh	2	2	4	--
North Sinai	3	2	5	--
South Sinai	1	1	2	--
Red Sea	2	1	3	--
New Valley	1	1	2	--
Total	557	643	1200	--
Total PSUs minus frontier governorates	**548**	**636**	**1184**	**100**

The 1998 Sample[1]

The 5,000 households that constitute the initial survey sample in 1998, were selected from a CAPMAS master sample prepared in 1995. The master sample consists of 750,000 households in 500 primary sampling units (PSUs) each consisting of 1,500 households.

The CAPMAS master sample was selected through a two-stage process. The country is first divided into two strata: urban and rural. Each stratum is in turn divided into sub-strata representing each governorate. All the villages (in the case of rural strata) or *shiyakhas* (urban quarter, in the case of urban strata) in each substratum were listed and assigned a weight based on their population. The first stage consisted of choosing the villages and *shiyakhas* that would be represented in the sample based on the principal of probability proportional to size. This meant that a *shiyakha* or a village is possibly selected more than once if its size warrants that. The selected *shiyakhas* and villages are then divided into PSUs of approximately 1500 housing units each; then one or more PSUs are selected from each *shiyakha* or village. The selected PSUs were then re-listed in 1995 to enumerate all the households selected. As shown in Table 6, the master sample contains 306 urban PSUs and 194 rural PSUs.

---

[1] This section builds on Assaad, Barsoum and Dang (2005)

For the 1998 survey sample, 200 PSUs were selected from the master sample, on the basis of the numbers shown in Table 6. The desired number of PSUs in each substratum was selected from the number available in the master sample using a systematic interval. Cairo and Alexandria were deliberately over-sampled and rural areas under-sampled at the time to increase the probability of obtaining women wage-workers in the private sector, which tend to be concentrated in metropolitan areas. [2]

Twenty-five housing units were then randomly selected from the 1,500 housing units listed in each of the selected PSUs. Table 6 also includes the expansion weight for each sampling stratum based on population estimates in June 1997.

Table 6 : 1998 Sample Strata and Distribution

| Region / Governorate | PSUs | | |
	# of PSUs in Master Sample	# of PSUs in 1998 Survey Sample	Ratio of Survey over Master
1) Greater Cairo:			
Cairo	40	33	0.83
Giza City	7	5	0.71
Shubra El Kheima	5	4	0.8
Total Greater Cairo Region	*52*	*42*	*0.81*
2) Alexandria Region:			
Alexandria	35	20	0.57
Suez	20	3	0.15
Port Said	20	2	0.1
Ismailia	10	3	0.3
Total Alexandria Region	*85*	*28*	*0.33*
3) Urban Lower Egypt			
Dumiat	9	2	0.22
Dakahlia	12	6	0.5
Sharkia	12	6	0.5
Kaliobia (excluding Shubra el Kheima)	7	3	0.43
Kafr el Sheikh	9	4	0.44
Gharbia	12	5	0.42
Menoufia	12	4	0.33
Beheira	12	5	0.42
Total Lower Urban	*85*	*35*	*0.41*
4) Urban Upper Egypt			
Giza (excluding Giza City)	14	2	0.14
Beni Sueif	10	5	0.5
Fayoum	10	5	0.5
Minia	10	5	0.5
Assiut	10	5	0.5
Sohag	10	5	0.5
Qena and Luxor	10	4	0.4
Aswan	10	4	0.4
Total Upper Urban	*84*	*35*	*0.42*
5) Rural Lower Egypt			
Dumiat	11	5	0.45
Dakahlia	13	3	0.23
Sharkia	13	4	0.31
Kaliubia	13	4	0.31
Kafr El Sheikh	11	4	0.36
Gharbia	13	3	0.23
Menoufia	13	3	0.23

---

[2] Suez Canal cities were over-sampled in the CAPMAS Master sample, therefore they appear to have been under-sampled in the study sample when in fact they were not.

Region / Governorate	PSUs		
	# of PSUs in Master Sample	# of PSUs in 1998 Survey Sample	Ratio of Survey over Master
Beheira	13	4	0.31
Ismailia	10	5	0.5
Total Lower Rural	*110*	*35*	*0.32*
6) Rural Upper Egypt			
Giza	14	6	0.43
Beni Sueif	10	3	0.3
Fayoum	10	2	0.2
Minia	10	3	0.3
Assiut	10	3	0.3
Sohag	10	2	0.2
Qena	10	3	0.3
Aswan	10	3	0.3
Total Upper Rural	*84*	*25*	*0.3*
TOTAL	**500**	**200**	**0.4**

The Attrition of some of the1998 households due to loss of identification data and the 2004 Pilot Sample

In 2004, we approached CAPMAS to conduct a pilot panel survey in preparation for the 2006 data collection process. We then learned that some of the questionnaires, which were the only source of household identification data, were lost. The cover pages of other questionnaires were scanned, which meant that the identification data was accessible for these households. Some questionnaires were also kept by CAPMAS. To rectify the loss of questionnaires, we did what we describe in Assaad, Barsoum and Dang (2005) as the matching process. We revisited households with lost questionnaires based on the 1998 sample lists. We used an instrument that collected information on the basic demographic, work and education characteristics of the household. The instrument also included questions about the dwelling unit characteristics. We then matched the data collected to the 1998 dataset for households in each locality to discern which data record belongs to which household. The data set only included household and individual characteristics and the names of localities (Shiakhas and villages). The success rate of this matching process was 63%. Households we could not match were not included in the panel sample, although their data was kept in the 1998 data set. These were considered as unreachable for the panel purpose.

An analysis of the attrition from the sample showed that it was essentially due to the random loss of identifying records rather than any systematic attrition process. No significant association was found between the probability of attrition and household and individual characteristics in 1998. Weights based on the probability of non-response were used to correct for attrition in the panel data

The following table illustrates the status of questionnaires per governorate and the final number of households whose identification data was retrievable before the 2006 data collection process.

During the pilot panel study, we collected data from 282 households in Cairo, Giza and Menoufia. These households were not revisited in 2006.

Governorate	Matched	Scanned	Available	Not Matched (lost households)	Households visited in the 2004 Pilot Survey*	Total 1998 sample at the beginning of the 2006 data collection
Cairo	0	735	49	3	99	685
Alexandria	275	117	0	55	-	392
Port Said	35	0	0	9	-	35
Suez	0	0	68	1	-	68
Dumiat	20	118	24	5	-	162
Dakahlia	181	0	0	38	-	181
Sharkia	80	0	108	54	-	188
Kaliobia	51	0	173	44	-	224
Kafr El Sheik	0	147	47	1	-	194
Gharbia	0	208	0	-	-	208
Menoufia	0	174	0	-	95	79
Behira	0	171	0	-	-	171
Ismailia	0	195	0	2	-	195
Giza	53	0	195	66	88	160
Beni Suef	183	0	0	12	-	183
Fayoum	26	0	126	13	-	152
Minia	189	0	0	8	-	189
Asuit	118	78	0	7	-	196
Sohag	41	120	0	11	-	161
Kena	0	148	0	-	-	148
Aswan	64	23	61	18	-	148
Luxor	21	0	0	2	-	21
Total	**1337**	**2234**	**851**	**349**	**282**	**4140**

* Out of these 282 households, only 247 households were actually re-interviewed in 2004 for reasons related to death, rejection of household to participate in the study, migration and inability to locate household.

Table 7 Number of printed questionnaires per category

Type of Questionnaire	# of Copies
Household questionnaires	9,000
Individual Questionnaires	13,500
Migration and non-work related Income questionnaires	9,000
Community Questionnaires	360