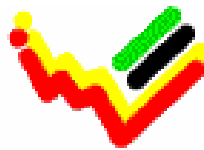# Report from a short-term mission on

EVALUATION OF THE INITIAL PHASES OF THE BUSINESS CENSUS

7-26 October 2002

TA for the 'Bridging Support Program to Strengthen the Institutional Capacity of the National Statistics, Mozambique

Kenny Petersson

_____
Instituto Nacional de Estatística

Kenny Petersson
Statistics Sweden
Kenny.Petersson@scb.se
+46 –19 17 65 62

Table of contents

List of abbreviations

CO	Scanstat Coordination Office in Statistics Denmark
CEMPRE	Censo de empresas 2002 (Business census 2002)
Danida	Danish International Development Assistance
DESE	Directorate of Statistics on Enterprises and Sector Economics
DISI	Department of Informatics and Information Systems
DPINE	Provincial delegation of INE
FUE	Ficheiro de unidades estatísticas (software for the business register bought from INE-P)
INE	Instituto Nacional de Estatística, Mozambique
INE-P	Instituto Nacional de Estatística, Portugal
Scanstat	Consortium between Statistics Denmark, Statistics Norway and Statistics Sweden
SCB	Statistics Sweden

Ms Irene Tuveng, Consultant and Mr Azarias Nhanzimo, Director of Sector Statistics and Statistics on Enterprises

1   EXECUTIVE SUMMARY

The National Statistical Institute (INE) is conducting a business census (CEMPRE) during the period October-December 2002 with a possible extension of the field-work period until January 2003. INE has received technical assistance for the planning of the survey from INE Portugal. INE Portugal is also assisting in the data processing by providing a routine for the data entry. The same routine will also be used for the maintenance of the business register in the future.

By making an evaluation during the first part of the field-work period and the start of the data entry, the Scandinavian Bridging Support Program gives assistance to INE in strengthening the routines and to assure that the continuance of the CEMPRE is done in the best way possible. As stated in the terms of reference (Annex 4) there are three objectives of the mission

- Evaluate the data collection process
- Discuss the data quality
- Evaluate the data processing routines

A forth objective was to give proposals for the future structure of the system for surveys of enterprises and establishments based on samples from the business register.

The main activities were to review the written documentation and discuss the plans as well as the day-to-day work with the staff of INE working with the CEMPRE. The consultant also assisted in preparing documents for the planning of the data processing. The field-work was monitored by field visits in Maputo Cidade and Maputo Província.

a) the data collection process

The evaluation could verify that the field-work operations were logistically well prepared and organised. A small pilot that was made immediately before the start of the main data collection gave a test of the organisation, training for the interviewers and indications on the main problems.

The main recommendation for this part is to prepare a small "pocket manual" as documentation of the changes and clarifications of the contents and objectives of the data collection that is expected to be the result of the first period of the data collection. Examples on how the operational definitions should be applied would also be desired to include as an annex to such a document.

Another observation is that there could have been need to prepare a "plan B" if the first experience shows that it will not be feasible to conduct the census according to the first plan within the time and budget restrictions that are given. The main strategy for such a plan would be to simplify the census for the smallest enterprises, and collect the most basic data, such as main type of economic activity already during the listing phase, and make the data entry directly from the results of the pre-listing without returning for a complete interview.

The main concern for the data collection is the difficulties to apply a uniform operational definition on which "enterprises" to include and which to exclude as too small.

b) The data quality

The contents of the CEMPRE should have been better prepared conceptually. If the concept is not completely clear it affects the quality of the result of the collection. The main conclusion is

that there should have been more time for the conceptual development and discussions on how to use each piece of information and also what a question costs in terms of extra visits to the enterprises and also if there are "immaterial costs", such as effects on the answers to other questions or the willingness to provide the information in general.

A general recommendation was to clarify the intended use of each variable and evaluate the data quality in relation to the planned use of it. The most cumbersome variable in the data collection is "Valor de vendas". The variable is mainly meant to be used for classification of the enterprises by size. The main recommendation is to classify by class of size already during the interview instead of asking in detail about the numbers, given the objective is to classify by size. The number of staff members is probably the best indicator on the size of the business in terms of production (value added). The only question about the size of the staff in the CEMPRE refers to November 2001. The recommendation was to classify the establishments by the number of staff members also in November 2002, at least into one of four size groups.

c) routines for the data processing

In a way it was too early to evaluate the data processing routines during the mission, because the planning of the data processing was very much related to the installation of the new software, FUE, for the data entry, which was bought from INE Portugal and installed 14-18 October. The evaluation of the plans for this part included a review of the facilities and recommendations on changes to be discussed between INE-M and INE-P. A strong recommendation was to include powerful tools to verify if an enterprise already exists in the register, to avoid that it is introduced as a new one if it already exists in the register.

During the mission INE decided to split the data entry of the three main questionnaires into two parts, one to be entered via the FUE and another for which INE was to design an own data entry routine. This action clearly indicates that the links between the questionnaires and the data entry routines had been very weak during the planning process.

c) proposals about the structure of the information system in the area of economic statistics in the future

The discussion on this part of the report focuses on how to facilitate the use of the new business register as frame for sample surveys. One recommendation was to prepare a table with all establishments. The first option would be to include all establishments in the table "Estab", which is now used only for establishments of enterprises with more than one establishment.

Another recommendation was to apply a more composite classification of the size of the business and the total investment it represents. That is because there are so many small businesses in the commercial and service sectors. These businesses are of minor interest for the business register and it will demand a lot of resources to keep the information for these establishments up to date. The interest in the CEMPRE is mainly to count these establishments by type of main economic activity. By using a composite size classification to identify small enterprises in the sectors of commerce and services, mostly operated by the family alone, without permanent structure or from the house where they live, it would be easier to "split" the register and focus on the larger businesses in the updating routines. Possible criteria to separate the "lowest" category could be if they have permanent staff for which they pay salaries.

Meeting in the he central office for the field work in Maputo City

2 RESUMO EM PORTUGUÊS

O Instituto Nacional de Estatística (INE) está a executar um censo do Empresas (CEMPRE) durante o período October-December 2002 com uma extensão possível do período do trabalho de campo até fevereiro 2003. O INE recebeu assistência técnica para a planificação de INE Portugal. INE Portugal também dá assistência ao processamento de dados fornecendo uma rotina para entrada de dados. A mesma rotina vai também ser usada para a manutenção do cadastro de empresas no futuro.

Por meio de executar uma avaliação durante a primeira parte do período do trabalho de campo e o começo do processamento, o projecto Escandinavo dá assistência ao INE para reforçar as rotinas e para assegurar que a continuação do CEMPRE vai dar resultados que servem bem o desenvolvimento das outras partes da estatísticas económicas de Moçambique.

Como indicado nos termos de referência (anexo 4) há três objectivos principais da missão

- avaliar o processo do levantamento de dados.
- discutir a qualidade dos dados
- avaliar as rotinas de processamento de dados

A missão foi feita durante três semanas em outubro de 2002 por Kenny Petersson, estatístico na área de "Labour Market and Education Statistics" do Statistics Sweden (SCB). A missão foi feita junto com o pessoal do pelouro económico, conforme os termos de referencia (Anexo 1).

As actividades principais foram discussões e entrevistas com o pessoal do INE. O consultor também preparou documentos para as discussões e sobre as rotinas para o processamento.

Os resultados/observações principais foram:

1) Recolha de dados.

As rotinas para recolha de dados são bem definidas e os manuais para o trabalho do campo conte as partes importantes. A recomendação principal é elaborar uma "guia do bolso" onde enfoca-se as partes mais importantes do manual e especialmente mudanças das rotinas para recolha que introduza-se para reforçar as rotinas quando mostra-se necessário.

A maior preocupação da recolha de dados é como, na prática estabelecer uma definição da limite entre os "estabelecimentos" que vão ser incluídos e quais que vão ser excluídos.

2) A qualidade dos dados

Observou-se que elaborou-se mudanças do conteúdo depois da última revisão dos questionários. Recomenda-se continuar avaliar a qualidade dos dados da recolha, clarificar o objectivo de cada variável e tomar acções para aumentar a qualidade quando é necessário para tornar possível usar os dados conforme o objectivo da variável. Se mostra-se evidente que não é possível atingir a qualidade necessário é melhor alterar o objectivo ou excluir a pergunta, para não executar trabalho redundante de recolha de dados que não pode-se usar.

A estrutura dos questionários é diferente comparado da estrutura do formulário para entrada de dados na versão do FUE que foi instalada 14-18 de outubro. A estrutura lógica do FUE (e também do CEMRE) é usar um questionário comum para a empresa e o estabelecimento que é a

sede da empresa quando existe somente um estabelecimento, e questionários anexadas para cada estabelecimento quando existe mais do que um estabelecimento por empresa.

O treinamento dos inquiridores não foi avaliado em detalhes mas discussões sobre exemplos mostrou que os questionários controlados foram bem preenchidos. A organização usando controladores do campo serve também para reforçar a qualidade do resultado das entrevistas. Entrevistas de controladores nas províncias Maputo Cidade e Maputo Província dão evidência que os controladores foram bem preparados.

3) Rotinas para processamento dos dados

Em princípio foi demais sedo avaliar as rotinas para o processamento durante a missão, porque o plano do processamento foi ligado da mudança para o software FUE. O processamento do CEMPRE vai dividir-se em duas partes separadas, uma para entrar-se directamente na aplicação para o FUE e uma outra para as outras variáveis dos questionários. A observação mais importante na área de processamento é que seria desejável salientar as ligações entre o desenho dos questionários e as rotinas para o processamento. Uma parte da missão foi utilizada para identificar as diferenças entre a primeira versão do FUE e os questionários principais e dar propostas sobre soluções para processar a parte do CEMPRE que não entra no FUE.

Por meio de prolongar a planificação do processamento seria possível elaborar testes da digitação antes de definir as versões finais dos questionários. Uma outra observação é que não elaborou-se testes para verificar se seria possível utilizar mais os dados no cadastro existente para tentar diminuir o peso de trabalho na etapa de digitação dos dados recolhidos. A razão mais importante é que o novo software para o cadastro, FUE, não foi instalada mais sedo durante o processo de planificação, e também que o INE prefere uma solução simples para a estrutura do processamento.

Um dos problemas no processamento é juntar/ligar estabelecimentos da mesma empresa. Para facilitar esta parte recomendou-se incluir um formulário separado na aplicação FUE para procurar empresas e estabelecimentos por meio de digitar qualquer parte dos nomes deles. Também seria possível usar a estrutura da empresa no velho cadastro. Uma empresa que tem estabelecimentos em mais do que uma província no velho cadastro, provavelmente também tem no novo.

4) dar propostas sobre a estrutura dos sistemas de informática na área de estatísticas económicas no futuro.

Esta parte dos termos de referência cobre primeiro propostas sobre a estratégia para utilizar o cadastro como fonte dos inquéritos mensais, trimestrais e anuais no futuro. A maior diferença entre o cadastro do FUE e o velho cadastro (em termos da estrutura) é que encontra-se os dados sobre estabelecimentos em duas tabelas separadas no FUE ("Empresa" e "Estab"), quando o cadastro existente tem somente uma tabela ("Testablishment"). Este pode ser resolvido por meio de incluir registos sobre todos os estabelecimentos na tabela "Estab" depois do fim do CEMPRE (ou elaborar uma cópia para facilitar amostragem, saída de estatística etc. por nível de estabelecimento). Para a manutenção do cadastro é importante facilitar o controle da cobertura e estabelecer rotinas permanentes (ou pelo menos anuais) para incluir novos estabelecimentos grandes no cadastro.

Seria bom incluir uma classificação separada para as pequenas "empresas" nos sectores comercio e serviços para tornar possível facilmente executar acções de actualização do cadastro sem incluir estas pequenas "empresas".

Agradeço todo o pessoal de INE, que eu tive a vantagem encontrar durante a estadia no INE e trabalhar junto com.

Interviewers and field supervisors in Maputo Cidade

3 INTRODUCTION

The main objectives for the mission were to make an evaluation of the initial phases of the CEMPRE for the Scandinavian Bridging Support programme, and also to assist INE in assuring that the field work operations in the CEMPRE are well organised and that the data collection routines are well coordinated with the quality goals and that the data handling routines and the routines for data processing are efficiently coordinated and well organised. The terms of reference for the mission are found in Annex 4.

The mission was conducted during the period 6-26 October 2002 by Mr Kenny Petersson, senior statistician in the Department for Labour Market and Education Statistics in Statistics Sweden (SCB).

The main counterparts were Mr Azarias Nhanzimo, Director of the Directorate of Statistics on Sectorial Economics and Enterprises, Mr Adriano Matsimbe, National Coordinator of the CEMPRE, Mr Cirilo Tembe, Deputy National Coordinator of the CEMPRE and Mr Calado Fijamo, Department of Informatics and System Design (DISI). The mission was planned in the Division for Economic Statistics in collaboration with Ms Irene Tuveng, long-term consultant in the Scandinavian Bridging Support Program.

The consultant would like to express his thanks to all officials and individuals met for the kind support and valuable information which he received during its stay in Mozambique, and which highly facilitated the work of the consultant.

This report contains the views of the consultant(s), which do not necessarily correspond to the views of Danida or INE.

Small workshop in Maputo

4   ACTIVITIES DURING THE MISSION

4.1   Review of the plans

The consultant first made a review of the written material, such as the plan of the census, the report from the mission from INE Portugal for design of the survey, the manuals for the interviewers and the manuals for the field supervisors. The first week was also used for review of the questionnaires and discussions of the field-work .

The review shows that the general plan for the logistics of the CEMPRE is the result of comprehensive work on , identification of tasks in the data collection process and preparations for the plan of operation for the field-work. Regarding the conceptual development and the design of design of the questionnaires there would have been need to clarify the objectives and strengthen the links to the data processing in a much earlier stage of the planning process. In principle, the whole plan for the data processing should be completed before the questionnaires are printed. The main observation was that this process should have started much earlier.

4.2   The pilot survey

A small pilot was conducted very soon before the start of the main data collection, mainly to test how the questionnaire was received by the enterprises. There were some changes in the questionnaires as a result of the pilot. The main result was that the question about the registration number NUIT was cut out and also the classification by "Grupo de contribuição Industrial" (GCI). The GCI was left in the questionnaire, but excluded from the interviews.

Due to the delay in the delivery of the software FUE for the data entry it was not possible to test the routines for the data entry on the pilot.

The solution should have been to start the whole process much earlier. The first version of the software should have been available in time for the pilot survey, and the time between the pilot and the main survey operation should have been sufficient to process the pilot, make a proper evaluation of the pilot and redesign the questionnaires accordingly. The pilot should also serve as a tool to verify if the estimated time schedule is realistic and if there is need for specific measure to conduct the census within the existing budget restrictions.

4.3   The old business register

The processing of the pilot survey could also have been used to verify if the data in the old business register could easily be used to avoid repeating the data entry of all the text information that is already in the old business register. During the mission, the consultant wrote a short memo on how to verify if it would be worth while importing the old business register into the new and organize the data entry as an update when the questionnaire refers to an enterprise that could be securely identified in the old register.

The "old" business register contains in total 28 449 establishments, whereof 24 320 are classified as active.

Table 1. Number of establishments in the existing "old" business register

Estado	Número de estabelecimentos
Activa	24320
Extinta	437
Nao confirmada	2607
Nova	306
Paralizada	778
Repetido	1

The register contains establishments both in the private sector and in the public administration. The active establishments were distributed by main type of economic activity according to the following:

Table 2. Number of active establishments by main type of economic activity.

Secção	Designação	Número
	Sem informação	9
A	AGRICULTURA, PRODUÇÃO ANIMAL, CAÇA E SILVICULTURA	697
B	PESCA	121
C	INDUSTRIAS EXTRACTIVAS	54
D	INDÚSTRIAS TRANSFORMADORAS	2 092
E	PRODUÇÃO E DISTRIBUIÇÃO DE ELECTRICIDADE, GAS E ÁGUA	67
F	CONSTRUÇÃO	381
G	COMÉRCIO POR GROSSO E A RETALHO etc.	7 152
H	ALOJAMENTO E RESTAURAÇÃO (RESTAURANTES E SIMILARES)	1 819
I	TRANSPORTES, ARMAZENAGEM E COMUNICAÇÕES	603
J	ACTIVIDADES FINANCEIRAS	174
K	ACTIVIDADES IMOBILIÁRIAS, ALUGUERES E SERVIÇOS PRESTADOS AS EMPRESAS	580
L	ADMINISTRAÇÃO PÚBLICA, DEFESA E SEGURANÇA SOCIAL OBRIGATÓRIA	415
M	EDUCAÇÃO	8 605
N	SAÚDE E ACÇÃO SOCIAL	719
O	OUTRAS ACTIVIDADES DE SERVIÇOS COLECTIVOS, SOCIAIS E PESSOAIS	505
Q	ORGANISMOS INTERNACIONAIS E OUTRAS INSTITUIÇÕES EXTRA-TERRITORIAIS	327
	Soma	24 320

In areas where the businesses are mostly hosted in well-defined sites and/or inside building or other permanent structures a high number of the establishments are expected to be found also in the "old" register, while the coverage of the old register is significantly lower in other areas. The table below shows the number of establishments by district in Maputo City. A substantial part of the population lives in the districts 2-3, but the number of registered establishments in these areas is relatively small.

Table 3. Number of establishments in Maputo City in the old business register.

CodProv	CodDist	Estabelecimentos	Distrito	População 1997
11	1	4 237	DISTRITO URBANO Nº 1	154 284
11	2	769	DISTRITO URBANO Nº 2	162 750
11	3	234	DISTRITO URBANO Nº 3	210 551
11	4	161	DISTRITO URBANO Nº 4	228 244
11	5	289	DISTRITO URBANO Nº 5	211 008
11	6	7	Catembe	
11	7	1	Inhaca	
		5 698		966 837

The address field is not always suitable to find the establishment. For 1214 of the 5698 "active" establishments in the business register for Maputo City, the address is just "Cidade de Maputo".

The estimates of the number of establishments that will be included in the CEMPRE and thereafter entered in the new business register vary between 40000 and 100000. The number depends very much on how many of the small businesses in the commercial sector that will be covered.

Considering the difficulties in verifying if an establishment is in fact the same as one of the existing, INE decided to use the old business register only as a tool for checking the listing phase of the field-work.

It might have been feasible to verify a large proportion of the larger enterprises in the old register, write the registration numbers in the old register on the questionnaires and use the CEMPRE as one among other sources for updating the business register. That would have enabled INE to reduce the amount of data entry via the keyboard, but such a solution would have demanded more planning and more time for testing. Besides INE has a very limited capacity for development and maintenance of computer systems and also limited experience in organizing such a large update of a computerized register.

The main use of the old business register was therefore only as a tool to verify if the establishment exists during the listing process

4.4 The target population

The target population is defined in the plan of the director of the CEMPRE. The main criteria is to include every workplace except

1. Families with household employees
2. "micro"-businesses with ambulatory "headquarters"
3. Small and medium-size establishments in the sector of agriculture, fishing and forestry.

For market areas it is decided to consider the whole market area to be one enterprise (and establishment?).

An operational definition that has been applied in the pre-listing phase is to include all small establishments who are due to pay municipal tax. It was not feasible to explore in detail how this works in practice. Many businesses are in fact very difficult to identify in the field, especially in the service sector. There will also be a number of borderline cases, for which there will be need for further guidance to the interviewer in order to establish a uniform implementation of definition all over the country. Some of these cases can be foreseen and

included in the training of the interviewer, but other cases will have to be decided upon on an ad hoc basis.

If might also be considered to simplify the interview considerably for the smallest establishments.

Typical cases that will need to be clarified how to treat are for example farmers who also process their products or provide services as their major activity. Farming is expected to be a frequent secondary activity in all rural areas.

The target population includes all parts of the state administration, though not the core target population. It should be further explored if the information that could be retrieved electronically, for example data on schools from the Ministry of Education, is sufficient for the business register.

4.5   The routines for the data collections

The consultant made interviews with the managers, the supervisors, the controllers and a sample of interviewers. There are detailed manuals both for the interviewers and for the controllers, which ere reviewed and discussed in relation to case studies. The details on the logistical part of the operations are also carefully planned and there is a separate manual for distribution of material, organisation of the field work teams, distribution of cars and other parts of the logistics for the operations. In the initial phase of the field work in a province, there are a number of meetings and other contact taken at the province level to inform other parts of the administration about the operations. The material with information about the census that is distributed in the surveyed areas at the start of the listing procedure seems to be well suited for its purpose. At national level INE has prepared the census by contacts with the media and the results has been publicity that has served to facilitate the field operations.

One day during the first week of the mission was spent of field visits, both to discuss with interviewers and to discuss possible problems in classification and the definition of the population of enterprises to be covered. The field-work was also followed by frequent visits to the headquarters of the data collection for Maputo Cidade, which was located in the same building as INE.

4.6   Pre-listing

The pre-listing phase is very important for the coverage of the CEMPRE. Before the listing starts, the INE distributes lists of the contents of the "old" business register order by district and address to the field work supervisors (controladores). Staff from the local administration are contributing to the listing phase by use of lists and/or knowledge of establishments that are due to pay local tax (taxa municipal). The main criteria for defining if a business is an "enterprise" or not, is if it is registered to pay this tax or not. Maybe the field work for CEMPRE may cause changes in these lists as well, and indirectly affect the field work for the CEMPRE. Such possible effects may have to be further discussed.

The supervisors enter the lists from the pre-listing into the computer to produce lists that are more convenient for checking if all the listed establishments have been interviewed. This is in a way a redundant part of the work, but there were no discussions about the need of this work or on how to ovoid it. The view of the consultant was that the lists that are filled manually are sufficient for checking the coverage.

4.7   The questionnaires

It was early decided to introduce a new software, FUE, for both data entry in the CEMPRE and for the maintenance of the business register. The design of the questionnaires was prepared to fit the data entry routine in the FUE software. The design was prepared by assistance from the consultant Ms Julia Cravo from INE Portugal and the three main questionnaires for the data collection were prepared in July 2002. After that INE Mozambique has removed the questions about the official registration number NUIT, the initial investment and annual salaries. A new question about HIV/AIDS has been added.

There could have been closer links to the data processing routines in the design of the questionnaires. In the "Empresa" questionnaire there could have been separate sections for the data related to the enterprise itself (the juridical unit) and those that are logically properties of the headquarters establishment (the geographical/organisational unit). The two questionnaires "Empresa" and "Instituições sem fins lucrativos e administração pública" could also have been better coordinated, for example by avoiding duplication of the codes for "Forma juridical".

There are also two additional questionnaires, one for enterprises in manufacturing industry and one for restaurants and hotels. The main objective of these questionnaires is to collect information requested by other ministries. For these questionnaires INE will develop an own data entry routine in Microsoft Access.

The main observation in relation to the questionnaires is that the objective of each question could be described more in detail. The discussion on this part is more elaborate in the section about data quality and in the separate document "Rotinas para o processamento do CEMPRE" (Appendix 5). The key message is that the concept has to be clear and all measures to simplify the data collection should be taken. If the objective of a question is only to classify the enterprise into 4 or 5 groups it is better to let the interviewer make the classification instead of collecting more detailed information and make the classification afterwards.

The contents of the data collection will also have to be questioned during the process. If the experience from the first weeks of the data collection shows that the quality of one variable will not be sufficient to enable the planned use of it, there is need to decide if the quality can be increased. Otherwise the question should be excluded as soon as possible in order to avoid redundant work in the data collection and in the processing.

4.8   The plan for the data processing

The new software, FUE, for the data entry in the CEMPRE is written in Visual Basic and linked to database in Microsoft Access. The abbreviation FUE is here used to refer to the software provided by INE Portugal, but in the future it will probably also refer to the business register itself.

The main structure of the census is prepared for the software FUE. The design of the questionnaires was prepared in July accordingly. The FUE was planned to be installed before the pilot survey, to be tested and refined during the processing of the pilot survey, but the whole process was delayed and the pilot survey was not processed. When the first version of the FUE was installed 14-18 October, there were therefore still a number of changes in the software that remained to be implemented.

All the data entry will be made in 6 computers in the INE headquarters, were the data entry staff are planned to work in two shifts. The field-work operations are planned to finish in December while the data entry is planned to be completed in February 2003. If 6 computers are used for data entry 12 hours per day until the end of February (about 90 working days) there are about 1100 working hours available, which means on average 9 establishments per hour if the total number is 100 000 establishments.

The result of the tests of the FUE during the period 14-18 October was that INE decided to postpone the start of the data entry until a new version had been delivered.

It was decided not to ask INE Portugal to include fields for the questions on the number of staff member who had died during 2001 and if the enterprise had a program to increase the awareness of HIV/AIDS in the main forms for the FUE software. Instead INE will prepare an parallel data entry routine and link the data to the FUE afterwards. The reason for not asking INE Portugal to include these changes was said to be political.

Some of the most urgent changes in the FUE software were made already during the mission on the installation of the FUE by the consultant Mr Fernando Carvalho. The remaining observations were compiled into a memo, discussed in INE-M and then a request for updating of the FUE was sent to INE Portugal. The new version was expected to be delivered within a few days only. In some cases it would be possible only to change the contents of a field that is not used for the CEMPRE, but a more tailor-made solution is always to be preferred.

The consultant participated in the tests of the FUE and communicated his observations to Mr Calado Fijamo, who prepared to request to INE-P.

As mentioned above, the questionnaires have been designed before the decision on how to use the data and how to make the data entry. An example is that the decision on whether to codify the location of the establishments on three or four administrative regional levels was taken during the last week of the mission.

4.9   The data entry routines

The new application for the business register, FUE, was installed during the second week of the mission. The FUE will also serve as tool for the data entry for the parts of the CEMPRE that go into the business register. After the installation, the consultant assisted in the testing process and prepared a document with observations and proposals for INE to refine and use for their contacts with the programmer in INE Portugal.

The work on preparing a manual for the data entry was not completed during the time of the mission. The consultant gave some proposals on the contents of the manual. The final version will not be produced until the software has been updated. There may be need for ad hoc solutions depending on how the new version of the software works.

4.10 The data quality

The data quality is a key issue. If the expected quality is not sufficient to serve the objective for which the question was designed, there is need for measures to improve the quality and if that is not feasible it is better to drop or redesign the question instead of proceeding with the data collection and then just not process the data.

During the process of defining the contents and designing the questionnaires it is very important to clarify the objectives of each piece of information. If a numeric variable is intended to be used only as an indication of the size of the number it is better define the groups in the questionnaire and present the options in the interview. Otherwise the interviewed person may feel uncomfortable if he or she is not able to know remember the correct number. Furthermore the data entry less time consuming and the risks of typing errors in the data entry will also be reduced. With a limited number of codes it is also easy to apply a test that the entered code is valid.

In the CEMPRE there are questions about the value of the sales during 2001 and the number of staff members in November 2001. The objective was to produce an indicator on the turnover per staff member, which is typically a case where the objective is to produce statistics. However, the number of staff members is also a general indicator on the size of the establishment, which may be needed for stratification when the business register will be used as frame for sample surveys. From the consultants point of view, the second objective is more important, and then it would be desired to collect as recent information as possible on the number of staff members.

The following list contains a short discussion of the most critical variables.

Field name or description	Observation/Comment
Localidade	Localidade is a text field. There is no strong reason to codify the location to more than three regional levels. Instead it is recommended to format the address field in a uniform way to show the regional hierarchy below the third level that way.
Cidade, vila ou Aldeia (Bairro)	See "localidade" above
Telephone	The telephone number has separate fields for the area code and the number, which means that only one telephone number can be entered.
Fax	Same as telephone.
Coordenates: Latitude and Longitude	Not needed for statistical purposes, since district is the lowest feasible regional level also when data are published as maps. May by a solution for identifying the establishment in the field given the updates are made by using GPS.
Grupo de contribição Industrial	The code remains in the quetionnaire, while it has been decided to drop teh question. Therefore the field is available to be used for an other classification. The proposal is to make a sub-classification of all enterprises in three groups. A.Large businesses. B Other formal businesses. C. small and semi-formal businesses.
Year of foundation	The year of foundation is a difficult variable and closely related to the definition of the enterprise. The year that is expected to be reported is when the current owner started to run the business. It is important to clarify for the interviewers how the year should be defined. If the purpose is mainly to verify if the business is well established or not, it may be sufficient to classify in broader groups, and at least give "before the year xxxx" as one option.
Number of months in action during the year 2001	It is recommended to verify if this variable will be useful for the task for which it was designed. If not, it could be excluded.
Year of change of the situation of the enterprise	Not planned to be included in the database. The question can be dropped immediately.
Type of secondary economic activities	The manual for the interviewer does not tell anything about exclusion of minor activities. In principle only the main economic activity will be used for the classification of the enterprise. The proposal is to ask only for activities that contribute with at least 10 percent of the value added. If the data refer to a small enterprise there is probably no need at all to ask about secondary economic activities
"Capital social"	The "Capital social" may not be a very good indicator on the size of the enterprise, but is officially recognized. If it is interesting to classify if an "enterprise" has a "Capital social" it would be required to specify how to discriminate between "zero" and unknown".
Percentage distribution of "Capital social" by category of ownership	To ask about single percentage units for each category of owner is probably more ambitious than necessary. For classification purposes it is probably sufficient to classify type of ownership into 7 categories, the three pure categories, the 3 combinations of two, all three categories and "unknown". The information is normally quite easy to retrieve.
Number of staff members November 2001	For classification purposes a more recent reference month would have been preferred.
Number of worker who died during 2001	It should be questioned if the question gives reliable information.
"Volume de negócios"	For classification purposes it would be sufficient to classify the volume into a limited number of categories.
Program against HIV/AIDS	If this remains the only question to enter in a separate data entry routine it is not recommended to establish a separate data entry routine. One solution could be to (temporarily) use the field for another variable, for example GCI or type of questionnaire.

4.11 Archive of questionnaires

The enterprise identity in the new business register is given by the data entry application FUE (as "Autonumber"), which means that the number and the order of the number is defined by the order of the data entry. Since 6 operators are planned to work simultaneously, the questionnaires of one operator will not be numbered in sequence. The operator writes the number on the questionnaire after it has been defined by the computer, to easier go back to the questionnaire to check the information in the computer or to check information that had not been entered in the computer, such as contact person or telephone number to the contact person. Probably it is easiest to archive the questionnaires in the order of the enterprise ID, or rather by province and enterprise ID, depending on how the data entry is organised.

4.12 The future business register

The new business register does not mean a significantly higher ambition in terms of the amount of data that is planned to be included for each enterprise and establishment, but the ambition is significantly higher in terms of coverage. Increasing the coverage may call for other routines for the updates. Already in the old register, there were problems with the updates and when the register is extended, the updates will demand more resources. Currently there are 1-2 persons in the INE headquarters working with the maintenance of the register. Therefore the major part of the work on the updating the register will have to be done in the province offices, which has also been the case for the annual updated that have been conducted earlier. One concern has earlier been that the ambition for the business register varies between the provinces. The solution may be to include (more) customized facilities in the register that are suitable for use in the province administration.

One such step, that is already taken, is to include codes for the localization of the establishments by "posto administrativo" (the third administrative regional level).

4.13 Composite classification of "small and semi-formal" businesses.

Many of the "new" establishments will be small businesses in the sectors of commerce and services, which are more or less out of scope for sample surveys etc. Therefore it would be a good solution to make a sub-classification of the establishments by size, to facilitate selective updates, listing etc.

Such a classification by size could be a composite measure considering the number of persons involved, the amount of money invested and the type of economic activity.

A separate category should be created for small businesses, operated only or mostly by members of the same family, without any larger investment, often without fixed structure or operated from the house where they live. Maybe the criteria for separating this category could be if there permanent employees. Such a classification should be implemented as soon as possible. The field GCI in the "Empresa" questionnaire could be used, but it is also a proposal to simplify the interviews for the small enterprises, and if possible make a shortened interview already during the pre-listing procedure.

It is recommended to make some theoretical development and field observations to develop the definition of the new category, here called "Small and semi-formal" businesses.

A. Pay municipal tax?	Yes
	No
B. Capital social	No
	Yes , < x million MZM
	Yes, > x million MZM
C. Permanent staff	Only family members
	At least one staff member with fixed salary and or employment contract
	1-4
	5 or more contracted employees
D. Physical structure	Ambulatory
	Only simple structure that needs only short time to build/restore
	Permanent structure in combination with the owner's house
	Permanent structure not combined with the owner's house and/or specially constructed for the business purpose
Juridical status	Officially recognised by all the administration.
	Officially recognised and authorized by part of the administration
Type of economic activity	Commerce or services
	Home production for direct sale to end-users
	More advanced structure

5 SUMMARY OF CONCLUSIONS AND RECOMMENDATIONS

The main objective of the evaluation was to identify if each part of the system for data collection and processing of the CEMPRE can or could have been improved in a specific way to contribute to the total result of the operation, -the new business register for Mozambique. The steps that have already been taken are more or less history, but in some cases there is still an option to take a step back and move forward along another track.

The main conclusion is that the CEMPRE is a little "off track", mainly due to the delay in the planning of the data processing. However, the track that has been chosen can still lead to the right goal, but it would have been an advantage if it had been paved more thoroughly.

The whole process of planning of the data entry, identification of output variables and planning of the data processing was delayed. It was still decided to proceed with the fieldwork according to schedule and therefore the planning of the data processing was not sufficiently integrated with the design of the questionnaires. Depending on the result of the discussions with INE Portugal on the design of the data entry routine for the FUE, the development of the data processing routines may therefore be more complicated than what would have been the case if the data collection routines and the data the processing routines had been more integrated from the beginning.

The pilot survey could have served as a better tool to improve the routines if it had been conducted one or two months earlier and used for testing of the new data entry routine. Now there was very little time to evaluate the contents before it was necessary to start printing the questionnaires for the main data collection.

One strong recommendation is to prepare a small "pocket manual" for the data collection to assure that the main decisions on changes in the data collection, that will be the result of the first weeks of the data collections, are communicated correctly to the interviewers and the field supervisors.

It is also recommended to clarify the objective of each variable, validate the expected data quality and explore the possibilities to improve the quality if it not expected to be acceptable. The action to take would in some cases be to try to simplify the question, if the conclusion is that it is not worth the effort to proceed with the current design. In some cases the conclusion may be that the data collection for a specific variable should be discontinued, -as was earlier the case for the variables NUIT and "Grupo de Contribuição Industrial".

To enable INE to split the data entry for the main questionnaires into two separate routines and combine the data afterwards, it is necessary to include the new enterprise number in the questionnaires. To include the number is also a good habit and necessary if INE wants to return to the questionnaire to check other information that is currently not included in the data entry, such as the name of contact person or just to verify if the data entry was made correctly. One reason to return to a questionnaire that has already been entered in the register is to match the questionnaire for an enterprise with the questionnaires for its establishments, which may arrive from another province long time after the data entry for the main questionnaire.

It is recommended to further explore if other possibilities to collect information about establishments belonging to the state administration would be satisfactory for the business register. There are for example more than 8000 schools that report to the Ministry of Education. In many remote areas the school may be the only work-place of any importance,

which means a high cost for a data collection that is not the core objective of the enterprise register.

Field supervision in Matola, Província de Maputo

Mr Calado Fijamo, Mr Cirilo Tembe and Mr Fernando Carvalho during the tests of the FUE software

APPENDIX 1. Persons met

<u>National Institute of Statistics</u>
Mr Valeriano de Levene   Vice-president of INE

<u>Directorate of Integration, Co-ordination and External Relations (DICRE).</u>
Luís Mungamba   Deputy Director
Calado P Fijamo   System development, DISI

<u>Directorate of Sector Economic Statistics (DESE)</u>
Azarias Nhanzimo   Director
António Júnior   Head of the Department for Goods and Environment
Adriano Atanásio Matsimbe   National Coordinator of the CEMPRE
Cirilo Tembe   Deputy National Coordinator of the CEMPRE
João Nhabete   Gabinete Central do CEMPRE
Maria Teresa Tovela   Controller (Field supervisor)
A number of other staff members not mentioned by name

<u>Delegação Provincial de Maputo Cidade</u>
Ms Natércia Macuáco   Supervisor
Mr Remígio Nkondya   Deputy Supervisor
A number of other staff members not mentioned by name

<u>Delegação Provincial de Maputo Província</u>
Ms Leia Macamo   Supervisor
A number of other staff members not mentioned by name

<u>INE Portugal</u>
Mr Fernando Carvalho   Consultant

<u>Scandinavian Bridging Support programme</u>
Mr Hans-Eric Altvall   Team-Leader
Ms Irene Tuveng   Consultant, Economic Statistics
Mr Bo Yttergren   Consultant, Information Technology

APPENDIX 2. List of Literature

INE: Censo de empresas. Plano Director do CEMPRE
INE: Censo de empresas. Manual do Inquiridor (Aug2002)
INE: Censo de empresas. Manual do Controlador(Aug 2002)
INE: Censo de empresas. Manual da Organização e Logísticas (Aug 2002)
INE: Censo de empresas. Relatório da realização do piloto(Sept 2002)
INE: Censo de empresas. Relatório da fase inicial do CEMPRE na província de Inhambane (Oct 2002)
INE: Censo de empresas. Relatório da fase inicial do CEMPRE na província de Gaza (Oct 2002)

INE Portugal: Report from mission on preparation of the Business Census (Julia Cravo, July 2002)
INE Portugal: Manual do utilizador do FUE (Rute Brandão, Jan 1998)

Mr Adriano Matsimbe, National Coordinator of the Business Census

APPENDIX 3. Programme for the Mission

A number of meetings were held during the mission, most of them informal.

APPENDIX 4. Terms of Reference

TERMS OF REFRENCE
within the Scandinavian bridging Support Program

For a short-term mission 7ᵗʰ of October – 25ᵗʰ of October
on

Evaluating the initial field-work of the Business Census

1. **Background.**

Mozambique is in need of a better Business register. The existing register is presumed to be of very low quality both in terms of coverage and quality of existing data. A Business register should contribute to a relevant and reliable description of the Mozambican economy. In a statistical system it is one of the three corner stones, as it provides a basis for all surveys within economic statistics. It is hence crucial that this register gives reliable information about the structure and development of the economy.

To help improve the Business register Instituto Nacional de Estatística (INE) is in the process of conducting a Business Census (CEMPRE). The main objective of CEMPRE is to provide updated information about enterprises and establishments in Mozambique.

2. **General objective of the mission**

The general objective of the mission is to evaluate the first phase of the Business Census. Not only as an isolated project, but as a part of the building of a better Business register. Based on the evaluation the short-term consultant will suggest a list of necessary steps to be implemented to assure that the continuance of the CEMPRE is done in the best way possible.

3. **Specific objectives of the mission**

The specific objectives of the mission will be to:
- **Evaluate the data collection process.** The total process (from the survey plan, interview procedures, to the logistics of the collecting process) should be evaluated.
- **Discuss data quality of different types of data collected.** This includes viewing the questionnaires, the manuals/instructions, and checking the interpretation of these documents amongst the interviewers.
- **Evaluate the data processing routines.** In terms of efficiency in the handling of data from CEMPRE and in terms of flexibility towards the "data-handling-system" of economic statistics at DESE as a hole.

4. **Expected results**

The mission should result in a report including;

- An evaluation of the first phase of the data collecting process, and suggestions to improvements
- An evaluation of how manuals and descriptions are understood amongst the data collectors, and expected effects on the quality of different data. Suggestions to improvements
- An evaluation of the data processing routines, and suggestions to improvements

This evaluation- and analyzing work will be based on the condition that the over all goal of the CEMPRE is to improve the Business Register at INE.

Other outputs of the mission will be:
- Improved capacity and competence amongst the INE-staff working with CEMPRE and the Business register

5. **Draft agenda for the mission.**

Date		Activity
7.	Oct.	Welcome and introduction
8.	Oct.	Working at INE. Preparing field trip
9.	Oct.	Fieldtrip in Maputo province. Viewing data collecting routines.
10.-11.	Oct.	Working at INE
12.-15.	Oct.	Fieldtrip in Inhambane province.
16.-18.	Oct.	Working at INE.
19.-20.	Oct.	Time off
21.-23.	Oct.	Working at INE
24.	Oct.	Presentation of the mission report
25.	Oct.	Reviewing comments to the report

6. **Tasks to be done by INE to facilitate the mission**

- Prepare ToR and the agenda for the mission
- Arrange necessary working condition for the consultant
- Provide the consultant with relevant background papers (plans, reports and documentation)
- Plan and organize the fieldtrip
- Arrange necessary meetings

7. **Name of Consultant and Counterparts**

The short-term consultant will be Mr. Kenny Petersson from Statistics Sweden (SCB). His counterparts will be:
    Dr. Azarias Nhanzimo – Director do Direcção de Estatísticas Económicas Sectoriais
    Dr. Adriano Matsimbe – Coordenador Nacional do CEMPRE
    Dr. Cirilo Tembe – Coordenador Nacional Adjunto

The consultant will have close assistance from Mr. Francisco Rodrigues from   Cap Verde who will be supporting the Business Census as a resource person in the period 7th to 25th of October 2002.

8. **Timing of the mission**

The mission will take place in the period from the 7th to the 25th of October 2002

9. **Finalization of the report**

The consultant will prepare a draft report to be discussed with INE before leaving Maputo. He will submit a final draft to INE for final comments within one week of the end of the mission. Statistics Denmark as Lead Party will print the final version within 3 weeks of the end of the mission.

These Terms of Reference were prepared by (date and name)

24 /09 /02   Director de DESE – Dr. Azarias Nhanzimo

Approved by/in the name of the President of INE (date and name)

24 /09 /02   Vice-President Área Económica – Dr. Valeriano Levene

APPENDIX 5. Rotinas para o processamento do CEMPRE

1   Antecedentes

Em princípio uma grande parte do processamento do CEMPRE é uma actualização do cadastro de estabelecimentos. Além disso o conteúdo do cadastro vai aumentar-se usando novas variáveis recolhidos pelo CEMPRE. O objectivo principal é melhorar a cobertura e a qualidade dos dados do cadastro.

O desenvolvimento da estratégia e das rotinas para entrada de dados para esta actualização eram atrasados, aguardando a instalação de um novo software, FUE.

O INE tem assistência técnica de INE Portugal para o processamento do CEMPRE. O FUE vai servir como ferramenta para uma parte da entrada de dados no CEMPRE bem como para gestão do cadastro no futuro. A instalação do FUE foi executado durante a semana 14-18 de Outubro 2002. A assistência técnica de INE Portugal também cobriu o conteúdo e o desenho dos questionários, principalmente para elaborar questionários que correspondem à estrutura das janelas para digitação no FUE.

O CEMPRE também conte variáveis que não entram no FUE. As expectativas do FUE foi inicialmente que a o INE pode mudar o FUE para usar como ferramenta para todos os dados nos formulários. Se a posição de INE Portugal é que não pode incluir estes campos na parte de entrada de dados, seria necessário elaborar uma nova aplicação para a entrada de dados para estes restante partes do CEMPRE. O INE Moçambique comprou a aplicação mas tem autorização muda-la.

Sobre a estratégia do processamento foi decidido entrar todos dados de novo pelo teclado. O INE não tem experiência da organização de grandes actualizações de cadastros deste tamanho e os recursos humanos na área de informática são limitadas. De organizar o processamento como uma actualização do cadastro existente para aproveitar dos dados do velho cadastro também exige mais tempo e para preparações e testes.

Elaborou-se um pequeno piloto para testar os questionários e para treinamento do pessoal. O piloto foi realizado somente uma semana antes de começar o levantamento principal, e não cobriu as rotinas para o processamento. O tempo entre o piloto e CEMPRE geral não permitiu aproveitar das experiências do piloto para melhorar a estrutura e o desenho dos questionários.

2   O conteúdo do CEMPRE

O conteúdo do CEMPRE serve no primeiro lugar os três objectivos principais de qualquer cadastro de empresas:
   A. Identificar as unidades económicas e físicas
   B. Classificar as mesmas por tipo de actividade principal, forma jurídica etc.
   C. Recolher dados para facilitar o contacto (endereço, número de telefone etc.)

Um objectivo secundário é recolher dados sobre programas de combate e sensibilização sobre HIV/SIDA.

As perguntas dos questionários são o resultado do processo:

   1) Qual é a informação desejável?
   2) Como definir e classificar cada parte desta informação?
   3) É provável que o entrevistado pode responder?
   4) É provável que a qualidade dos dados vai ser suficiente?
   5) Existe um risco que a pergunta vai afectar numa maneira negativa, a entrevista em geral e/ou as respostas de outras perguntas no questionário?
   6) É comparativamente caro/trabalhoso recolher esta informação?
   7) Qual é a importância relativa desta informação?

Muitas vezes a conclusão é que uma pergunta desejável não pode entrar no questionário.

Os questionários do CEMPRE mostram os resultados deste processo, mas seria desejável incluir mais testes dos questionários preliminares antes de fixar o desenho definitivo daqueles.
Testes servem, com grande peso, como ferramentas para justificação para excluir perguntas que são o resultado de pedidos externos, quando as testes mostram que não é possível atingir uma qualidade suficiente para servir os objectivos da pergunta.

No caso de CEMPRE foi decidido excluir as perguntas sobre o NUIT e "Grupo de Contribuição Industrial" depois do piloto.

A estrutura geral dos questionários "Empresa" e "Estabelecimento" foi elaborada usando o levantamento para o FUE de Cabo Verde 1997 como padrão. O questionário para "Instituições sem fins lucrativos e Administração Pública" é uma versão do formulário "Empresa", mas utiliza-se também para unidades pendentes, por exemplo centos de saúde e a administração provincial.

O questionário "Empresa" conte também informação sobre o estabelecimento que é a sede da empresa, e quando existe somente um estabelecimento, não utiliza-se o questionário "Estabelecimento". A estrutura do questionário "Empresa" não mostra directamente se a natureza da informação é sobre a empresa ou sobre o estabelecimento sede.

Os outros questionários contem informação auxiliares sobre estabelecimentos das actividades económicas de indústria, restaurantes e hotéis.

3 Objectivo e qualidade de cada variável

Algumas das variáveis pertencem a mais do que um dos grupos por tipo de objectivo. "Província" utiliza-se para classificação e também para identificação. É importante clarificar o objectivo de cada variável e definir a exigência sobre a qualidade considerando o objectivo.

Se toma-se a localização como exemplo, tem que definir qual é o nível administrativo mais baixo para digitar e usar como fonte de tabelas estatísticas. No caso do CEMPRE faz-se digitação da localização por nível de província, distrito e posto administrativo. A outra parte da informação sobre a localização do estabelecimento utiliza-se para identificar o estabelecimento, mas para facilitar a utilização dos dados para identificação e também necessário discutir como usar estes. Se um objectivo é tornar possível elaborar listagens dos estabelecimentos por endereço, é necessário preencher e digitar os endereços numa maneira uniforma.

Toda a discussão sobre o objectivo de cada campo do questionário tem que ser concluída antes de elaborar o desenho final do questionário.

No CEMPRE existe perguntas sobre o valor de vendas do ano 2001 e o número do pessoal em novembro 2001. O objectivo foi tornar possível elaborar cálculos do indicador "Volume de negócios por membro do pessoal 2001". Um outro objectivo para recolher dados sobre o número do pessoal é indicar o tamanho do estabelecimento e o valor da produção. Para este objectivo seria melhor recolher dados mais recente como possível sobre o pessoal, para indicar a situação actual. Também encontra-se estabelecimentos novos, que não foram activas no ano 2001. Para estes recolha-se dados sobre o número de pessoal actual, mas faz-se a digitação no mesmo campo que utiliza-se para o número de pessoal 2001 e somente indicar a diferença no campo "Observações". A informação do campo "Observações" não permite facilmente discriminar entre os dois casos no processamento.

Um outro exemplo é "Ano de fundação". Conforme o manual o ano de fundação indica o ano de fundação da empresa ou o estabelecimento e não com o da situação perante a actividade. Significa normalmente o ano quando o empresário actual foi registrado como dono da empresa ou o estabelecimento, mas pode talvez referir-se ao ano quando o edifício foi construído ou quando começou actividades do mesmo tipo de actividade económica da mesma localização. Por meio de clarificar o objectivo para os inquiridores, o campo vai ser mais útil para o seu objectivo. Se o objectivo somente é indicar se é uma nova empresa (ou um novo dono) pode mudar o desenho e talvez agrupar as respostas. Talvez não é importante se o ano é 1986 ou 1987. Talvez é suficiente classificar como "Antes de 1990" para facilitar as entrevistas.

Estes são somente exemplos das etapas do desenvolvimento do conteúdo do inquérito, que são importantes terminar antes de definir o desenho final dos questionários.

Se as conclusões das primeiras semanas de recolha de dados mostra que não é possível atingir a qualidade necessária para o objectivo para uma variável específica é necessário decidir como tratar. Em princípio existe duas opções, - melhorar a qualidade ou alterar os objectivos.

Se a conclusão é que custo para melhorar a qualidade não corresponde à valor da informação é melhor excluir a pergunta imediatamente (para evitar trabalho redundante para os entrevistados, inquiridores e digitadores).

4 Actualização do conteúdo do levantamento

Elaborou-se alterações do levantamento, que não foram introduzidos como novas versões dos questionários. Por isso, existe diferenças entre os questionários o e conteúdo do CEMPRE.

Uma documentação detalhada sobre o conteúdo é uma ferramenta importante para evitar confusão, primeiro e mais importante na recolha de dados, mas também para o desenvolvimento das rotinas para o processamento.

Recomenda-se elaborar três documentos deste objectivo:

1. Uma guia do bolso para os inquiridores e controladores com informação sobre todas as alterações introduzidos. Seria bom incluir as partes dos manuais que não mudam-se e que contem informação importante durante o trabalho no terreno.
2. Uma listagem de todas as variáveis com códigos validos. Este documento serve como manual para codificação e para os digitadores.
3. Uma versão mais detalhada sobre as variáveis que também dá informação sobre a estrutura da base de dados, os nomes dos campos na base de dados, os nomes do campos correspondentes no velho cadastro (e também os códigos correspondentes no velho cadastro se são diferentes).

E importante tratar estes três documentos como uma unidade para evitar coincidência entre a recolha de dados e o processamento.

Para preparar o trabalho indicado no ponto dois comparou-se a estrutura do cadastro e o conteúdo do CEMPRE para elaborar propostas sobre a estrutura do novo cadastro.

5 Estrutura lógica da aplicação

As entidades principais são EMPRESA, a unidade económica/jurídica e o ESTABELECIMENTO, a unidade de local de trabalho. A maioria das empresas tem somente um estabelecimento, que é também a sede da empresa. Se a empresa tem mais do que um estabelecimento, o estabelecimento sede é um estabelecimento entre os outros.
A estrutura lógica é que cada estabelecimento pertence a uma empresa.

Na aplicação FUE, a tabela "Empresa" conte também informação do estabelecimento sede (por exemplo códigos regionais do local de trabalho) e a tabela "Estab" conte somente informação sobre estabelecimentos quando existem mais do que um estabelecimento da empresa.

Recomenda-se tratar empresas e estabelecimentos como unidades separadas e incluir a informação sobre o estabelecimento sede na tabela "ESTAB" (correspondente à tabela Testablishment no cadastro). Para realizar esta alteração e necessário desactivar uma regra de validação que diz que não aceite somente um estabelecimento por empresa.

6 O FUE e o processamento do CEMPRE

Um problema no processamento do CEMPRE é que não foi clarificado antes de começar imprimir os questionários, se/que e como usar o aplicativo do FUE para digitação dos dados dos três questionários principais "Empresa", "Estabelecimento" e "Instituições sem fins lucrativos e Administração Pública". Desta maneira seria possível diminuir a exigência de mudanças no

programa para digitação, FUE, e evitar soluções provisórias, como usar um campo designado para uma outra variável.

Num documento separado elaborou-se uma comparação entre o conteúdo do CEMPRE e as facilidades do FUE para avaliar as opções para o processamento e preparar um pedido para INE Portugal sobre alterações no FUE para facilitar a digitação dos dados do CEMPRE.

Recomenda-se pedir o INE-P alterar o FUE para servir como ferramenta para a digitação de toda a informação dos questionários "Empresa", "Estabelecimento" e "Instituições sem fins lucrativos e Administração Pública".

Para tornar possível procurar um questionário qualquer por exemplo para juntar os dados do FUE e os dados dos questionários auxiliares foi decidido preencher o número interno do FUE no encabeçamento do questionário "Empresa" e cada outro questionário separado. O mesmo número vai utilizar-se para a digitação dos questionários auxiliares.

7   A organização física dos questionários

O código da empresa que mostra-se no écran quando a digitação está confirmada e que o digitador preenche no questionário é o melhor critério também para ordenar os questionários, quando o objectivo e comparar os dados dos questionários e os dados digitados ou ver a informação no questionário que não foi digitado. Para outras tarefas seria também desejável ordenar por exemplo por nome da empresa. Porque o código da empresa é um número que aumenta-se automaticamente é a ordem de digitação que dá o número. Significa que se deseja por exemplo ordenar por distrito e nome, recomenda-se ordenar os questionários antes de digitar. Quando 6 digitadores trabalham simultaneamente significa também que o computador dá o próximo número do ordem para um questionário de um outro digitador. Por isso é ainda mais importante que o digitador não lembre-se preencher o número no questionário.