Potential for Small Area Estimation

and Poverty Mapping

in Timor-Leste

Feasibility Report
Phases 1 and 2

Professor Stephen Haslett
Dr Geoffrey Jones
Massey University
New Zealand

20 December 2008

Abstract

Small area estimation (sae) is a mathematical technique for extracting more detailed information from existing data sources by statistical modelling. The estimates are often mapped, so the technique is often generically called poverty mapping. These maps and estimates (together with estimates of accuracy) are key information in aid allocation within a country. They are also increasingly important inputs to negotiations on allocation of international aid to particular countries.

Assessing feasibility of small area estimation for Timor-Leste is the topic of this report. The general conclusions are that:

- Small area estimation using the World Bank method, based on the Timor-Leste Survey of Living Standards (TL-SLS 2007) and the Timor-Leste Population and Housing Census (TLS) 2004 is unlikely to be warranted, because the two data sources do not match sufficiently well. The differences are due to an unusual combination of circumstances including:
    - internal migration between 2004 and 2007, for example internally displaced persons from the conflict in Timor-Leste which began in April 2006,
    - differences in structure between the detailed questions asked in the census and survey questionnaires.
- Nevertheless, if small area estimates of poverty are required soon, there may be benefit in researching and using alternative small area models based on the Timor-Leste Survey of Living Standards (TL-SLS 2007) only, since such models do not use or rely on the census questionnaire or data. Such alternative models may still be able to give poverty and kilocalorie estimates at a finer geographical level (via statistical modelling) than is possible using traditional design-based sample survey estimation methods. However, testing this possibility using statistical modelling was not possible as part of this feasibility study, because the Timor-Leste government, citing confidentiality, was not willing to release the TL-SLS 2007 data for the necessary analysis within the July 2008 to November 2008 timeframe of this feasibility study.
- Small area estimation done properly requires a high level of mathematical and statistical expertise. Release of the necessary data to bona-fide researchers under the usual internationally accepted conditions for detailed unit-record information from surveys and censuses, is required and remains a policy matter that requires urgent resolution by the Timor-Leste government. Unit record data release, without names and detailed addresses, will be required for the TL-SLS 2007 before any small area estimation of poverty at all (or even a complete feasibility study) will be possible in Timor-Leste.
- The Timor-Leste national census (TLS) is scheduled to be repeated in 2010. If it were to use sufficiently similar core questions to TL-SLS, and the TL-SLS were also to be repeated around the same time as the census, this would give an excellent basis for small area estimation of poverty (poverty incidence, gap and severity, plus kilocalorie consumption) for Timor-Leste post-2010, using the World Bank and/or a range of other small area estimation methods, provided unit record survey and census data plus geographical coding information were made available to researchers through the Timor-Leste Direcção Nacional de Estatística (DNE) and the necessary local expertise at DNE was also an integral part of this exercise..

Executive Summary

1. Small area estimation (sae) is a mathematical technique to extract more detailed information from existing data sources by statistical modelling. The methodology is important because it produces finer level information than is possible for a sample survey analysed by standard methods, for poverty related variables that are not collected in the census. Small area estimates are often illustrated geographically using poverty maps. The cost of small area studies can be saved many times over by having this better poverty information at a finer level for use in aid allocation. The method is also important for providing comparisons of poverty at a fine level in different countries.

2. This report provides a brief summary of a study, undertaken by statisticians from Massey University, New Zealand, to assess feasibility for small area estimation of poverty and poverty mapping in Timor-Leste at district or sub-district level. This report includes background and assessment, but does not include any statistical analysis of survey data which was to form part of this study, because the Timor-Leste government, citing confidentiality, was not willing to release the required unit-record survey data from the Timor-Leste Survey of Living Standards (TL-SLS 2007) within the July 2008 to November 2008 timeframe of this assessment. The feasibility study also does not include small area estimates of poverty based on both survey and census data; even had the necessary data been available, such estimation and the consequent poverty maps were always intended to be a separately funded exercise.

3. The particular aspects of poverty that are considered in the feasibility report are poverty incidence, gap and severity relative to the national poverty line, with an additional focus on undernourishment (as measured by kilocalorie consumption). While the last of these four strongly reflects the food security interests and concerns of the sponsor of the feasibility study, which is the World Food Programme, this report also recognises the importance to a wide range of international aid agencies of all four measures, and the centrality of the first three to sound economic measurement of poverty.

4. For Timor-Leste, the data sources considered were the Timor-Leste Population and Housing Census (TLS) which was conducted in 2004, the Timor-Leste Comprehensive Food Security and Vulnerability Analysis (CFSVA) survey conducted December 2005 to January 2006, and the Timor-Leste Survey of Living Standards (TL-SLS) conducted January 2007 to January 2008.

5. Because the required survey data was not released by the Timor-Leste government during the timeline for this feasibility study, this feasibility report could not include preliminary statistical model fitting using the 2007 TL-SLS. There are consequently no analyses in this report that provide detailed assessment, via statistical modelling, of poverty mapping feasibility.

6. Small area estimation done well requires a high level of mathematical and statistical expertise. Agreed release of the necessary data to bona-fide researchers under the usual internationally accepted conditions for detailed unit-record information from surveys and censuses, is a prerequisite and a policy matter that requires urgent resolution by the Timor-Leste government. Unit-record (ie household and individual level) data release, without names and detailed addresses, but including detailed geographical coding information, and sample survey weights, will be required for the TL-SLS 2007 before any small area estimation of poverty at all (or even a complete feasibility study) will be possible in Timor-Leste.

7. Even without unit record data however, it is clear that small area estimation using the World Bank method, based on the Timor-Leste Survey of Living Standards (TL-SLS 2007) and the Timor-Leste Population and Housing Census (TLS) 2004 is not warranted, because the two data sources do not match sufficiently well. The differences are due to an unusual combination of circumstances, particularly:
   - internal migration between 2004 and 2007, for example internally displaced persons from the conflict in Timor-Leste which began in April 2006,
   - differences in structure between the detailed questions asked in the census and survey questionnaires.

8. The Timor-Leste Comprehensive Food Security and Vulnerability Analysis survey (CFSVA) was conducted by the World Food Programme between December 2005 and January 2006 so that, like the census, it precedes the April 2006 conflict. In principle, this removes the first of the concerns about using the 2004 census and the TL-SLS 2007, although using the CFSVA and the census could only give pre-2006 conflict estimates, and the unit-record 2004 census data would still be required from the Timor-Leste government. The CFSVA is however not suitable for small area estimation for two reasons. Firstly, it cannot be used for poverty incidence, gap and severity since it only collects food expenditure. Secondly, it collects expenditure information on certain foods, but no information on quantities consumed. Indeed, WFP's own CFSVA report comments that kilocalorie information is 'unavailable'.

9. Nevertheless, despite these somewhat negative findings, there may still be benefit in using alternative small area modelling based only on the Timor-Leste Survey of Living Standards (TL-SLS 2007) were this survey data to become available, since such models do not use or rely on links to the census questionnaire or data. Such alternative models may still be able to give poverty incidence and kilocalorie estimates for 2007 at a finer geographical level (via statistical modelling) than is possible using traditional design-based sample survey estimation methods. Such alternative small area estimation methods, since they rely only on survey data, would not be compromised by the major structural changes in Timor-Leste between the 2004 census and the TL-SLS in 2007.

10. The Timor-Leste census is scheduled to be repeated in 2010. If it were to use sufficiently similar core questions to TL-SLS, and the TL-SLS were also to be repeated around the time of the 2010 census, this should in principle give a strong basis for small area estimation of poverty (poverty incidence, gap and severity, plus kilocalorie consumption) for Timor-Leste, post-2010, using the World Bank and/or a range of other small area estimation methods.

11. For poverty estimates using the World Bank small area estimation method, conclusions about use of the 2010 census remain subject to resolution of some minor geographical coding issues. An integrated system linking these geographical levels and the availability of this coding information to all members of the research team undertaking the small area estimation will be required. The complications with boundary changes at enumeration area (EA) level between the TL-SLS 2007 and the 2010 census together with the three year time interlude after TL-SLS 2007 and the planned census, make these two data sources a somewhat unattractive option for small area estimation of poverty using the World Bank method.

12. No poverty estimates using small area estimation techniques were intended to be produced as part of this feasibility study based on the 2004 census, even if the TL-SLS 2007 data had been available. As noted above, such estimates at a minimum require the TL-SLS 2007 data, the 2004 census data, and further funding beyond the feasibility phase. Beyond the 2010 census, the new Population and Housing Census (TLS) data, new TL-SLS data, and the integrated geographical information will also be required.

13. If small area estimates are required in 2009, the TL-SLS 2007 data will need to be released by the Timor-Leste government, after which funding should be sought immediately from international donors for a small area estimation study in Timor-Leste based only on the TL-SLS 2007 data. Because no relevant statistical modelling of the TL-SLS 2007 data has been carried out to date, it is recommended any such project be divided into two phases, the first of which would extend the current study by assessing feasibility of alternative small area estimation methodologies and include the initial modelling of the TL-SLS 2007 data that is missing from this feasibility report.

14. The completion of this report follows consultation with the Timor-Leste Direcção Nacional de Estatística, other officials and members of the government at the Ministério de Finanças, UNFPA, the World Bank, and the World Food Programme (WFP) - which commissioned this research - in Timor-Leste, 3 - 14 November 2008. The authors are grateful for these contributions. Viewpoints expressed in this report do not however necessarily reflect those of all or any of the organisations consulted.

1.   Introduction

The small area estimation feasibility project was intended to be in two phases, which together provide an assessment of the potential for small area estimation of kilocalorie consumption in Timor-Leste. Neither phase was to include provision of small area estimates. As originally specified, the essence of the two phases was:

Phase One:

-       Analysis of existing research and information on food security necessary for statistical assessment of kilocalorie consumption in Timor-Leste, based on reference material to be supplied by the World Food Programme and existing knowledge of small area estimation methods. Note that the nutritional aspects of methods of estimating kilocalories from expenditure data were not to form part of the assessment.
-       Analysis of relevant Timor-Leste questionnaires from:
        (a) The population census (Census of Population and Housing, July 2004),
        (b) The national household income and expenditure survey (Timor-Leste Survey of Living Standards 2007), and
        (c) The Timor-Leste Comprehensive Food Security and Vulnerability Analysis (CFSVA: Dec2005-Jan 2006).
-       Identification and listing of questions asked in (a), (b) and (c) above that *prima facie* are similar enough to be used for small area estimation of poverty incidence, gap and severity, and kilocalorie consumption, with the investigation based on the English versions of questionnaires only.

Phase Two:

-       Identification of variables potentially useful for small area estimation of poverty incidence, gap and severity, and kilocalorie consumption.
-       Developing and testing preliminary statistical regression models (including estimation of variance components) based only on the survey data, which was to be supplied by WFP.
 -      Identification, in conjunction with the Timor-Leste Directorate of Statistics (Direcção Nacional de Estatística, Ministério de Finanças, DNE), of any Administrative Unit (area) code changes that may complicate any later analysis of the statistical relationship between survey and census data (outside the scope of this study). This was to be based on survey and census information supplied by DNE to WFP.
-       Comment on the potential impact of these statistical analyses on the feasibility of small area estimation of poverty incidence, gap and severity and kilocalorie consumption.

Phase two of this feasibility project involved in-country travel to Timor-Leste, planned in coordination with the WFP Country Office.

Completion of this report has included consultation and discussion in Timor-Leste with the following people and organisations:

Direcção Nacional de Estatística, Ministério de Finanças

Elias dos Santos Ferreira,
  Project Manager / Head of Research / Acting Head, National Statistics Directorate
Cristina Dasilva-Cruz, Communications and Management Facilitator

Ministério de Finanças

HE Rui Manuel Hanjam,
  Vice-Minister Economic and Development, and Acting Minister of Finance

UNFPA / Direcção Nacional de Estatística

Frederick Okwayo, Chief Technical Advisor, Census Project
Anouska Charles, Programme Coordinator, Population and Development Strategies

World Bank

Antonio S. Franco, Country Manager , Timor-Leste
Homa Z. Fotouhi, Senior Operations Officer, Timor-Leste
Gaurav Datt,
  Economist, Poverty Reduction and Economic Management,
  East Asia and Pacific Region, World Bank Office, Sydney, Australia

United Nations World Food Programme

Joan Fleuren, Representative and Country Director, Timor-Leste
Angelline Rudakubana, Deputy Country Director, Timor-Leste
Flavia Dasilva, National Programme Officer,
  Vulnerability Analysis and Mapping Unit, Timor-Leste
Michael Sheinkman *, Senior Regional Programme Advisor,
  Vulnerability Analysis and Mapping (VAM), WFP Bureau for Asia, Thailand

Note that, while additional people were approached, the limited period of the visits to Timor-Leste during November 2008 and the priorities of members of the Timor-Leste government and ministries meant not all were available. In particular, while meetings held were very helpful, the resources available from the Direcção Nacional de Estatística during this period limited detailed discussion, particularly of translation issues for census and survey questionnaires and changes to geographical coding over time. Nevertheless, every effort has been made to ensure as full a range of opinion as possible was canvassed and considered in the preparation of this feasibility report.

* Consultation on Timor-Leste undertaken in Thailand.

Background

2.1   Small Area Estimation

Small area estimation is a mathematical and statistical method that models data collected from one or more data sources, to produce estimates, for example of poverty, that are more accurate at small area level than using only data collected from each small area. The additional accuracy is achieved in many such models by "borrowing strength" for the estimate for a particular small area by using information from areas to which it is similar. Some small area estimation techniques combine data from different sources. Usually for small area poverty estimates, a statistical model is fitted to survey data collected around the same time as the census, and this model is used to predict a variable not collected in the census, based on shared variables that are collected in both survey and census. In poverty studies, the most usual variable predicted is income or expenditure (or its logarithm) based on a model which includes education, age of household members, number of people in the household and type of house construction, among other variables. The poverty estimates are often mapped in detail, which is why this technique is sometimes given the generic title, poverty mapping. The maps can make interpretation simpler, but the central point is not the maps *per* se, but that poverty can be assessed at a much finer level at a much lower cost than by increasing the sample size for the survey sufficiently or rerunning the census. The statistical modelling has a cost, of course, but this can be saved many times over by having better information at a finer level and maps for use in aid allocation.

Numerous small area estimation methods are detailed in Rao (2003). One method, documented, supported and advocated by the World Bank for small area estimation of income or expenditure poverty is that developed by Elbers, Lanjouw and Lanjouw (2001, 2003). This is now available as free software (PovMap – Zhao, 2006) from the World Bank website. By fitting a statistical model for expenditure (or income) on the logarithmic scale to the sample information, applying the model to the census data to predict expenditure (or income) at household level for all households, and summing back-transformations of the predictions, small area estimates using the World Bank method for poverty incidence, gap and severity can be derived, and mapped at small area level. Variations of the Elbers, Lanjouw and Lanjouw (ELL) method have been implemented for the World Bank in a number of other countries including Thailand (Healy, 2003), South Africa (Alderman et al., 2002), Brazil (Elbers et al. 2001), the Philippines (Haslett and Jones, 2005), and for the World Food Programme in Cambodia (Fujii, 2002), Bangladesh (Jones and Haslett, 2003) and Nepal (Jones, Haslett and Parajuli, 2006)

The Timor-Leste Population and Housing Census (TLS) was conducted in 2004, and the Timor-Leste Survey of Living Standards (TL-SLS) in 2007. This sets the stage for feasibility assessment of small area estimation of poverty. There has been no previous small area estimation of poverty in Timor-Leste.

One complication for small area estimation in Timor-Leste using theses two data sources and the World Bank method is the conflict that began in April 2006. The World Bank small area estimation method uses a statistical model fitted to survey data at household or individual level which is then applied to census data. This is why unit-record (ie household or individual level data) is necessary. Ideally survey and census are conducted at the same time, because the viability of the method is affected by changes that occur between data collection for census and survey (or survey and census if they are conducted in the opposite order). For Timor-Leste, the interval is three years which is rather long for use of the World Bank method, and the situation is additionally complicated by the conflict occurring between the two periods of data collection, especially given the number of internally displaced persons

(IDPs) that the conflict generated (variously assessed at between 30,000 and 100,000 or approximately 3% to 10% of the Timor-Leste population). This issue is discussed in greater detail later in this report.

The particular aspects of poverty that are considered in this feasibility study are poverty incidence, gap and severity relative to the national poverty line, with an additional focus on undernourishment (as measured by kilocalorie consumption). While the last of these four strongly reflects the food security interests and concerns of the sponsor of the feasibility study, which is the World Food Programme, this feasibility report also recognises the importance to a wide range of international aid agencies of all four measures, and the centrality of the first three to sound economic measurement of poverty.

2.2   General Background

The Millennium Development Goals (MDGs) provide a context for small area estimation, since local estimates of poverty and their updates have potential to provide fine-detailed, national monitoring information against which progress can be measured.

The República Democrática de Timor-Leste and United Nations Country Team (2004) *Timor-Leste Millennium Development Goals Report* provides a summary of progress toward the Millennium Development Goals in Timor-Leste up to 2004. The report covers historical background, development context, status on major goals, targets and indicators, challengers in addressing MDGs, and next steps. The status covers the eight Millennium Development Goals: eradicate extreme poverty and hunger, achieve universal primary education, promote gender equality and empower women, reduce child mortality, improve maternal health, combat HIV/AIDS, malaria and other diseases, and partnerships for development. Although it preceded availability of the 2004 census data, it was noted even then that 'There are significant discrepancies in the information on some indicators emerging from the surveys and those generated through administrative channels' (See p52). To the extent that this issue remains, it is additional justification for research on and provision of small area estimates of poverty in Timor-Leste.

The Asian Development Bank (2005) report, *ADB Country Strategy and Program Update: Democratic Republic of Timor-Leste (2006-2008)*, outlines the ADB funding and development priorities for Timor-Leste for the period 2006-2008. It covers country strategy, current trends and issues, implementation of country strategy and program, and portfolio management issues. It notes that 'no effective local government system has been established', p1. The 2005 Timor-Leste government 'describe[d] 2006-2008… as a period of consolidation as it deals with a decrease in external assistance', although that decrease was reversed in the aftermath of the 2006 conflict. For 2004, it also notes 'Timor-Leste's low institutional capability and absorptive capacity', that 'government officers lack experience, their offices are underresourced and have few technically able staff", and 'few staff that are able to absorb technical training'. The ADB report was written three years ago, but given the conflict that began in 2006, such problems must remain at least in part. Change takes time, so that even given improvements to the situation have been possible since 2004, strengthening government capacity to co-ordinate and manage remains an issue. Small area estimation techniques require a strong mathematical background and considerable time to understand in detail and implement. Nevertheless, improved technical skills within government is a potential side-benefit of any full small area estimation study, given sufficient government interest, co-operation, release of data, and commitment of human resources. Personnel with local knowledge remain essential for detailed analysis of survey and census questionnaire content.

The BELUN, FONGTIL and National Dialogue Team of the Office of the President of the Republic 2006 report *National Dialogue* (*Open Space*) held on 7-8 April discusses the 'open

space' meeting of NGOs and the Timor-Leste government to discuss NGOs and government working together; community development with active communication; development of the local economy; strengthening NGOs; increased access to information; strengthening governance; collaboration between NGOs and government to strengthen security and development; strengthening the health sector; reinforcing the water and sanitation sector; support for youth; opportunities for disabled people; strengthening human rights; and development of the technology sector. This potentially useful planning initiative was undermined by the complications and aftermath of the internal conflict in Timor-Leste which began in late April of 2006.

More directly related to food and food security there are a number of relevant reports.

The 2006 República Democrática de Timor-Leste: Ministry of Agriculture, Forestry and Fisheries report, *National Food Security Policy for Timor-Leste* links access, availability, stability, and effective utilisation to food security policy in Timor-Leste and the Millennium Development Goals.

Fedele and Horjus's July 2006 paper, *Timor-Leste: Comprehensive Food Security and Vulnerability Analysis (CFSVA): Conducted in December 2005-January 2006*, is a publication of the Vulnerability and Mapping Unit, World Food Programme. It provides an analysis of food security in Timor-Leste undertaken in December 2005-January 2006, before the unrest which began on 28 April 2006. This report is based on a sample survey of 1700 households across 163 sucos in 13 districts. It outlines WFP's Protracted Relief and Recovery Operation (PRRO), together with survey results and commentary on the socio-economic situation (population and demographics; economic characteristics including living conditions, main income sources and household expenditures); household food security and vulnerability (including availability and access to food); food consumption; sources of food insecurity problems; food utilisation health with a focus on maternal and child nutritional status; and household-level food insecurity profiling.

Forsén, Sanogo, Ayanahati,, Rasenen, and Dasilva provide a 2007 World Food Programme local update, *Dili Emergency Food security Assessment, Timor-Leste, September 2007*. Their report examines food availability, access and usage in Dili for different livelihood groups and geographical locations; estimation of the number of people who are food insecure via a sample of 613 households; assessment of whether food aid is an appropriate response and links to ongoing programmes and longer term food security. The results showed that 27% of households were spending over 65% on food (poor food security), 37% have average food access (50%-65% spent on food) and 36% had good food access (<50% spent on food); the sample chose approximately 300 households in sucos and 300 households containing internally displaced persons (IDPs). The report notes, as a result of the political unrest that began in April 2006 and was still ongoing in September 2007, that, 'the estimated figure of internally displaced persons (IDPs) remaining in the camps in Dili are (sic) approximately 30,000 with an additional 50-60,000 in districts'.

The World Food Programme currently supports an extensive school feeding programme in Timor-Leste, which recognises the link between food and education. The 2006 República Democrática de Timor-Leste, Ministry of Education and Culture, *Report of the First Annual Joint Review for the Education Sector*, outlines a two-day meeting in Dili, 18-19 October 2006, on development of Timor-Leste educational policy. It presents 'a bleak picture of the education scenario', p11. The meeting outlined two objectives: having a baseline and sub-sector indicators; and designing a three year rolling plan for primary, secondary and senior secondary education. Curriculum development at pre-secondary and senior secondary levels was recognised as a priority. It was also recognised that 'careful and systematic work…. [was required to] change the overall picture of the education program', p11. The report is however

a set of conference proceedings rather than a fully integrated education programme and framework, so that it lacks detail on the links that are required between education and other sectors of the economy.

On a more positive note, the Seeds of Life / Fini ba Moris *Annual Research Report 2006* recognises the need for and reports on agricultural research in Timor-Leste. This publication is a summary of the research programme and results for agriculture in Timor-Leste from the Seeds of Life (Fini ba Moris) programme within the Timor-Leste Ministry of Agriculture, Forestry and Fisheries. This report covers research on maize; sweet potatoes; cassava; rice; peanuts; jak, sword and velvet beans; and includes effects of fertilizer, weeds, and pests on yields.

2.3 *Timor-Leste Censuses and Surveys*

Beyond this general background, there have been a series of reports more directly related to small area estimation of poverty. These document the data collections necessary for the statistical modelling necessary to produce sound small area poverty estimates.

One core data collection required for small area estimation of poverty via the World Bank method is a national census.

The first Timor-Leste census was held in 2004. A brief sketch of the preliminary planning information is in the October 2003 National Statistics Office Timor-Leste document, *2004 Census of Population and Housing: Project Plan 2004 Census Project.* This report outlines the planning of and preparation for of the 2004 census, covering background, objectives, the basic approach, assumptions, organisation, training, project reporting, project milestones, dependencies, risk analysis, quality issues, resourcing and budget. The report is very brief and contains only some of the information necessary to run a sound census but is nevertheless a useful document for outlining the approach taken to some of the issues that needed to be resolved before the full census could be undertaken.

There is also a sequence of reports with results from the 2004 census. These include:

-   The Direcção Nacional de Estatística, Ministério de Finanças 2004 report, *Timor-Leste 2004 Census Project: Census Dictionary*, which notes, 'The 2004 Census Dictionary is a comprehensive reference guide designed to assist users of 2004 Census data to develop a better understanding of the conceptual issues underlying the data. This dictionary contains definitions of census terms and explains the concepts relevant to census collection, processing and output of data. Many of the entries in this Dictionary refer directly to questions on the 2004 Census Questionnaire. The census gathers information on a number of topics about persons, families and households.' The census dictionary is an essential reference source for assessment of feasibility of small area estimation using the World Bank method.
-   The 2006 National Statistics Directorate / United Nations Population Fund report, *Timor-Leste Census of Population and Housing 2004: Atlas*, which is mainly based on the 2004 census. It contains maps of and details about physical and administrative structure, general population characteristics, age structure, economy and employment, languages and education, and fertility and mortality.
-   The National Statistics Directorate 2006 report, *National Priority Tables: Census of Population and Housing 2004*, published in conjunction with the United Nations Population Fund. This report contains extensive tables from the 2004 census.
-   The 2006 National Statistics Directorate publication, *Population Projections 2004-2050: Analysis of Census Results, Report 1: General, Census of Population Census of Timor-Leste 2004*, again published jointly with the United Nations Population Fund.

This report contains population projections from the 2004 Timor-Leste census of Population and Housing. A notable feature is that because there is rapid population growth due to high fertility and lowering mortality, approximately half the population is 20 years of age or younger.

There is also more recent and detailed information in atlas form. The 2008 OCHA / UNMIT Integrated Humanitarian Coordination Team publication, *Timor-Leste District Atlas: Profiling Districts of Timor-Leste*, Version 2, was published by the UN Office for the Co-ordination of Humanitarian Affairs. This is a finely detailed atlas of Timor-Leste at district level including district profiles, which contains sufficient detail to designate resources (eg primary, pre-secondary, secondary, schools, and advanced education centres; airports, secondary airports and airstrips; hospitals, community health centres, and health posts; ports; roads; bridges; and elevation), plus tabulated information on land area; population by sex; average household size; languages spoken; flood rating; food security in 2005/6; and livelihoods. Information was from a variety of sources including the Timor-Leste National Statistics Directorate (DNE), WHO, UNPOL / PNTL through UNMIT, FAO, WFP, and OCHA. The 2004 census and the TL-SLS 2007 are specifically mentioned as major data sources.

There have also been a number of sample surveys conducted in Timor-Leste. Reports on the most important of these include:

-   The 2003 República Democrática de Timor-Leste, ADB, JICA, UNDP, UNICEF, UNMISET and the World Bank report, *Poverty in a New Nation: Analysis for Action*, which was the first published report on poverty and living standards, and was the outcome of three data collection activities: a Suco Survey, a Participatory Potential Assessment and a Living Standards Measurement Survey (TLSS 2001). The TLSS 2001 surveyed 1,800 households from 100 sucos, covering around one percent of the population. Fieldwork was completed over a period of three months, August to November 2001. While it was a major information source at the time, the TLSS 2001 survey is no longer sufficiently current for inclusion as a data source for small area estimation.
-   The 2006 Direcção Nacional de Estatística, Ministério de Finanças manual, *Projeto de avaliação da pobreza - Estudo sobre os padrões de vida de Timor-Leste: TL-SLS – 2006: Interviewer's manual*. This report is the detailed manual for the Timor-Leste Survey of Living Standards (TL-SLS) begun in 2006, suspended, and then reinitiated in 2007 (by carrying out the fieldwork essentially *de novo*). It covers background, survey methodology (including the intended survey period, team approach, questionnaire, visits, and integration of computer-based quality controls as an integral part of field operations), organization of the field work, role of interviewer (including designation, duties and responsibilities of the interviewers, and relationships to supervisors and data entry operators), general rules for interviewing (how to conduct an interview, ask questions, record responses, and general instructions), important concepts and definitions, and detailed instructions and clarifications of meaning for each question in the questionnaire. This survey manual and the 2004 census dictionary contain essential information for linking and matching of survey and census questionnaires that is necessary for small area estimation using the World Bank method.
-   The 2008 Direcção Nacional de Estatística, Ministério de Finanças report, *Final Statistical Abstract: Timor-Leste Survey of Living Standards 2007*, which contains a detailed summary of the survey design for the *Timor-Leste Survey of Living Standards 2007* (TL-SLS 2007) together with sets of tables relating to demographics; housing; access to facilities, durable goods; education; health; employment; social capital; and subjective well being. These tables provide information necessary for

sound implementation of the Timor-Leste government's national priorities which cover:

- Public Safety and Security
- Social Protection and Solidarity
- Addressing the Needs of Youth
- Employment and Income Generation
- Improving Social Service Delivery
- Clean and Effective Government

The TL-SLS 2007 itself is intended to provide the necessary survey-based material for the second National Development Plan, the Poverty Estimates, Profiles and Assessment and also the indicators to be used to measure and monitor the success of the plan. The survey is already being used for the MDG Poverty Reduction Report and will contribute to the update of the next Human Development Report (UNDP) and the Consumer Price Index (CPI). These planned reports, taken in conjunction, should improve the integration of efforts to reduce poverty. In the longer term planning and further integration would benefit from having sound small area estimates of poverty (ie poverty incidence, gap and severity) and undernourishment (ie kilocalorie consumption).

The topics covered in TL-SLS 2007 are very broad, and include much of what would be covered by surveys such as a Demographic and Health Survey, a Multiple Cluster Indicators Survey and a typical labour force survey. The TL-SLS was launched on 27 March 2006, and designed to run over a full calendar year to account for seasonal variation. However, after about two months of fieldwork, there was the April 2006 outbreak of conflict, and the survey had to be suspended. It resumed in January 2007. Survey operations ran continuously until conclusion in January 2008. Households which were interviewed in 2006 were revisited during December 2007-January 2008. For such households (around 300 in number), the final TL-SLS data contains only the information from the revisits. The TL-SLS sample by design has two components: a cross-sectional component of 4500 households representing the current population of Timor-Leste, and a panel component of 900 households, which contains half of the 2001 TLSS sample - randomly selected and re-interviewed. The panel is intended for evaluation of changes in living conditions between 2001 and 2007.

The cross-sectional survey provides estimates for rural and urban areas of each of five recently defined groups of districts or Regions:

- Region 1: Baucau, Lautem and Viqueque;
- Region 2: Ainaro, Manufahi and Manatuto;
- Region 3: Aileu, Dili and Ermera;
- Region 4: Bobonaro, Cova Lima and Liquiçá; and
- Region 5: Oecussi.

The cross-sectional sample was selected in two stages:

- In stage one, 300 Census Enumeration Areas (EAs) were selected as the primary sampling units (psus).
- In stage two, 15 households are selected in each EA.

The design has ten strata, which are the urban and rural areas in each of the five regions. The first stage of sampling used the 1,163 Census Enumeration Areas (EAs) that were generated from the 2004 Census which was used as a sample frame. Within each stratum, EAs were selected with probability proportional to size (pps) with-replacement, using the number of households reported by the census as a size measure. Smaller EAs were not appended to neighbouring EAs, nor were larger EAs subdivided to give more uniformly sized primary sampling units. At the second sampling stage, an exhaustive household listing operation was

undertaken in all selected EAs and this was used as the detailed household-level sample frame. Sample households in each EA were selected by systematic equal-probability sampling. The sampling fraction was comparatively large in some strata, which meant certain large EAs were selected more than once by the pps procedure that was used at the first sampling stage. This is the reason the cross-sectional sample consists of only 269 (rather than 300) different EAs. For multiply-selected EAs, a multiple of 15 (rather than 15) households were. Sucos from 2001 were redefined, and 60 of 442 the new sucos were classified as urban.

2.4   Pros and Cons of and Possibilities for Small Area Estimation of Poverty in Timor-Leste

For sound small area estimation using the World Bank method, integration of survey and census information is required. The preceding summary is intended to provide documentation of the background detail which is necessary for developing any local estimates of poverty at small area level. The actual level at which it will be possible to provide small area estimates in Timor-Leste is at present unknown, both because the TL-SLS 2007 data has not been released for preliminary assessment, and also because full integration with the census data for using the World Bank method, or using alternative small area methods based only on survey data (rather than fitting preliminary statistical models to survey data) fall beyond the scope of this study. Estimating poverty at small area level will be contingent on funding being made available, based on the recommendations in this feasibility assessment report.

One current hindrance to small area estimation by any method is that, while there was general recognition from the people and organisations consulted of the need for small area estimates of poverty in Timor-Leste for planning and resource allocation, there is also a perception that there is a more immediate need for improved communication between and integration of the activities of organisations (government, ministries, UN and NGOs) operating in Timor-Leste. As a consequence, even if small area estimation using the World Bank method were possible using the TL-SLS2007 and the 2004 census, it may have to wait for several years of further infrastructural development in Timor-Leste. Having available resources from within the Directorate of Statistics, given their personnel's local expertise and familiarity with questionnaires (including their languages and translation), form an essential and relatively time-consuming component to sound small area estimation using the World Bank method. For alternative small area estimation methods using survey data alone, the requirement for local expertise is not as pressing, since no matching of questions in survey and census is required. Subject to preliminary methodological research proving favourable, using such alternatives plus the TL-SLS 2007 data represents a logistically viable option even now.

A general comment about the relationship between small area estimation and mapping may also be warranted. Small area estimation of poverty, especially if extended from poverty incidence gap and severity, plus kilocalories, to stunting, underweight and wasting in children (as in Jones, Haslett and Parajuli, 2006) provides a detailed perspective on the spatial distribution of poverty. Other variables are also important however (eg health information, rainfall, and other Geographical Information System (GIS) data), even if these cannot be produced at such a fine level as variables collected in the census. For most users of this information, an atlas of maps is much more useful than a detailed technical report on small area estimation methodology, even if the technical report also contains finer level tabulated detail. Such a detailed, technical, small area estimation report, which would be the next stage after this feasibility assessment report should funding become available, is however essential, as it provides a clear indication of the methodological foundation for small area maps (often called poverty maps) that are included in the atlas. Without sound use of small area methodology, and publication of the technical report that outlines that methodology, the utility of the more generally-used atlas containing small area estimate maps must remain in doubt.

3 Data and other Requirements for Feasibility Assessment of Small Area Estimation of Poverty in Timor-Leste

3.1 Core Requirements

The World Bank method for small area estimation of poverty needs both survey and census data, which have been collected at approximately the same time. As a preliminary to statistical model fitting, a clear link via the questionnaires between what seems to be the same variables collected in census and survey needs to be established, and the actual average values or proportions for variables that prima facie match need to be checked to ensure that potential matched variables are not significantly different in census and survey.

The Timor-Leste Population and Housing Census (TLS) was conducted in 2004 and the Timor-Leste Survey of Living Standards (TL-SLS) in 2007. The time difference between these two data collections, given events and the number of internally displaced persons (IDP) in Timor-Leste between these two dates, represent a complication for small area estimation of poverty using the World Bank method, because the implicit assumption in this method is that the structure and location of the population, and the relationship between the matched variables in each data source, remains unchanged in the period between census and survey. Consequently, even prima facie, any use of the 2004 census and the 2007 TL-SLS for small area estimation of poverty using the World Bank method would be tentative at best, and would need to proceed with caution.

Neither census 2004 nor the 2007 living standards survey data were available from the Timor-Leste government during the period of this feasibility study (July 2008 to November 2008). Absence of data has compromised the current feasibility assessment for small area estimation of poverty in Timor-Leste. Were the required data to be come available, an extension of this assessment is warranted at least to the stage of fitting preliminary statistical models to the TL-SLS 2007.

Without survey data, it has not been possible to fit models (to the TL-SLS 2007) in the parallel manner to what has been done recently as part of small area estimation of poverty feasibility assessment for WFP in other countries (eg Cambodia and Bhutan), where government co-operation and support has been more tangible. See Haslett and Jones (2008) *Potential for Small Area Estimation and Poverty* at *Constituency and at Gewog / Town Level in Bhutan: Feasibility Report Phases 1 and 2,* and Haslett, and Jones (2008) *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia: Feasibility Report Phases 1, 2 and 3.* Having no TL-SLS 2007 data means there are no analyses in this report for Timor-Leste that provide a detailed statistical modelling assessment of poverty mapping feasibility.

Also, unlike Bhutan and Cambodia, there has only been very limited direct contact available with the Direcção Nacional de Estatística, Ministério de Finanças (DNE) during the two week (3 November – 14 November 2008) in-country assessment period for this report. It has consequently not been possible to discuss the tables (Appendix A, Tables A1-A15) that detail potential matching variables in census and survey with the central government statistical agency. The matching of variables is an aspect of the small area methodology using the World Bank method for which local expertise is indispensable, and detailed discussion was to have formed a substantial part of the current assessment. Essentially, it is only personnel in the organisation conducting survey and census that are able to answer the subtle questions about matching of questions in census and survey, and provide the necessary links to the

versions of questionnaires not in English actually used to conduct the data collections. Tables A1-A15 should consequently be viewed as provisional.

Nevertheless, Tables A1-A15 still provide a strong initial indication of connections between variables in the census and the living standards survey. Although they apply to the 2004 census and the 2007 survey, they will also be of use should small area estimation of poverty and poverty mapping at a fine level using the World Bank method be postponed until after the next census (planed for 2010) and the next living standards survey, because the questionnaire for the census is being planned now (ie in 2008) and (with the provisos detailed in the appendix) will be very similar to the 2004 census. There is also no current evidence that the 2007 TL-SLS questionnaire would be modified substantially. This variable-matching issue that is detailed in tables A1-A15 is not unique to Timor-Leste. Ideally, exactly the same question with exactly the same categories would be asked in the survey and census, but in practice internationally there are almost always differences.

The scope for the feasibility assessment did not include detailed variable-matching and statistical testing of the matching of survey and census data. This would have required both survey (TL-SLS 2007) and census (TLS 2004) data at unit-record (i.e. household and individual level) without names and addresses, which was not available.

A further aspect requiring substantive input from Direcção Nacional de Estatística (DNE) is the detailed geographical coding within Timor-Leste, and the changes to this coding over time. As far as we can assess without such local input, such coding changes appear to be relatively minor. This finding must however remain tentative, given the limited extent of the contact available with Direcção Nacional de Estatística during this feasibility study. What is required for small area estimation of poverty by the World Bank method is that changes to names, identification numbers and boundaries can be tracked in detail over time, so that the clustering used for the whole country at survey design stage can be linked almost exactly with the equivalent geographical units in the census. This requirement means that an integrated system linking geographical levels and the availability of this coding information to all members of the team undertaking any future small area estimation will need to be authorised before small area estimation of poverty by the World Bank method. Ideally this information will also be available at feasibility assessment phase when using 2010 census and the next TL-SLS.

3.2   Alternative Data Sources

For Timor-Leste, the main data sources considered were the Timor-Leste Population and Housing Census (TLS) which was conducted in 2004 and the Timor-Leste Survey of Living Standards (TL-SLS) conducted January 2007 to January 2008. Additional possibilities were the earlier Timor-Leste Living Standards Survey from 2001, which was not suitable because it is now seven years old, the Timor-Leste Comprehensive Food Security and Vulnerability Analysis (CFSVA) survey conducted December 2005 to January 2006, and the planned 2010 census.

The Timor-Leste Comprehensive Food Security and Vulnerability Analysis survey (CFSVA) was conducted by the World Food Programme between December 2005 and January 2006, preceding the April 2006 conflict. In principle, this removes the concern about the effects of internally displaced persons (IDPs) that was raised about the TL-SLS 2007 (point 6 above). However, the CFSVA is not suitable for small area estimation of poverty and kilocalorie consumption two reasons. Firstly, it cannot be used for poverty incidence, gap and severity since it only collects food expenditure. Secondly, it collects only limited information on food intake and expenditure information on certain foods, but no information on quantities consumed. Even if it were possible in principle, conversion of expenditure to kilocalories

would introduce an additional complexity and unknowable measurement errors. Indeed WFP's CFSVA report comments explicitly that kilocalorie information is ' not available'. See Fedele and Horjus (2006) *Timor-Leste: Comprehensive Food Security and Vulnerability Analysis (CFSVA)*, p15.

One further option is to use the TL-SLS 2007 in conjunction with the planned 2010 census. The option has some attractions, provided that there is relative stability in Timor-Leste over the period between the two data collections, especially since it would in principle allow use of the World Bank method for small area estimation. Caution is however warranted. There would still be three years between survey and census, which as noted earlier in relation to the 2004 census and the TL-SLS 2007 is rather long. There would also be major and time-consuming geographical coding issues to resolve, as the TL-SLS 2007 has used enumeration areas (EAs) that are essentially those from the 2004 census, but all EAs (which are below suco level and consist of about 100 households) will need to be redefined for the 2010 census. Every EA (or part EA or household) in Timor-Leste in 2004 will need the geographical information to be derived that links it to the 2010 EA definitions before small area estimation using TL-SLS 2007 and the 2010 census would be feasible.

4 Core Methodology

The first phase of the feasibility study has examined data sources and identified possible variables useful for small area estimation of poverty and undernourishment. These are tabulated in Tables A1-A15. For model fitting to survey data, these variables comprise:

- a target variable, here denoted y, recorded for each household in the survey and from which household-level poverty or undernourishment status can be inferred; this variable is log-transformed household per capita consumption expenditure (for poverty) or household per adult equivalent daily calorie intake (for undernourishment);

- a set of auxiliary variables, denoted x_1, x_2, ..., x_p, either recorded at household level in both survey and census or else missing from the survey but available from the census or another external source as an average over a geographic area; for example household size (`hhsize`) is a household-level variable available in both the census TLS 2004 and the TL-SLS 2007; posto- or suco-level averages have not yet been included because neither the census nor the survey data were made available;

- a set of regional indicators common to the survey, census and other external data sources; these are required so that area-level information can be merged appropriately with the survey and census databases, and so that small area estimates can be produced at appropriate geographic levels; for Timor-Leste these are indicators for region, district, sub-district (posto), suco and enumeration area (EA)..

The ELL method uses the auxiliary data to infer the value of the target variable y, and hence poverty, malnutrition or undernourishment status, for every household in the census via a statistical model

$$y_{vh} = \beta_0 + \beta_1 x_{1vh} + \beta_2 x_{2vh} + \cdots + \beta_p x_{pvh} + \eta_v + \varepsilon_{vh} \qquad (1)$$

where y_{vh} denotes the value of the target variable in the hth household of the vth cluster. Clusters are defined as the primary sampling units (psus) from the survey design – here the enumeration areas (EAs). The unknown parameters $\beta_0, \beta_1, \cdots, \beta_p$ are estimated using the survey data. The disturbance terms η_v, ε_{vh} represent unexplained variation at primary sampling unit (psu) and household level respectively; these are treated in the modelling as random effects with mean zero and variances $\sigma_\eta^2, \sigma_\varepsilon^2$. These variance components are also estimated from the survey data, and can be extended to include variance components at small area level too\.

The quality of the small area estimates produced by the ELL method, measured in terms of the standard errors of the estimates, depends crucially on the quality of the model linking the target and auxiliary variables. We now briefly examine this issue, indicating what would be required for a proper evaluation.

5   Measures of Model Performance

There are two aspects of model performance commonly used to evaluate linear models like (1). The first measures the proportion of the variability in the target variable (y) explained by the predictors (x); this is commonly denoted R^2. The second, the mean squared error (MSE), measures the overall size of the unexplained variation. In the ELL method however there is no direct link between these and the precision of the final small area estimates; this is partly because the variable of interest at small area level is not y itself but a nonlinear function of y (for example *poverty incidence* is the proportion of households in an area for which y is below a poverty line). Nevertheless, previous experience with the ELL method, combined with theoretical considerations, enable some statements linking the two to be made.

In successful applications of the ELL method to poverty estimation, the R^2 value of the model for log-transformed per capita expenditure tends to be about 50% or higher. R^2 always increases as more x variables are added to the model, but that a point of "diminishing returns" sets in after which an increase in model complexity gives only a negligible improvement in R^2. Furthermore such apparent improvements in the predictive power of the model may be spurious, holding for the estimation data but not for future predictions. The target, of around 50% for consumption poverty, should be achieved with a modest number p of auxiliary variables in relation to the size of the estimation dataset. Calorie intake tends to be harder to predict than consumption expenditure (perhaps, in part, because of measurement errors); there is less experience here, but (because of a different balance between the types of unexplained variation as outlined below) reasonably successful small area estimates have been produced using models with an R^2 of around 30%.

In (1) the unexplained variation is decomposed into psu-level (η_v) and household-level (ε_{vh}) effects. When the model is used to predict y for the census households, simulated values of these random effects are included to incorporate the uncertainty in these predictions. When the household-level predictions are amalgamated to small area (posto) level to produce the final estimates, these effects will tend to "average out", reducing the variability at small area level; the extent to which this happens depends on the number of census households, and the number of census psus, in the small areas. Since the number of households is generally much larger than the number of psus, the crucial factor affecting precision is usually the size of the psu-level effects. Thus while both σ_ε^2 and σ_η^2 should preferably be small, it is particularly important that the latter should be as small as possible.

Table 1. Summary measures for expenditure models from recent SAE exercises.

	p	R^2	σ_η^2	σ_ε^2
Cambodia2002 PPenh	38	0.48	0.099	0.359
Cambodia2002 urban	40	0.70	0.034	0.354
Cambodia2002 rural	58	0.54	0.141	0.236
Bangladesh2003	31	0.59	0.025	0.107
Philippines2004	68	0.74	0.039	0.113
Nepal2005	37	0.55	0.027	0.194

Table 1 gives the values of these summary measures for the models used in some recent sae poverty-mapping exercises. Variance decomposition of carefully fitted regression models using the TL-SLS would be required in order to assess the situation for Timor-Leste.

6   Structure of Survey and Census – Further Detail

Another important factor governing the achievement of sufficient precision in small area estimates of poverty by the ELL method relates to the number of psus in each small area. In our experience, the average number of psus in the small area for which poverty estimates are produced is typically 20-30. Any areas where the number of psus falls significantly below this number are likely to be too small to get reliable estimates, unless the regression modelling can reduce psu-level variation below what is usually achieved. The limited data sources currently available to us (in particular no available TL-SLS 2007 data) do not enable us to examine this issue properly for Timor-Leste. Table 2 summarizes the currently available information on the structure of the TLS 2004 census and TL-SLS 2007 survey.

Table 2. Structure of TLS2004 and TS-SLS2007

	Region	District	Posto	Suco	EA	households
Census	5	13	68	442	1163	194962
Survey	5	13	-	-	269	4477

We can see that the average number of psus (or enumeration areas, EA) in a sub-district (or posto) is approximately 17. However, we also need to be able to look at the overall distribution of EAs per posto, and it is likely that there will be at least some postos with small numbers of EAs. Without TL-SLS 2007 data, we are also not able to judge how many of the 68 postos were sampled in the survey. Though the absence of survey data in some small areas does not preclude their estimation by the ELL method, it would rule out estimation by other, survey-only, small area estimation methods unless the auxiliary variables for the unsampled postos were available from another source.

7   Conclusions

No survey or census data were available for this feasibility assessment.

Even without unit record data however, it is clear that full small area estimation using the World Bank method, based on the Timor-Leste Survey of Living Standards (TL-SLS 2007) and the Timor-Leste Population and Housing Census (TLS) 2004 is unlikely to be warranted, because the two data sources do not match particularly well. The differences are due to an unusual combination of circumstances including:
- internal migration between 2004 and 2007, for example internally displaced persons from the conflict in Timor-Leste which began in April 2006,
- differences in structure between the detailed questions asked in the census and survey questionnaires.

No other suitable data sources are currently available.

Using the World Bank method with the TL-SLS2 2007 and the planned 2010 census may be another option. However, as for the 2004 census and TL-SLS 2007 there would still be three years between survey and census. In addition as a prerequisite there will be major and time-consuming geographical coding issues to be resolved at enumeration area (which consists of about 100 houses) for the whole of Timor-Leste.

If small area estimates are required before the 2010 census and the follow up TL-SLS are completed however, one available option is to consider using alternative small area modelling based only on the Timor-Leste Survey of Living Standards (TL-SLS 2007). Given this data were to become available, this set of methods are viable because such models do not use or rely on the census data or matching of the variables in its questionnaire with that of the survey. Such alternatives may well be able to give poverty incidence and kilocalorie estimates for 2007 at a finer geographical level (via statistical modelling) than is possible using traditional design-based sample survey estimation methods. The extent to which such alternatives would provide improved poverty estimates when compared with the direct survey estimates is however likely to be more limited than that claimed for the World Bank method. There are a number of aspects to this difference, but the most obvious is that with only one rather than two data sources, improvements to accuracy (and hence how small the area is at which small area estimates are possible) are more limited. However, the advantage of these alternative small area methods, since they rely only on the survey data, is that they would not be compromised by the major changes in Timor-Leste between the 2004 census and the TL-SLS in 2007. If small area estimates are required in 2009, the TL-SLS 2007 data will first need to be released by the Timor-Leste government, after which funding should be sought from a variety of international donors for such a small area estimation study in Timor-Leste based only on the TL-SLS 2007 data. Because no relevant analysis of the TL-SLS 2007 data has been carried out to date, it is recommended any such project be divided into two phases, the first of which would extend the current study by developing and assessing the theory behind the methodology and would also fit the provisional statistical models that were intended to be part of the current feasibility study.

The Timor-Leste census is scheduled to be repeated in 2010. If this census were to use sufficiently similar core questions to TL-SLS, and the TL-SLS were also to be repeated around the time of the census, in principle this would provide a strong basis for small area estimation of poverty (poverty incidence, gap and severity, plus kilocalorie consumption) for Timor-Leste post-2010, using the World Bank and/or a range of other small area estimation

methods. This new Population and Housing Census (TLS) data, new TL-SLS data, and the integrated geographical information would be required. If the 2004 questionnaire census were to be used for the 2010 census with the addition of information on a number of variables commonly used in poverty mapping: Number of Rooms, Fuel for Cooking, Lighting Source, Drinking Water, Toilet Type, Radio Ownership, Telephone Ownership, and the TL-SLS 2007 questionnaire were to be essentially reused, this may suffice, although (where they are candidate variables for small area modelling) asking the same questions with the same categories in both the new census and any new survey is strongly recommended.

Completion of this report follows consultation with the Timor-Leste Direcção Nacional de Estatística, other officials and members of the government at the Ministério de Finanças, UNFPA, the World Bank, and the World Food Programme (WFP) - which commissioned this research - in Timor-Leste, 3 - 14 November 2008. The authors are grateful for these contributions. Viewpoints expressed in this report do not however necessarily reflect those of all or any of the organisations consulted.

Bibliography

Alderman H., Babita M., Demombynes G., Makhata N. and Ozler B. (2002) "How low can you go? Combining census and survey data for mapping poverty in South Africa", *Journal of African Economics*, **11**, 169-200.

Asian Development Bank (2005) *ADB Country Strategy and Program Update: Democratic Republic of Timor-Leste (2006-2008)*, Asian Development Bank, September 2005, 39 pages.

BELUN, FONGTIL and National Dialogue Team of the Office of the President of the Republic. (2006) *National Dialogue* (*Open Space*) 7-8 April 2006, Presidencia de República, 102 pages.

Direcção Nacional de Estatística, Ministério de Finanças (2004) *Timor-Leste 2004 Census Project: Census Dictionary*, Direcção Nacional de Estatística, 42 pages.

Direcção Nacional de Estatística, Ministério de Finanças (2006) *Projeto de avaliação da pobreza - Estudo sobre os padrões de vida de Timor-Leste: TL-SLS – 2006: Interviewer's manual*, Direcção Nacional de Estatística, 84 pages.

Direcção Nacional de Estatística, Ministério de Finanças (2008) *Final Statistical Abstract: Timor-Leste Survey of Living Standards 2007*, July 2008, Direcção Nacional de Estatística, Ministério de Finanças, Timor-Leste, x + 246 pages.

Elbers, C., Lanjouw, J.O. and Lanjouw, P. (2001) *Welfare in villages and towns: micro-level estimation of poverty and inequality*, unpublished manuscript, The World Bank.

Elbers, C., Lanjouw, J. and Lanjouw, P. (2003) "Micro-level estimation of poverty and inequality", *Econometrica*, **71**, 355-364.

Elbers, C., Lanjouw, J.O., Lanjouw, P. and Leite, P.G. (2001) *Poverty and inequality in Brazil: new estimates from combined PPV-PNAD data*, unpublished manuscript, The World Bank.

Fedele, S. and Horjus, P. (2006) *Timor-Leste: Comprehensive Food Security and Vulnerability Analysis (CFSVA): Conducted in December 2005-January 2006*, Vulnerability and Mapping Unit, World Food Programme, July 2006, 87 pages.

Forsén, Y., Sanogo, I., Ayanahati, L., Rasenen, A-L., Dasilva, F. (2007) *Dili Emergency Food security Assessment, Timor-Leste, September 2007*, World Food Programme, 55 pages.

Fujii, T. (2002) *Estimation of Poverty Rates at Commune-Level in Cambodia – using Small-Area Estimation Technique to Obtain Reliable Estimates*, Ministry of Planning and the United Nations World Food Programme, October 2002, vi+62 pages+CD of estimates.

Haslett, S. and Jones, G. (2008) *Potential for Small Area Estimation and Poverty at Constituency and at Gewog / Town Level in Bhutan: Feasibility Report Phases 1 and 2,* Report to the World Food Programme, 11 November 2008

Haslett, S. and Jones, G. (2008) *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia: Feasibility Report Phases 1, 2 and 3,* Report to the World Food Programme, 20 October 2008

Haslett, S. and Jones, G. (2005) *Estimation of Local Poverty in the Philippines*, Philippines National Statistics Co-ordination Board / World Bank, November 2005.

Healy A.J., Jitsuchon S. and Vajaragupta, Y. (2003) "Spatially disaggregated estimation of poverty and inequality in Thailand", preprint.

Jones, G. and Haslett, S. (2003) *Local Estimation of Poverty and Malnutrition in Bangladesh*, Bangladesh Bureau of Statistics and United Nations World Food Programme.

Jones, G., Haslett, S. and Parajuli, D. (2006) *Local Estimation of Poverty, Undernourishment and Malnutrition in Nepal*, The Central Bureau of Statistics of His Majesty's Government of Nepal / United Nations World Food Programme.

National Statistics Directorate / United Nations Population Fund (2006) *Timor-Leste Census of Population and Housing 2004: Atlas,* National Statistics Directorate, September 2006, 89 pages.

National Statistics Directorate (2006) *National Priority Tables: Census of Population and Housing 2004*, National Statistics Directorate / United Nations Population Fund.

National Statistics Directorate (2006) *Population Projections 2004-2050: Analysis of Census Results, Report 1: General, Census of Population Census of Timor-Leste 2004*, National Statistics Directorate / United Nations Population Fund, 78 pages.

National Statistics Office Timor-Leste (2003) *2004 Census of Population and Housing: Project Plan 2004 Census Project,* Document of the Census Technical Committee, 22 October 2003, 11 pages.

OCHA / UNMIT Integrated Humanitarian Coordination Team (2008) *Timor-Leste District Atlas: Profiling Districts of Timor-Leste*, Version 2, UN Office for the Co-ordination of Humanitarian Affairs, August 2008, iv + 31 pages.

Rao, J. N. K. (2003) *Small Area Estimation*, Wiley, New York.

Rao, J. N. K. (1999). Some recent advances in model-based small area estimation, *Survey Methodology* 25, 175-186.

República Democrática de Timor-Leste and United Nations Country Team (2004) *Timor-Leste Millennium Development Goals Report*, February 2004, 53 pages.

República Democrática de Timor-Leste, ADB, JICA, UNDP, UNICEF, UNMISET and the World Bank (2003) *Poverty in a New Nation: Analysis for Action*, May 2003, xvi + 138 pages.

República Democrática de Timor-Leste: Ministry of Agriculture, Forestry and Fisheries (2006) *National Food Security Policy for Timor-Leste*, iii + 33 pages.

República Democrática de Timor-Leste, Ministry of Education and Culture (2006) *Report of the First Annual Joint Review for the Education Sector*, 18-19 October 2006, Dili, Ministry of Education and Culture, 71 pages.

Seeds of Life / Fini ba Moris (2006) *Annual Research Report 2006*, Ministry of Agriculture, Forestry and Fisheries, República Democrática de Timor-Leste, x + 58 pages.

You, Y. and Rao, J. N. K. (2002). A pseudo-empirical best linear unbiased prediction approach to small area estimation using survey weights, *Survey Methodology* 30, 431-439.

You, Y., Rao, J. N. K. and Kovacevic, M. (2003). Estimating fixed effects and variance components in a random intercept model using survey data, *Statistics Canada International Symposium Series - Proceedings*.

Zhao, Q. (2006) *User Manual for PovMap*, The World Bank.

http://siteresources.worldbank.org/INTPGI/Resources/342674-1092157888460/Zhao_ManualPovMap.pdf

Appendix A

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

General Notes for Tables A1-A15:

1. Tables A1-A15 provide the detail necessary for preliminary matching of variables in the Census of Population and Housing Timor-Leste (TLS) 2004 and the Timor-Leste Survey of Living Standards (TL-SLS) 2007.

2. For the Census we have only considered the Household Questionnaire. Separate questionnaires were used for hotels and institutional residences such as hospitals, prisons etc.

3. The Census records information on household members present on census night (or are away for a period of less than 48 hours), although some basic information is available on usual members who are temporarily absent for longer periods (Part 6: Usual residents absent on census day). The survey in contrast collects information for all usual household members.

4. Political upheavals have reportedly resulted in the displacement of significant numbers of people, so that some households present in 2004 no longer existed in 2007. This displacement may affect the matching of census and survey information, particularly in those districts where a significant amount of displacement occurred.

5. Questionnaire details have been compared in these tables via their English versions. The 2004 census was also printed and administered in Portuguese, Tetun, and Bahasa Indonesia; the 2007 TS-SLS survey was administered in Tetun. For final matching decisions (between survey and census) local knowledge of the questions actually asked is essential. The tables below list agreement or otherwise in principle. Even where questions are identical in English, they may not be in Tetun. Further statistical checking that

similar proportions of people in survey and census respond to each apparently equivalent category will be required after the small area estimation feasibility assessment, and before fitting the final small area models to be applied to census data.

6. The "Other (specify)" coding used in some census and survey questions will need to be clarified to ascertain if/where this code has been used to create new codes using specific categories.

7. It is highly recommended that in future, if further small area estimation using a combination of both data sources is desired, survey and census questions should be standardized to use the same wording and categories.

8. The majority of the TL-SLS 2007 survey questions are not collected in the census. With the exception of expenditure related information, such non-matching questions are irrelevant to small area estimation using variants of the World Bank methodology for small area poverty estimation, which model expenditure using survey data and use variables that match between survey and census for prediction of expenditure at household level, and which are then combined to small area level.

9. There are migration/displacement related questions in both 2004 census (Household Questionnaire Part 7, Individual particulars, place of residence q1-6) and TL-SLS 2007 (Section1: Household Information, part A: Household roster q11-16) but the questions are different and code matching would be difficult.

10. There are a few variables: Religion, Mental and Physical Disability: that are recorded in the 2004 census but not in the TL-SLS 2007. These variables, and any others for which the matching is felt to be inadequate, could potentially be averaged at EA or suco level and merged with the existing data to enhance the modelling.

11. The 2004 census did not collect information on a number of variables commonly used in poverty mapping: Number of Rooms, Fuel for Cooking, Lighting Source, Drinking Water, Toilet Type, Radio Ownership, Telephone Ownership. This information is expected to be available from the next census in 2010.

12. Tables A1-A15 below are organised in the same order as the questions appear in the 2004 census questionnaire.

Table A1:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

House occupation status – Household level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*

Ownership category	Household questionnaire part 3: Dwelling and household information, q1.	1 Individual and/or family owned property
		2 Community or Suco owned property
		3 Government owned property
		4 Church property
		8 Other
			Section 2: Housing, part A: Description of the dwelling, q1, q5, q6.	Own the dwelling:
				  1 yes
				  2 no
				Ownership status:
				  1 Lease/Rent
				  2 Official House
				  3 Rent Free
				  4 Other (Specify___)
				Rent/lease from whom:
				  1 Relative
				  2 Private Employer
				  3 State
				  4 Private Person
				  5 Other (___)
					Rental categories difficult to match. Can derive a variable for owner-occupied or not. Ownership status may be uncertain because of the current status of property legislation, and because of the abandonment of properties due to civil unrest.

Table A2:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Wall material – Household level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Wall material	Household questionnaire part 3: Dwelling and household information, q2.	1 Brick 2 Concrete 3 Unbaked Brick 4 Wood 5 Bamboo 6 Rattan 7 Tin 8 Mud 9 Other (Specify___)	Section 2: Housing, part A: Description of the dwelling, q1.	1 Concrete/brick 2 Wood 3 Bamboo 4 Corrugated iron 5 Clay 6 Other	Some but not all categories are identical; amalgamation of categories likely to be required to match census and survey information

Table A3:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Roof material – Household level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Roof material	Household questionnaire part 3: Dwelling and household information, q3.	1 Concrete 2 Wood 3 Bamboo/thatch/grass 4 Corrugated iron 5 Tiles 6 Asbestos 8 Other	Section 2: Housing, part A: Description of the dwelling, q2.	1 Concrete 2 Wood 3 Metal Sheets/Zinc 4 Tile 5 Sugar Palm Fibre 6 Leaves 7 Other (Specify___)	Some but not all categories are identical; amalgamation of categories likely to be required to match census and survey information

Table A4:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Floor material – Household level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Floor material	Household questionnaire part 3: Dwelling and household information, q4.	1 Concrete or Tile 2 Wood 3 Soil 8 Other	Section 2: Housing, part A: Description of the dwelling, q3.	1 Marble/Ceramic 2 Floor Tile/Cement 3 Concrete/Brick 4 Wood 5 Bamboo 6 Earth/Clay 7 Other (Specify___)	Some but not all categories are identical; amalgamation of categories likely to be required to match census and survey information

Table A5:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Livestock ownership – Household level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Ownership of livestock	Household questionnaire part 3: Dwelling and household information, q5.	Number of: Chickens Pigs Sheep Goats Horses Cattle Buffalo	Section 9: Farming, Livestock, Forestry and Fisheries, part G: Livestock, q3.	Number of: 8001 Buffalo 8002 Bali cow 8003 Cow 8004 Horse 8005 Pig 8006 Goat 8007 Sheep 8008 Chicken 8009 Duck 8011 Other (Specify___)	Use to create separate variables for ownership of matching animals, either numerical, binary or categorical. Needs exploration of the data to decide on the best strategy. A possible complication is the shared ownership of livestock.

Table A6:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Crops grown – Household level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*

Crops grown	Household questionnaire part 3: Dwelling and household information, q6.	Tick all that apply:
		01 Rice
		02 Maize
		03 Cassava
		04 Vegetables
		05 Fruit (temporary)
		06 Fruit (permanent)
		07 Coffee
		08 Coconut
		09 Other temporary crops
		10 Other permanent crops
			Section 9: Farming, Livestock, Forestry and Fisheries, part B: Crops harvested, q1.	No/Yes:
				01 "Gogo" rice
				02 Rice
				03 Maize
				04 Cassava
				05 Coffee (Cherries)
				06 Coffee (Dry beans)
				07 Kidney beans
				08 Sweet potato
				09 Potato
				10 Taro (Talas/Kontas)
				11 Squash
				12 Mung bean
				13 Soy bean
				14 Coconut
				15 Peanuts
				16 Vegetables
				17 Bananas
				18 Other fruit
					Some collapsing of categories may achieve compatible variables such as:
					   Rice
					   Maize
					   Cassava
					   Vegetables
					   Fruit
					   Coffee
					   Coconut
					
					May be of limited use for modelling as amounts cultivated are not recorded.

Table A7:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Household size, Head of household - Household level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Household size	Household questionnaire part 4: People who usually live in this house, total of respondents. Also consider part 6: Usual residents absent on census day.	Numeric: 1, 2, 3,…..	Section1: Household Information, part A: Household roster, total of respondents	Numeric: 1, 2, 3,…..	Need to watch during modelling for household sizes in census that are very much in excess of those in the survey. Also check matching of eligibility criteria for household membership, particularly treatment of absentees.
Household head	Household questionnaire part 4: People who usually live in this house, q6	"Relationship to head of household" is coded, with "01" for "Head".	Section1: Household Information, part A: Household roster, q2.	Code '01' denotes head of household	This variable is useful for deriving personal characteristics of the head as household-level variables. Need to check for possible inconsistencies in the treatment of an absent head of household.

Table A8:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Sex, Age – Individual level

Topic	TLS 2004 Question	TLS 2004 Categories	TL-SLS 2007 Question	TL-SLS 2007 Categories	Notes
Sex	Household questionnaire part 4: People who usually live in this house, q5.	1. Male 2. Female	Section1: Household Information, part A: Household roster, q1.	1. Male 2. Female	Combine with age, head of household to give proportion of adult females, female head.
Age	Household questionnaire part 4: People who usually live in this house, q3.	Numerical values: Age at last birthday (years).	Section1: Household Information, part A: Household roster, q5.	Numerical values: Age in completed years. For children < 12 years, age in years and months is given.	Ignore "month" information. Age is used to derive age-related variables such as proportion of children under 7, proportion aged 7 to 14.

Table A9:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Marital Status – Individual level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Marital Status	Household questionnaire part 4: People who usually live in this house, q7.	1. Never married 2. Married 3. Widowed 4. Divorced 5. Separated	Section1: Household Information, part A: Household roster, q7.	1. Married 2. Divorced 3. Separated 4. Widow or widower 5. Never married	Categories appear identical, except for permutation of order. Combine with head of household to give marital status of head. "Married" includes de facto union. There may be difficulties in the definition of "Separated", and in the treatment of cases where a spouse is missing because of displacement.

Table A10:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Spoken Language – Individual level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*

Mother Tongue	Household questionnaire part 7: Individual particulars, q8.	01 Tetun
		02 Baequeno
		03 Bunak
		04 Fatalucu
		    ....
		20 Waweloi
		21 Bahasa Indonesia
		22 Portuguese
		23 Inggris
		24 Other
			Section1: Household Information, part A: Household roster, q9.	01 Portuguese
				02 Tetun
				03 Indonesian
				04 English
				    ....
				35 Tokodede
				36 Waima'a
				50 Malay
				60 Chinese
				70 Other
					Too many categories to use all. Might be useable if some categories can be collapsed. Perhaps use at head-of-household level. Could be problems where a head of household grew up with one language but has adopted another within their current household

Languages	Household questionnaire part 7: Individual particulars, q9.	0. Do not speak, read or write
		1. Speak only
		2. Read only
		3. Speak and read only
		4. Speak, read and write in:
		    Portuguese;
		    Tetun;
		    Bahasa;
		    English.
			Section1: Household Information, part A: Household roster, q10.	Tetun
				Indonesian
				Portuguese
				English
				
				    1. Yes
				    2. No
					For each language, answers 1, 3 or 4 in TLS 2004 should correspond to 1 in TL-SLS 2007. Use at head-of-household level.

Table A11:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Literacy – Individual level

Topic	TLS 2004 Question	TLS 2004 Categories	TL-SLS 2007 Question	TL-SLS 2007 Categories	Notes
Literacy - reading	Household questionnaire part 7: Individual particulars, q9.	0. Do not speak, read or write 1. Speak only 2. Read only 3. Speak and read only 4. Speak, read and write in: Portuguese; Tetun; Bahasa; English.	Section 5: Education, part A: General education, q3.	Can read a letter: 1. Yes, without difficulty 2. Yes, with difficulty 3. No	Need to combine survey answers to get overall literacy. Then assume that literacy is achieved only in the four languages given. Use to derive household level variables: literacy of head, proportion literate (aged 6+). Unclear how survey with/without difficulty corresponds to census responses.
Literacy - writing	Household questionnaire part 7: Individual particulars, q9.	0. Do not speak, read or write 1. Speak only 2. Read only 3. Speak and read only 4. Speak, read and write in: Portuguese; Tetun; Bahasa; English.	Section 5: Education, part A: General education, q4.	Can write a letter: 1. Yes, without difficulty 2. Yes, with difficulty 3. No	Need to combine survey answers to get overall literacy. Then assume that literacy is achieved only in the four languages given. Use to derive household level variables: literacy of head, proportion literate (aged 6+)

Table A12:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Education – Individual level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*

Schooling	Household questionnaire part 7: Individual particulars, q10.	Number of years at primary school.	Section 5: Education, part A: General education, q5.	Attended school
				  1. Yes
				  2. No
					Use to derive whether head of household has no schooling.

Education level	Household questionnaire part 7: Individual particulars, q10-q14.	Number of years of:
		  Primary school
		  High school
		
		Received diploma for high school
		
		Number of years of:
		  Tertiary education
		
		Received university degree or certificate
			Section 5: Education, part A: General education, q5.	Highest grade completed:
				
				Level Class
				Pre-School 1 1
				Primary 2 1-6
				Pre-Secondary 3 1-3
				Secondary 4 1-3
				Academy 5 1-3
				University 6 1-7
				Vocational 7 1-6
				Non-Formal 8
					Care will need to be taken to match the census and survey response categories. The education system has undergone significant transformation in the past 30-40 years, and grade level criteria may not be consistent across time. Perhaps create broad categories for head of household, eg
					• completed primary;
					• completed secondary

Table A13:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Employment – Individual level

Topic	TLS 2004 Question	TLS 2004 Categories	TL-SLS 2007 Question	TL-SLS 2007 Categories	Notes
Worked in the past week	Household questionnaire part 8: Employment information, q2.	Number of days worked in past week (for ages 15+)	Section 8: Employment, part C: Unemployment, q1.	Number of days worked in past week (for ages 10+)	Non-zero values indicate employment – but what is included as employment? Is it seasonal? Need to check for statistical match. Could use at head-of-household level, or proportion of adults (15+) in employment.
Employment type	Household questionnaire part 8: Employment information, q1.	1 Government (incl Police, Military, Teacher). 2 UN Organisation. 3 Worked in NGO (paid or voluntary). 4 Private Industry (paid or voluntary). 5 Self employment working in own business. 6 Subsistence farming and fishing	Section 8: Employment, part B: Wage Employment, q4.	1 A rural public works program 2 A state-owned enterprise 3 A private individual 4 A private company, enterprise or cooperative 5 The government, public sector or army 6 An NGO 7 Other (Specify:_____)	Some categories appear to match, but need to check statistical comparisons. Could use at head-of-household level.

Table A14:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Occupation – Individual level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Occupation	Household questionnaire part 8: Employment information, q3.	01 Armed Forces 10 Legislators, Senior Officials, Managers including Suco and Aldeia Chiefs 22 Health Care Professionals 23 High School Teachers, University Lectures .... comprising 37 categories	Section 8: Employment, part A: Jobs during the past 12 months, q1.	01 Professional and Technical Experts 02 Managerial, Administrative and Decision-Making Staff 03 Clerical 04 Sales Workers 05 Service Workers .... comprising 13 categories	Too many categories with poor matching between census and survey. Could use at head-of-household level if broader categories can be found that agree in statistical matching. Census records for main occupation only; further work may be necessary to identify the main occupation from the survey.
Industry	Household questionnaire part 8: Employment information, q4.	05 Agriculture 10 Forestry 15 Fishing 20 Mining, Quarrying, Oil 25 Manufacturing .... comprising 21 categories	Section 8: Employment, part B: Wage Employment, q2.	01 Agriculture, Livestock 02 Forestry 03 Fishing 04 Hunting 05 Mining And Quarry .... comprising 16 categories	Some categories appear to match, but need to check statistical comparisons. Could use at head-of-household level.

Table A15:

Links between Census of Population and Housing Timor-Leste (TLS) 2004 and Timor-Leste Survey of Living Standards (TL-SLS) 2007

Fertility Information – Individual level

Topic	*TLS 2004 Question*	*TLS 2004 Categories*	*TL-SLS 2007 Question*	*TL-SLS 2007 Categories*	*Notes*
Child mortality	Household questionnaire part 9: Fertility information, q8-q9.	Number of offspring born alive but subsequently died; Total number of offspring born alive.	Section 7: Fertility, q6-q7.	Number of offspring born alive but subsequently died; Total number of offspring born alive.	Not necessarily "child" mortality as the offspring could have died in adulthood. Census includes women aged 15+; survey is for women aged 10-49. Could use the age variable to restrict to, say, women aged 15-30.