Potential for Small Area Estimation of Malnutrition at District and Commune Level in Cambodia

Feasibility Report

Phases 1 and 2

Professor Stephen Haslett
Dr Geoffrey Jones
Dr Marissa Isidro

Massey University
New Zealand

19 November 2010

Summary

This report extends an earlier 2008 feasibility report, *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia* (which focused on poverty incidence gap and severity) to include stunting, underweight and wasting in children under five years of age. Since the publication of the 2008 report, additional unit-record data has become available from the Cambodia Socio-Economic Survey (CSES2009), and the Cambodia Anthropometric Survey (CAS2008), and these two surveys form the basis for the current study. Note that unit-record data from the General Population Census of Cambodia (GPCC2008), is not yet available and will be required to extend results in these feasibility reports to production of any poverty related estimates via small area estimation and to produce maps.

The analyses undertaken both here and in the 2008 feasibility report fit statistical models to survey data, and assess feasibility of small area estimation and mapping only from diagnostics from these models, without using census data. Neither of the two feasibility studies, as for those from other countries, is itself able to produce any small area estimates, because to do so requires use of unit-record census data which is not yet available from the 2008 Cambodian census. As a supplement to the analyses for CAS2008 and CSES2009 detailed here, however, useful comparison of survey model diagnostics has been made with countries where small area estimation and mapping has been completed in order to support the conclusions drawn on small area estimation feasibility.

The general conclusion of this report is that, for Cambodia, updated small area estimation of stunting and underweight, but not of wasting, seems to be feasible at district or commune level. In practice this means that some smaller communes may need to be combined with neighbours. This conclusion supplements that in the earlier feasibility report that small area estimation of poverty incidence, gap, and severity is feasible at a similar level of disaggregation.

Executive Summary

1. Small area estimation 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. 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.

2. This report provides a brief summary of the two phases of a 2010 feasibility study, undertaken by staff from Massey University, New Zealand in conjunction with the Cambodia National Institute of Statistics and the World Food Programme, to assess feasibility for small area estimation of stunting, underweight and wasting in children under five years of age at district or commune level. This report includes and hence supersedes the earlier 2010 Phase 1 report, and should be read in conjunction with the 2008 feasibility report, *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia*, which focused instead on poverty incidence, gap, and severity.

3. For Cambodia, the data sources considered for use in small area estimation are the population census (conducted in 2008), the Cambodia Socio-economic Survey (CSES) which was most recently conducted in 2009, and the Cambodian Anthropometric Survey 2008. Use of the Seila commune database, more recently known as the National Committee for Decentralisation and Deconcentration (NCDD) database, is also discussed. For the preliminary statistical model testing however, the feasibility report focuses only on the unit record data for CSES2009 and CAS2008, and the questionnaires for CSES2009 and CAS2008 and the 2008 census.

4. Some additional research may be needed later re-checking area boundaries at district, commune and possibly village level when using CAS2008, CSES2009 and the 2008 census data for small area estimates. The boundaries have been revised with the 2008 census but not retrospectively. Tight control and recording of all such boundary changes, and care in coding region, district, commune and possibly village for every household in the census is essential. This caution applies even though the 2008 census boundaries were used for both CAS2008 and CSES2009, since even minor boundary differences can induce area matching problems for small area estimation techniques using census data.

5. This feasibility report includes details of the preliminary statistical model fitting using the 2009 CSES and the 2008 CAS, which is Phase 2 of the WFP Massey University contract.

6. The general conclusion from this detailed statistical modelling is that small area estimation of stunting and underweight for Cambodia seems generally to be feasible at commune level, although the diagnostics for the models fitted to stunting and underweight are not as favourable as those fitted in the earlier feasibility study to poverty incidence, gap and severity. In practice this means that estimates for some smaller communes may need to be combined with neighbours. The models fitted to wasting are not currently adequate, although some improvement may be possible (as for the statistical models fitted to all six variables) once it is possible to reassess after adding village and commune level means for additional key variables from the census. This addition of census means is recommended for all the statistical models detailed in the two feasibility reports, as it has improved models - sometimes markedly - in other countries, for a variety of small area estimation of poverty measures.

7. As for the earlier 2008 feasibility study, no poverty estimates have been produced as part of this feasibility study. Such estimates require further funding beyond this feasibility phase, and availability of clean unit-record level 2008 census data.

8. The focus of the current report on stunting, underweight and wasting in children under five years of age strongly reflects the food security interests and concerns of the sponsor of the feasibility study, which is the World Food Programme. However WFP also recognises the centrality of and the importance to a wide range of international aid agencies of poverty incidence, gap and severity because these three measures together form a basis for sound measurement of economic poverty. For this reason, should small area estimation later be used in conjunction with the 2008 census data to produce poverty related measures on a local scale, this report recommends that maps should be produced for five of the six measures (i.e. stunting, and underweight – but possibly not wasting - in children under five years of age, plus poverty incidence, gap and severity).

9. The completion of both this and the earlier 2008 feasibility report follows extensive consultation with the National Institute of Statistics (NIS), other staff from the Ministry of Planning (MoP), International Fund for Agricultural Development (IFAD), Statistics Sweden, and World Bank, and World Food Programme (WFP) - which commissioned this research. For the current report, this consultation in Phnom Penh was during the period 24 May to 2 June 2010. The authors are grateful for these extensive contributions. Viewpoints and opinions expressed in this report do not however necessarily reflect those of all or any of the people or organisations consulted.

1 Introduction

This assessment is in two phases, which together provide a feasibility assessment of the potential for small area estimation of measures of malnutrition such as stunting, underweight and wasting in Cambodia. None of the two phases includes provision of small area estimates, however.

Phase 1, completed in 17 May 2010 – 13 August 2010, involved:

- Analysis of existing research and information on food security necessary for statistical assessment of malnutrition data for small area estimation in Cambodia based on reference material to be supplied by the World Food Programme and existing knowledge of small area estimation methods.
- Analysis of relevant Cambodia questionnaires from the Population Census 2008 and the Cambodian Anthropometric Survey (CAS) 2008 to be supplied by the World Food Programme.
- Identification and listing of questions asked in the census, and the sample survey (CAS 2008,) that *prima facie* are similar enough to be used for small area estimation of malnutrition (underweight, wasting, and stunting). This investigation will be based on English versions of questionnaires only.

Phase 2, completed in 14 August 2010 – 19 November 2010

- Identification of variables in both survey(s) and census(es) that are potentially useful for small area estimation of malnutrition (underweight, stunting, and wasting) in conjunction with the Cambodian National Institute of Statistics.
- Develop and test preliminary statistical regression models (including estimation of variance components) based only on the survey data, which will be supplied by WFP.
- Advise Cambodian National Institute of Statistics (NIS) and WFP how to identify Administrative Unit (area) code changes that may complicate analysis of the statistical relationship between survey and census data. This will be based on survey and census information supplied by NIS and WFP.
- Comment on the potential impact of these statistical analyses on small area estimation of malnutrition (underweight, wasting, and stunting).
- Preparation of the final report.

Completion of Phases 2 and 3 of the 2008 feasibility study *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia* included consultation and discussion with the following people and organisations:

National Institute of Statistics
H. E. San Sy Than, Director General, NIS, Ministry of Planning
Ing Sokun, Officer, General Statistics, NIS, Ministry of Planning
Khin Song, Deputy Director, General Statistics, NIS, Ministry of Planning
Nguon Sovann, Vice Chief, Bureau of Survey Planning, NIS, Ministry of Planning
Saint Lundy, Deputy Director, General Statistics, NIS, Ministry of Planning
Sin Serey Vuth, Chief, Bureau of Statistical Information, NIS, Ministry of Planning

National Strategic Development Plan (NSDP) Secretariat, Ministry of Planning
H. E. Tuon Thavrak, Director General, General Directorate of Planning, Ministry of Planning
Dr Hildegard Lingnau, Senior Advisor to Ministry of Planning, Centrum für Internationale Migration und Entwicklung (CIM)
Ramanathan Natarajan, International Consultant

Statistics Sweden – International Consulting Office
Sten Backlund, Chief Advisor, International Capacity Building Project at NIS
Agneta Sandqvist, Advisor Household Surveys, International Capacity Building Project at NIS
Lars Soderberg, Advisor ICT, International Capacity Building Project at NIS

United Nations Development Programme / International Fund for Agricultural Development
Ung Dara Rat Moni, IFAD/UNDP Advisor

United Nations World Food Programme
Thomas Keusters, Country Director, Cambodia
Bradley Busetto, Officer-in-Charge, Cambodia
Coco Ushiyama, Deputy Country Director, Cambodia
Kim Ratha, Senior Programme Assistant, Vulnerability Analysis and Mapping (VAM), Cambodia
Michael Sheinkman, Senior Regional Programme Advisor, Vulnerability Analysis and Mapping (VAM), WFP Bureau for Asia, Thailand

World Bank
Dr James Knowles, International Consultant.
Tim Conway, Senior Poverty Specialist
Neak Samsen, Poverty Specialist

The additional people consulted in Cambodia during the current 2010 feasibility study detailed in this report are:

UNICEF
Joel Conkle, Nutrition Specialist

Statistics Sweden – International Consulting Office
Dr Tiina Orusild, Senior Advisor Statistical Methodology, SCB/National Institute of Statistics

United Nations World Food Programme, Cambodia
Jean-Pierre de Margerie, Country Director, Cambodia
Kurt Burja, Programme Officer
Yav Long, National Program Officer, Vulnerability Analysis and Mapping
John Jeong, Programme Officer

2 Background

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. For example, census and new survey information may be combined to update estimates from the original census. Alternatively, and this is more usually the case for malnutrition 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 variables that are collected in both survey and census.

One of the recent studies involving small area estimation of malnutrition estimates in Cambodia is the *Micro-level Estimation of the Prevalence of Stunting and Underweight Among Children in Cambodia* from the Ministry of Health, Cambodia / World Food Programme / Measure DHS+ - ORC Macro (2003). This study uses the World Bank method for small area estimation to provide preliminary small area estimates for stunting and underweight in children. The statistical models used are not given, and the detailed methodology is not discussed, but maps are provided at commune level and averages of estimated accuracy of the small area estimates (as measured by their estimated standard errors given the fitted regression model is correct) are provided with discussion.

The World Bank method popularly known as the Elbers Lanjouw and Lanjouw (ELL) method has been commonly used in small area estimation of poverty measures. In poverty studies, the most usual variable predicted is 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. Poverty incidence, gap and severity are derived from the household level predictions of per capita expenditure. 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 sufficiently or rerunning the census. The statistical modelling has a cost, of course, but this is much lower than for a survey that is sufficiently large that it can produce estimates at this fine level. The cost of small area estimation can be saved many times over by having better information at a finer level and maps for use in aid allocation.

The initial, national, small area estimation of poverty in Cambodia was undertaken by Fujii (2002) for the World Food Programme, with support from the World Bank, using the 1998 population census and the 1997 Socio-economic Survey (CSES). By fitting a set of separate statistical models for expenditure on the logarithmic scale to sample information within strata for the CSES, applying these multiple models to the census data to predict expenditure at household level for all households, and summing transformations of the predictions, small area estimates of poverty incidence, gap and severity were derived, and mapped at commune level. The methodology used was a

standard application of the World Bank method (Elbers, Lanjouw and Lanjouw, 2001, 2003), which is now available as free software (PovMap – Zhao, 2006) from the World Bank website. 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 Bangladesh (Jones and Haslett, 2003) and Nepal (Jones and Haslett, 2006)

More recently, Pinney (2007) has undertaken a small area estimation exercise in Cambodia to update Fujii's estimates. Pinney has used the 2003/4 CSES and (rather than the population census, which as is common internationally is only conducted every ten years) has also used the commune database, also known as the Seila database or Seila commune database, or the National Committee for Decentralisation and Deconcentration (NCDD) database. The NCDD database is an annual census of villages and provides household information on a limited number of variables, which restricts the strength or predictive capacity (as measured by the percentage of variance that can be explained, usually denoted R^2) for statistical modelling, or predictions based on it. Pinney fits a multiple regression to the CSES data based on the variables also in the NCDD database, but without including the random effects (which would allow estimates of standard error via modelling of an additional commune or village level random component, fitted for example using the bootstrap as in ELL). The methods used by Pinney are potentially useful for providing an update to the 1997/8 estimates of Fujii, but the limited number of variables available for modelling may limit utility. The lack of information about standard errors is also a restriction, because poverty estimates are consequently of uncertain accuracy, so that it must remain unknown whether the method can provide sound poverty estimates at commune or district level.

The April 2007 World Food Programme report, *Integrated Food Security and Humanitarian Phase Classification (IPC) Pilot in Cambodia*, provides the most complete currently available comprehensive food security and vulnerability analysis. It has a direct focus on food, reflecting WFP's mandate. It contains a series of useful maps in appendices, including expenditure poverty (from CSES 2003/4) and underweight, stunting, and wasting in children. See also map on p44 – "Integrated Food security and Humanitarian Phase Classification (valid until 31.08.07) in Cambodia (as of 26.02.07)". None of these maps is however at commune level, so the need for small area estimates of poverty remains. It has a useful reference list but no statistics, or relevant methodological details or content, although see Section 1.2 Methodology, which outlines a "meta analysis approach".

This report and the *Micro-level Estimation of the Prevalence of Stunting and Underweight among Children in Cambodia* mentioned above warrant general comment about the relationship between small area estimation and mapping. 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 and Haslett, 2006), provides a detailed perspective on the spatial distribution of poverty. Other variables are also important however (e.g. health information, rainfall, and other Geographical Information System (GIS) data), even if these cannot be produced a such a fine level. 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 it also contains finer level tabulated detail. The detailed small area report 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 a more generally-used atlas must remain in doubt.

In September 2007, the *Statistical Master Plan for Cambodia* was published by the National Institute of Statistics, Ministry of Planning. This document outlines the development of statistical functionality at NIS. Page 20, as part of section 6.3 “Censuses and surveys”, contains detail on CSES as point 95, and Demographic and Health Surveys (DHS) as point 94. On page 21, there is Table 2, “Indicative Timetable for censuses and household surveys 2006-2015”. Small areas, but not small area estimation, are mentioned in item 89, p19.

3 Data Sources and Requirements

Application of the ELL method requires data of three types:

- a target variable *y* from which area-level estimates of the quantity of interest can be derived, available from a survey;
- a set of auxiliary variables, denoted x_1, x_2, ..., x_p; that are related to *y*, and available in both survey and census;
- 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.

We consider below the availability and quality of each of these components in the CAS2008 and CSES2009 surveys.

3.1 Height and weight data

The malnutrition indicators *stunting*, *underweight* and *wasting* are calculated for children less than five years old based on their height, weight and age. The three target variables from which area-level malnutrition indicators are derived are:

- standardized height-for-age (denoted *ZH*); a child is *stunted* if their *ZH* value is below –2, and severely stunted if below –3; stunting can be regarded as evidence of chronic malnutrition;
- standardized weight-for-age (denoted *ZW*); a child is *underweight* if their *ZW* value is below –2, and severely underweight if below –3; underweight reflects both chronic malnutrition and acute malnutrition: it is a current condition resulting from inadequate food intake, past episodes of under-nutrition or poor health conditions;
- standardized weight-for-height (denoted *ZWH*); a child is *wasted* if their *ZWH* value is below –2, and severely wasted if below –3; wasting can be an indicator of acute malnutrition.

Anthropometric measurements (height in cm, weight in kg) were made on children under 5 years old in both CAS2008 and CSES2009. Conversion to standardized z-scores involves subtracting the corresponding mean measurement and dividing by the standard deviation, the age-specific mean and standard deviation being obtained from a reference population (WHO Multicentre Growth Reference Study Group, 2006). This was accomplished using a macro provided by the WHO. The accompanying documentation suggests biologically plausible ranges for each of *ZH*, *ZW*, *ZWH*, and the WHO program flags any such values. We decided to eliminate from the survey data any observations outside of these limits (this approach was also taken by Fujii et al., 2004).

Table 3.1 shows, for each dataset, the percentage of the data flagged for being outside of the plausible range. (Because some data were missing for one variable but not the other, and because some observations were flagged more than once, the percentages do not add to 100).

Table 3.1. Percentage of data flagged as being outside the "biologically plausible" range

Flags	CAS2008	CSES2009
None	97.70	90.84
ZH	1.75	8.91
ZW	0.23	0.29
ZWH	1.23	4.43
Any	2.30	9.16

Although the weight measurement seems to be of high quality in both surveys, there is clearly a problem with the height measurements in CSES2009, resulting in nearly 10% of the data being eliminated as implausible.

Figure 3.1 examines the non-integer parts of the height and weight measurements in each survey. In CSES2009 many height measurements have been recorded to the nearest whole number, and many weights to the nearest 0.5. There has been some rounding of heights in CAS2008, but the problem is much less severe.

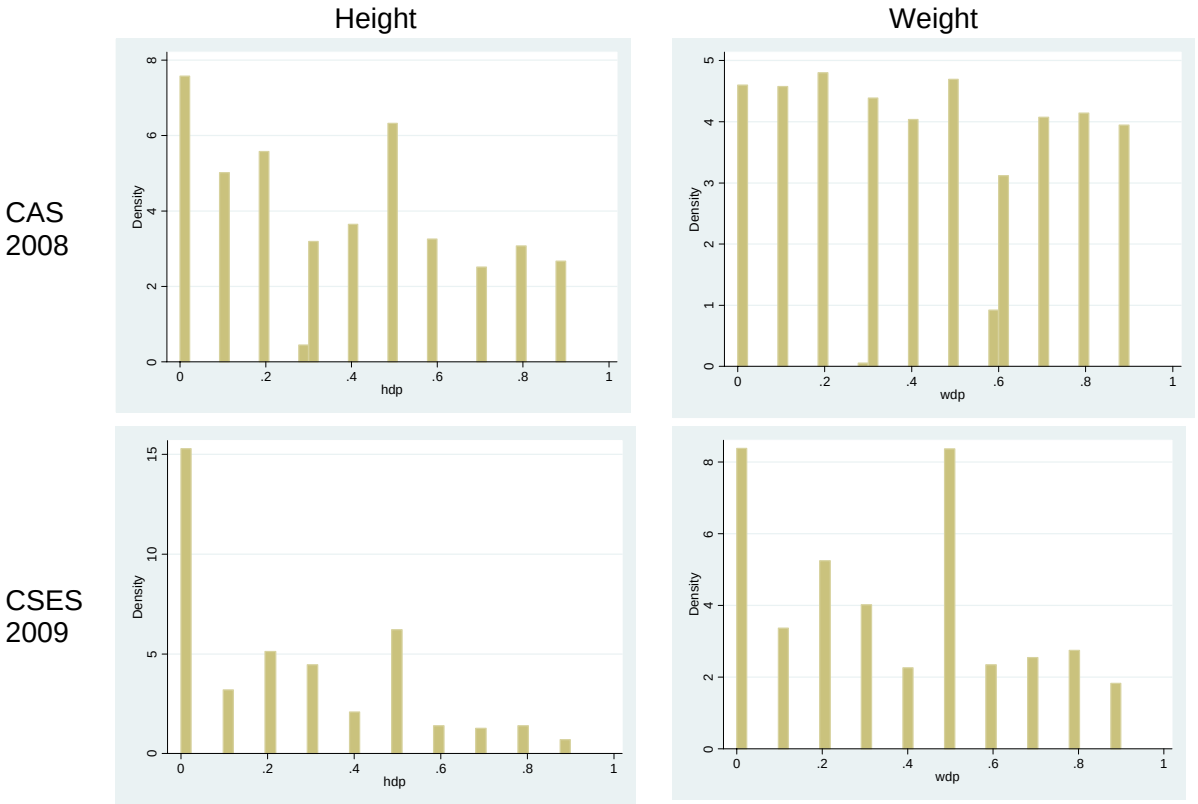

Figure 3.1. Non-integer parts of height and weight measurement

3.2 Auxiliary data

The modelling of the target variables (*ZH*, *ZW* and *ZWH*) in the ELL method uses auxiliary information which may be one of two types:

- variables identified in Phase 1 as being available in both the survey and census;
- variables missing from the survey data but available from the census or another external source as an average over a geographic area.

Variables of the second type, though potentially very useful in the modelling, are not available at this stage as they require summarising and merging census or other data at an appropriate level, so this feasibility study focuses only on variables identified as being in both the census and one or other of CAS2008 and CSES2009. A complete breakdown and description of these is given in Appendix A. For convenience we summarise in Appendix B the final lists of variables used for modelling, from CAS2008 in Table B1 and from CSES2009 in Table B2. Note however that household size, although identified as being present in CAS2008, was not in the CAS dataset.

These variables are also of two types: numerical variables such as household size (`hhsize`) that take numerical values within a range and categorical variables such as `roof` in CAS2008 that take one value from a limited range of categories (`rudimentary/natural/finished`). In the latter case one value is taken as the baseline or default, and binary (0/1) variables are created for the other values or levels. For example the categorical variable `roof` yields two binary variables `roof_natural` and `roof_finished`. in CAS2008, with `roof_rudimentary` as the default. The same variable in CSES2009 has seven levels so yields six binary variables, because the categories used in this survey match more closely the ones used in the census.

Tables B1, B2 show that there is a much richer set of auxiliary variables in CSES2009 than in CAS2008, so *prima facie* we would expect to get better models for predicting the target variables using this dataset.

3.3 Regional indicators

The hierarchy of regional and administrative areas for Cambodia are the provinces, districts, communes and villages, provinces being the largest. Table 3.2 summarises the number and size of each, based on the 1998 census (2008 data not available at this time).

Table 3.2. Structure of 1998 Census giving number and average size (households/villages) for each unit

	province	district	commune	village	hh
Number	24	180	1594	13750	2162086
Mean hh	90086.9	12011.6	1356.4	157.2	
Mean v	572.9	76.4	8.6		

It is hoped that the small-area technique will provide sufficiently precise estimates of malnutrition indicators (proportions stunted, underweight or wasted) at commune level. The primary sampling units (PSUs) for both surveys were villages. Because the anthropometric measurements relate to individual children, both surveys have a nested structure: children within households within villages within communes.

Table 3.3. Structure of CAS2008

	province	district	commune	village	hh	child
Number	24	185	709	760	7268	8557
Mean #child	356.5	46.3	12.1	11.3	1.2	
Mean #hhold	302.8	39.3	10.3	9.6		
Mean #vill	31.7	4.1	1.1			

Table 3.4. Structure of CSES2009

	province	district	commune	village	hh	child
Number	24	169	613	693	3993	4869
Mean #child	202.9	28.8	7.9	7	1.2	
Mean #hhold	166.4	23.6	6.5	5.8		
Mean #vill	28.9	4.1	1.1			

Tables 3.3 and 3.4 show the nested structure of each survey after the biologically implausible and missing values have been eliminated. CAS2008 has a lot more children and households, but approximately the same average number of children per household. CAS has slightly more villages sampled, and more households per village. Both have similar numbers of sampled villages per commune, and per district, but CAS2008 appears to cover all districts whereas CSES2009 does not. This structure has an important bearing on model fitting and calculation of standard errors, as we discuss later.

We also need to consider the survey stratification in calculating standard errors. The strata for CSES2009 are the provinces subdivided into urban and rural; the CAS2008 strata do not seem to map directly to any structures in the census. CSES2009 has five strata with only one PSU; CAS2008 has 44 such. This is a problem when calculating valid standard errors for the estimated parameters in model-building; some strata will have to be merged, and this being an exercise reliant on local knowledge will best be done later in collaboration with the National Institute of Statistics.

It is important that the codes used to identify these regional indicators match in the census and survey, at least down to commune level and preferably to village level. This is so that we can merge commune- or preferable village-level auxiliary data from other sources with the survey data to assist in the final model-building. We have not been able to check this thus far, as the census data are not available to us, but we understand that the codes are compatible with the 2008 Census for both surveys. Nevertheless careful checking of this will need to be done in conjunction with the National Institute of Statistics.

4 Modelling

The basic ELL method uses the auxiliary data to infer the value of the target variable *y* for every household in the census via a statistical model

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

where y_{vh} denotes the value of the target variable in the *h*th household of the *v*th village. 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 village 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.

However, since we are dealing with child-level data, some extension of this is advisable. Instead of a two-component error variance structure (household within village) we should consider at least three (child within household within village). But this makes the implicit assumption that there is no commune-level component in the unexplained variation in (1), so we could try extending to four components (child_household_village_commune). This could have an important bearing on the standard errors of the final small-area estimates (see Haslett and Jones, 2010a, in press). In fact, there are some children in the same household born to different mothers and this information is available in CAS20008, so in theory one could use a five-level structure (child_mother_household_village_commune). In our modelling we explore the feasibility and desirability of these alterative error structures.

4.1 Estimation methods

The data used to estimate the model (1) typically comes from a two-stage survey design: in the first stage “primary sampling units” (PSUs) are randomly sampled within each stratum, after which individual households are selected at random within each sampled PSU. The PSUs usually represent natural clusters of households within the population; in CAS 2008 and CSES2009 the PSUs are villages. One consequence of this design is that individual households can have different representational values within the population; this is reflected in the different “survey weights” given to each household. Statistical methods exist for taking into account the survey weights and the two-stage structure of the sample.

Early implementations of ELL used ordinary least squares (OLS) regression (i.e. ignoring the clustering and the weights) when selecting the variables to be included. The chosen model was then re-estimated using a form of generalized least squares (GLS) in which estimated variance components (for $\sigma_\eta^2, \sigma_\varepsilon^2$) were used to re-weight households within clusters. This GLS estimation is included in the PovMap software provided by the World Bank. An alternative is to perform the model selection using a survey regression method specifically designed for the purpose and available in some common statistical programs such as Stata. It is sometimes found that variables which seem to have a significant effect in OLS estimation lose their significance when the survey design is accounted for. There are other, more sophisticated technologies available (e.g. You et al, 2003) for

estimating the final model and its variance components. The differences obtained from these various methods do not seem to be great. Because the present exercise is exploratory, we consider here only Stata's survey regression method for model fitting, using the survey weights and robust standard errors but ignoring the stratification and clustering because some strata have only one PSU. We then decompose the residuals from the fitted model by fitting an unweighted random effects model, with four components (child_household_village_commune) if possible.

4.2 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 stunting is the proportion of children in an area for which *ZH* is below – 2). 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. For malnutrition indicators however, sufficiently precise estimates of stunting, wasting and underweight have been obtained using models where the R^2 was as low as 20% (Jones and Haslett, 2006) because of the different relative contributions of each of the variance components.

It should be noted that 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 aim should be to achieve a reasonable R^2 with a modest number *p* of auxiliary variables in relation to the size of the estimation dataset, preferably without subsetting the data before fitting models.

In the extended version of (1) the unexplained variation is decomposed into effects at three or four levels (commune_village_household_child). Denote the variances of these components by σ_c^2, $\sigma_v^2, \sigma_h^2, \sigma_e^2$ respectively. When the model is used to predict *y* for each child in the census, simulated values of random effects at each level are included to incorporate the uncertainty in these predictions. When the child predictions are amalgamated to small-area (commune) 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 numbers of children, households, and villages in the small areas. Since the numbers of children and households are generally much larger than the number of villages, the crucial factor affecting precision is usually the size of the village-level effects. Any commune-level effect, if included, would not average out at all. Thus it is particularly important that σ_c^2 and σ_v^2 should be as small as possible.

Table 4.1. Modelling results for round one of Fujii et al (2004). Here σ_u^2 represents total error variance.

	Height (cm)				Weight(kg)			
Zone	R^2	*p*	σ_v^2/σ_u^2	σ_u^2	R^2	*p*	σ_v^2/σ_u^2	σ_u^2
Urban	0.46	28	0.00	19.37	0.53	28	0.00	1.20
Plain	0.45	43	0.00	19.19	0.40	34	0.00	1.60
Tonlesap	0.43	55	0.00	18.51	0.40	40	0.00	1.46
Coastal	0.65	19	0.00	17.25	0.63	19	0.00	1.13
Plateau	0.47	32	0.00	16.58	0.47	32	0.00	1.26

Table 1 gives the values of these summary measures for the models used in round one of the malnutrition mapping exercise reported by Fujii et al (2004) based on Cambodia Health and Demographic Survey of 2000. Here different models have been fitted in each of five ecozones, rather than a single overall model which has a strong influence on sample size available for each fit. The resulting commune-level poverty incidence estimates were eventually superseded by a round two analysis using a more sophisticated methodology with three levels of variance components (adding household-level) and a multivariate approach in which height and weight were modelled simultaneously. No diagnostic measures were reported for this second round. Note that Fujii et al (2004) standardized height and weight by transforming first to the equivalent 24-month old healthy female, but this should be equivalent to modelling the standardized values *ZH* and *ZW* since the methodology is affine invariant. In our experience the R^2 reported by Fujii are remarkably high, and the village-level variances remarkably low.

4.3 Variable selection

There is typically a large set of auxiliary variables available for inclusion in the model (1). It is not good practice to include all of these *x* variables as the complete set would be highly multicollinear, giving low precision in the estimates, and possibly many spurious relationships leading to bias, especially if the survey data is subset before model fitting. A model selection procedure is needed to decide which variables to include and to avoid this "over-fitting". We want to achieve a reasonably high R^2 and low with only a moderate number *p* of variables.

The model may also include "interactions", allowing the effect of one variable to change with different values of another variable. For example the effect of household size may vary between urban and rural areas; to accommodate this we would include an interaction variable for `hhsize` and `rural`, which we denote `hhsizeXR`, rather than fitting separate rural and urban models.

Since the commune- and village-level error components play the largest part in the precision of the final small-area estimates, we try to get σ_c^2 and σ_v^2 as small as possible.

This is often aided by using area-level variables such as GIS data and census means. Care must be taken however not to fit too many of these in relation to the number of

communes or villages in the survey data, as spuriously optimistic results may be obtained. This use of census means has not been possible at this stage, as the 2008 census data is not yet available in this form.

Some approaches to the model selection problem have first created a maximal pool of potential *x* variables by adding quadratic and cubic terms for all numeric variables and all possible two-way (sometimes even three-way) interactions, then running an automatic model selection method (stepwise regression) that lets the computer choose the model. The result can be over-complicated and hard to interpret; it also increases the possibility of spurious relationships in the model. We tend to prefer a more cautious approach, starting with a relatively simple and interpretable model and judiciously adding nonlinear and interaction terms where they seem necessary and where the results seem to be plausible

Some implementations of the ELL method have used different models in different strata, these strata being defined by the survey sample design or, in the case of a survey design with many small strata, as accumulations of geographically contiguous strata. Previous surveys in Cambodia have used Phnom Penh, Other Urban, and Rural as strata, but in recent years there as been a trend towards using many small strata. One possible subdivision for modelling purposes is to use the five ecozones Urban, Plain, Tonlesap, Coastal, Plateau as was done in Fujii et al (2004). One drawback with this use of many models is that there may be few surveyed villages in some of these areas, leading to the possibility of spurious relationships and over-optimistic results. This is perhaps a partial explanation for the remarkable summary statistics presented in Table 4.1 for the first round regression models of Fujii et al (2004). We prefer to fit a single model for the whole country and introduce regional variations in parameters using interaction terms as appropriate.

5. Results

Full details of the variables fitted in the modelling of the three target variables in each of the two surveys are presented in Appendix C. These should be regarded as indicative for the purposes of the feasibility study rather than final models suitable for the production of small-area estimates. Here we present and discuss summary statistics for the models.

Table 5.1. Modelling results for CAS2008.

Target	*p*	R^2(%)	σ_c^2	σ_v^2	σ_h^2	σ_e^2
ZH	34	19.9	na	0.0758	0.4072	1.4035
ZW	26	15.6	na	0.0673	0.2723	0.8167
ZHW	25	3.4	0.0059	0.0780	0.1876	1.0795

Table 5.1 gives the number of *x* variables used (*p*) the R^2 achieved, and the estimated variance components for each of the models fitted using CAS2008. It is difficult to estimate the variance components because most communes had only one sampled village and most households had only one eligible child (see Table 3.3). For most of our models a variance structure with four components could not be estimated; for those when it could, the estimated commune-level component was small with a confidence interval suggesting that it could be negligible, as is desirable for accurate small area estimation. Thus when the four-level structure was inestimable, we reverted to a three-level structure without the commune level.

It is interesting to note that the largest component is that of variation between children in the same household (σ_e^2). This concurs with the finding of Jones and Haslett (2006) in Nepal. Since there are few child-level auxiliary variables, this component will not be reduced by further modelling. In any case, the child-level covariates available (age, sex) may not be particularly useful in predicting area-level aggregates if their distributions do not vary significantly between areas. Since this within-household variation will average out across small-areas, the models used for mapping malnutrition indicators should probably not require, or be expected to achieve, particularly high R^2. This makes the R^2 values presented in Table 4.1 for the first round regression models of Fujii et al (2004) even the more remarkable.

Regarding the village-level components (σ_v^2), these are at present a little high. However the current models do not include commune- and village-level means, derivable from the census and other databases but not currently available. Inclusion of such variables in the final modelling exercise would be expected to lead to a modest increase in R^2 and perhaps a considerable drop in village-level variance.

Nevertheless, the results for the wasting indicator (*ZWH*) do not look promising. For some reason there was little explanatory power for weight-for-height in the CAS2008 auxiliary data. There is the potential for improvement when area-level means are made available, but the improvement would have to be dramatic for commune-level estimates of wasting to be feasible from this source. This is in marked contrast to the results of Jones and Haslett (2006) in Nepal, and we are looking for an explanation.

Table 5.2. Modelling results for CSES2009.

Target	p	R^2(%)	σ_c^2	σ_v^2	σ_h^2	σ_e^2
ZH	55	8	0.0180	0.3515	0.4206	2.4583
ZW	56	12.5	na	0.1169	0.2489	1.0681
ZHW	52	6	na	0.3399	0.3074	2.0964

Table 5.2 gives the corresponding summary of models fitted using the CSES2009. Although there is a much richer set of auxiliary variables in this dataset, doubts were expressed earlier in section 3.1 about the quality of the anthropometric measurements, particularly with regard to height (and by implication weight-for-height). The modelling largely confirms these reservations. Despite using 55 variables, we were unable to explain more than 8% of the variability in standardized height-for-age. Even the model for weight (*ZW*) does not reach the R^2 achieved from CAS2008, despite using many more variables. The village-level variance components are also significantly larger than in the CAS2008 models. (It is perhaps possible that some of the variability at this level might be due to differences between the people doing the measuring). Again, some of this variability might be reduced by the inclusion of area-level means.

Surprisingly perhaps, the R^2 for standardized weight-for-height (*ZWH*) is higher with the CSES2009 data than with CAS2008, albeit with many more *x* variables. This minor advantage is outweighed by the much larger village-level error component in the CSES2009 model, so that for *ZWH* neither the model fitted to CSES2009 nor that fitted to CAS2008 is currently adequate as a basis for small area estimation.

6. Conclusions

The modelling results for standardized height-for-age and weight-for-age using the CAS2008 data suggest that small-area estimation of stunting and underweight could yield reasonably precise estimates of stunting and underweight at commune level, at least for most communes. This will depend on the availability of further commune- or preferably village-level auxiliary information with matching regional codes, as might be provided from the census data or other matching village-level databases. The average number of villages per commune is around nine (Table 3.2), which is quite small compared to our usual experience with poverty and malnutrition mapping in which the targeted small areas are typically 20 or more PSUs. It is therefore particularly important here to minimise residual village-level variation in the models. It may be necessary to combine some of the smaller communes in the final stage in order to achieve a useful level of precision.

The CSES2009 data does not seem as useful, particularly for stunting where the height measurements seem to be of dubious quality. However it would be possible to proceed independently with both data sources, combining the estimates in the final stage according to their relative precision, and thereby produce a single set of estimates with more precision than those obtained by using CAS2008 alone.

Small-area estimation of wasting does not seem to be feasible because of the poor explanatory power of the fitted models for standardized weight-for-height. We do not know why this is. It might be useful anyway to include wasting along with stunting and underweight in a full malnutrition mapping exercise, to see how much improvement in modelling can be made when area-level means are available, to evaluate the performance of the technique in situations where R^2 is very low, and to investigate the reasons why, in our data, there is so little relationship between wasting and other variables.

References

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.

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.

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.

Fujii T., Lanjouw P., Alayou S. and Montana L. (2004) Micro-level estimation of prevalence of child malnutrition in Cambodia. World Bank Report.

Haslett S.J. and Jones G. (2010a) Small-area estimation of poverty: the aid industry standard and its alternatives. *Australian & New Zealand Journal of Statistics*, in press.

Haslett. S.J. and Jones, G. (2010b) *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia*, report to World Food Programme, December 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.J. (2006) Small area estimation of poverty, caloric intake and malnutrition in Nepal, Published: Nepal Central Bureau of Statistics / World Food Programme, United Nations / World Bank, September 2006, 184pp, ISBN 999337018-5.

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

Ministry of Health, Cambodia / World Food Programme / Measure DHS+ - ORC Macro (2003). *Micro-level Estimation of the Prevalence of Stunting and Underweight Among Children in Cambodia*

National Institute of Statistics, Cambodia (2007) *Statistical Master Plan for Cambodia*, National Institute of Statistics, Ministry of Planning, v+40pages, September 2007.

Pinney, A. (2007) *Initial calculation of small area estimates of poverty using CSES 2003/04 and Commune Database 2002-2006*, draft report: UNDP / National Institute of Statistics. July 2008.

WHO Multicentre Growth Reference Study Group (2006). WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: World Health Organization; pp 312. : http://www.who.int/childgrowth/publications/en/

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

Note: A more extensive reference list is given in Haslett and Jones (2010b) *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia*, report to World Food Programme, Dec 2008.

Appendix A

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009 and between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

General Notes for Tables CSES A1-A3, CSES B1-B16, CAS A1 and CAS B1-B8

1. Tables CSES A1-A3, CSES B1-B12 provide the detail necessary for preliminary matching of variables in the 2008 General Population Census of Cambodia, GPCC2008 (for which the relevant questions are almost identical to the 1998 census) and the Cambodian Socio-economic Survey (CSES). For the CSES, the 2009 questionnaire is detailed below. As for the census, CSES questionnaires are very similar, although in years between 2003/4 and 2009, the CSES (when implemented, as in 2007) uses a smaller sample (n=3600 versus n=15.000 approximately) and a reduced bank of questions.

2. Tables CAS A1, B1-B8 provide the detail necessary for preliminary matching of variables in the 2008 General Population Census of Cambodia, GPCC2008 (for which the relevant questions are almost identical to the 1998 census) and the Cambodian Anthropometric Survey (CAS 2008).

3. Census Forms A and B were not completed simultaneously. Form A was collected several days previous to Form B (ref: 2008 Census Enumerators' Manual p3). This can lead to some mismatching of census households between Form A and Form B.

4 Questionnaire details have been compared in these tables via their English and Khmer versions; through extensive discussions with Cambodian National Institute of Statistics (NIS) staff. For final matching decisions (between survey and census) local Khmer knowledge of the questions actually asked will continue to be essential. The tables below list agreement or otherwise in principle. Note that even where questions are identical in English, they may not be in Khmer. Even if identical in Khmer, 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 the 2008 census data.

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

6. Standardizing by using the same categories and identical questions in census and survey is highly recommended in future, if further small area estimation using a combination of both data sources is planned.

7. The following parts of the census are not directly relevant to matching with the CSES 2009 or CAS 2008 survey questionnaires:
Form B, Household Questionnaire Part 1
Charge Register
List of Defence Establishments
List of Police Headquarters
List of Major Institutions
List of Villages with Boat Population
List of Remote Villages difficult to Reach
Training Centres at District Headquarters
Number of Trainers / trainees
List of Census Officers
Enumerator Summary Account for Questionnaires Received by Enumerator
Supervisors’ Summary
Receipt of materials Given
List of Filled-in Records Submitted by Enumerator
Commune Population Statement (Provisional)
District Population Statement (Provisional)
Appointment Order for Enumerators / Supervisors
Appointment Order for the Census Officer

8. The majority of the CSES 2009 and CAS 2008 survey questions are not collected in the census. With the exception of expenditure related information from CSES 2009 (for poverty incidence, gap and severity) and height, weight and age of children under five (for stunting, underweight and wasting), such non-matching questions are irrelevant. This is because small area estimation using variants of the World Bank methodology for small area poverty estimation models expenditure using survey data and only uses survey variables that match between survey and census for the prediction of expenditure (and hence poverty incidence, gap and

severity) at household level, which are then combined to small area level. A parallel comment applies to modelling stunting, underweight and wasting.

9. Physical/mental disability, basis of house occupancy (e.g. own, rent), , and ownership of radio, television, telephone (fixed), cell phone, personal computer, bicycle, motorcycle, car / van, boat, tractor are collected for the 2008 census, which were not collected for the CSES 2003/4, were collected in CSES 2009, and so have been added to the list of matching variables outlined in the earlier 2008 report, *Potential for Small Area Estimation and Poverty Mapping at District and Commune Level in Cambodia.*

10. There are migration related questions in both 2008 census (Form B Household Questionnaire Part 2, Individual particulars columns 9-12) and 2009 CSES (Section 03) but the questions are different and coding match would be difficult without the extensive local knowledge at the Cambodian National Institute of Statistics (NIS).

11. Tables CSES A1-A3, CSES B1-B16, CAS A1 and CAS B1-B8 below are organised essentially in the same order as the questions appear in the 2008 census questionnaires

CSES 2009

Table CSES A1:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Wall material – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Wall material	Form A: Houselist, p1, column 3.	1. Bamboo / thatch / grass / reeds 2. Earth 3. Wood / plywood 4. Concrete / brick / stone 5. Galvanised iron / aluminium / other metal sheets 6. Asbestos cement sheets 7. Salvaged / improvised materials 8. Other (specify)	Section 04 Housing: Q4	1. Bamboo / thatch / leaves / grass 2. Wood or logs 3. Plywood 4. Concrete / brick / stone 5. Galvanised iron or aluminium or other metal sheets 6. Fibrous cement / asbestos 7. Makeshift, mixed materials 8. Clay / dung with straw 9. Other (specify)	Some but not all categories are identical; amalgamation of categories likely to be required to match census and survey information

Equivalence Table: ***Wall Material***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Wall material	1	1
	2	8
	3	2, 3
	4	4
	5	5
	6	6
	7	7
	8	9

Notes: Logs are included are implicit in the 2008 census category, but explicitly in CSES 2009 category 2

Table CSES A2:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Roof material – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Roof material	Form A: Houselist, p1, column 4	1. Bamboo / thatch / grass 2. Tiles 3. Wood / plywood 4. Concrete / brick / stone 5. Galvanised iron / aluminium / other metal sheets 6. Asbestos cement sheets 7. Plastic / synthetic material sheets 8. Other (specify)	Section 04 Housing: Q5	1. Thatch / leaves / grass 2. Tiles 3. Fibrous cement 4. Galvanised iron or aluminium 5. Salvaged materials 6. Mixed but predominantly made of iron / aluminium, tiles or fibrous cement 7. Mixed, but predominantly made of thatch / leaves / grass or salvaged materials 8. Concrete 9. Plastic sheet 10. Other (specify)	Some but not all categories are identical; amalgamation of categories likely to be required to match census and survey information

Equivalence Table: ***Roof Material***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Roof material	1	1
	2	2
	3, 8	10
	4	8
	5	4
	6	3
	7	9

Notes

1. Bamboo (which is included in census category 1, but not in CSES 2009 category 1) is a rare roof material in Cambodia.
2. Category 3 in the census (Wood and plywood) included under category 10 (Other) in CSES 2009.
3. Category 4 in the census includes brick and stone which are uncommon as roofing materials, so category 4 in the census has been matched with category 8 in CSES 2009.
4. Add census categories 5, 6, 7 to CSES category 10. Alternatively, census categories 1 and 5 are matched respectively to CSES 2009 categories 7 and 6 respectively, and census category 5 is matched to CSES 2009 category 10.

Table CSES A3:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Floor material – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Floor material	Form A: Houselist, p1, column 5	1. Earth / clay 2. Wood / bamboo planks 3. Concrete / brick / stone 4. Polished stone 5. Parquet / polished wood 6. Mosaic / ceramic tiles 7. Other (specify)	Section 04 Housing: Q6	1. Earth / clay 2. Wooden planks 3. Bamboo strips 4. Cement / brick / stone 5. Parquet / polished wood 6. Polished stone / marble 7. Vinyl 8. Ceramic tiles 9. Other (specify)	Some but not all categories are identical; amalgamation of categories likely to be required to match census and survey information

Equivalence Table: ***Floor Material***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Floor material	1	1
	2	2, 3
	3	4
	4	6
	5	5
	6	8
	7	7, 9

Notes 1. Mosaic in category 6 of the census is not explicit in CSES 2009 category 8
2. Vinyl in CSES category 7 means plastic, and has been coded to census category 7 (Other).

Table CSES B1:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Household size, Head of household - Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Household size	Page 3, Form B: Household questionnaire Part 2, derivable as sum of number of entries (i.e. names) in column 2.	Numeric: 1, 2, 3,…..	Section 01"Initial visit" A "List of household members", derivable from this "Listing of household members", column 2.	Numeric: 1, 2, 3,…..	Usually need to watch during modelling for household sizes in census that are very much in excess of those in the survey, especially if a 'household size squared' term is included in the survey-based model.
Household head	Page 1 Form B, Household Questionnaire, column 3; also collected in Form A, column 8.	In Form A, name and sex of head of household only is collected. In Form B, "Relationship to head of household" has the instruction "(Write in words)" below it on the questionnaire.	Section 01"Initial visit" A "List of household members", via "Relationship to head of household", column 6.	Code '01' denotes head of household	The name and sex of head of household is available from census Form A, and name and other personal characteristics including sex from Form B. Note that census question is asked twice. Personal characteristics of head of household including name are also available from the CSES questionnaire.

Table CSES B2:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Sex, Age – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Sex	Page 3, Form B: Household questionnaire Part 2, column 4.	1. Male 2. Female	Section 01 "Initial Visit", A, "List of household members", derivable from "Sex", column 3	1. Male 2. Female	Hence proportion of males or proportion of females (or proportions of males and females within age ranges using sex, and age below)
Age	Page 3, Form B: Household questionnaire Part 2, column 5.	Codes 0-97 correspond to age in completed years, 98 to 98 or more	Section 01 "Initial Visit", A, "List of household members", derivable from "Age in completed years", column 5	Codes 0-95 correspond to age in completed years, 96 to age 96 or more, and * to don't know	Take care with '*' code for CSES; can cross check age against date of birth question for CSES only

Table CSES B3:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Marital Status, Languages Spoken – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Marital Status	Page 3, Form B: Household questionnaire Part 2, column 6	1. Never married 2. Married (i.e. currently married) 3. Widowed 4. Divorced 5. Separated	Section 01 "Initial Visit", A, "List of household members", column 9	1. Married / living together 2. Divorced / separated 3. Widowed 4. Never married / never lived with partner	"Living together" in CSES 2009 is "never married" in census. It is unclear how to resolve this, except by having only three categories: widowed, divorced / separated, other. However then, married and not married are in the *same* category, "other".
Languages spoken	Page 3, Form B: Household questionnaire Part 2, column 7, which allows a maximum of three	29 options including all six of those specified in CSES 2009	Section 01 "Initial Visit", A, "List of household members", column 12a, 12b, 12c, i.e. maximum of three Note: Ethnicity is collected in CSES 2009 but not in census	0. None 1. French 2. English 3. Chinese 4. Vietnamese 5. Thai 6. Lao 7. Other (specify)	The CSES 2009 question is, "Can you speak other languages than Khmer?" – all are foreign; the census 2008 question is, "Mother Tongue". This means, since most people's primary language is Khmer, that this question cannot be matched.

Table CSES B4:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Literacy – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Literacy	Page 4, Form B: Household questionnaire Part 2, column 13 (a) and (b)	Yes 2. No.	Section 02 "Education and Literacy", (2) and (3)	1. Yes 2. No	Census asks "read and write with understanding" in two parts: (a) Khmer (b) any other language. CSES 2009 has separate questions for reading (column 2) and writing (column 3) "a simple message" in any language.

Table CSES B5:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Education – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Education	Page 4, Form B: Household questionnaire Part 2, column 14 (a) and (b)	(a) Attending school 1. Never 2. Now 3. Past (b) Highest grade - Never attended 00 No class completed 01 Class 1 completed …. 12 Class 12 completed 13 Lower 2ndary diploma 14 2ndary school baccalaureate holder 15 Technical / vocational pre-2ndary diploma 16. Technical / vocational post-2ndary diploma 17. Undergraduate 18. Graduate 19. Postgrad and above 20. Other (specify)	Section 02 (6)	Highest grade 90 None 98. Don't know 88 No class completed 00. Preschool / kindergarten 01 Class 1 completed …. 12 Class 12 completed 13 Lower 2ndary school certificate 14 Upper 2ndary school certificate 15. Technical / vocational pre-2ndary diploma 16. Technical / vocational post-2ndary diploma 17. College / university undergraduate 18. Bachelor degree (B.A., B.Sc. etc.) 19. Masters degree (M.A., M.Sc. etc.) 20 Doctorate degree (PhD) 21. Other (specify)	The equivalent to census 14(a) - the attending school question – can in principle be derived from the CSES 2009 The codes on highest grade are very similar except for codes -, 00, 88, 90, 98, and census codes 17 and 18 representing combinations of CSES 2009 categories

Equivalence Table: ***Education***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Education	-, 00	90, 88, 00, and possibly 98
	01	01
	02	02
	03	03
	04	04
	05	05
	06	06
	07	07
	08	08
	09	09
	10	10
	11	11
	12	12
	13	13
	14	14
	15	15
	16	16
	17	17
	18	18
	19	19, 20
	20	21

Table CSES B6:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Employment – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Employment	Page 4, Form B: Household questionnaire Part 2, column 16-23. Questions asked are: Main activity Employment period Occupation Employment status Industry, trade or service Sector 2ndary economic activity Place of work or school Proportion of year employed - B(17)	See census questionnaire page 4, Form B	Sections 15& 16, with focus on Section 16 – Activities in the last 12 months Questions asked are For the last 7 days - Economic activity Number of occupations Primary occupation Secondary occupation Type of economic activity Hours worked For the last 12 months - Section 16 Main activities 1ary (2)and 2ndary (8) occupations Kind of economic activity Proportion of year employed – (3).	See CSES questionnaire Form 3, Section 13, A-C, for details	The questions in the employment category in census and CSES 2009 are similar but not identical. For example, the periods for which the questions apply are different, as do the times of year to which they apply since CSES 2009 and census were not conducted at the same time of year. Some additional information may be able to be extracted, but to simplify primary emphasis should be placed on the 12 month period and whether employed or not, plus agricultural links. Further advice from NIS also advisable.

Table CSES B7:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Fertility - Individual level

Note: Fertility measures in the 2008 census include: number of children born alive, number of children living, and how many children have died. Although this fertility information was also collected in CSES 2003/04, it was not collected in CSES 2009. (For the survey that did collect these fertility measures, see tabulations for the CAS 2008 questionnaire below.

Table CSES B8:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Housing conditions: Source of light – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Main source of light	Page 6, Form B: Household questionnaire Part 4, column 2	1. City power 2. Generator 3. Both 1 and 2 4. Kerosene 5. Candle 6. Battery 7. Other (specify)	Section 04 Housing, Q7	1. Publicly owned electricity / city power 2. Generator 3. Battery 4. Kerosene lamp 5. Candle 6. None 7. Other (specify)	Similar categories. Check whether percentages match after amalgamating categories for both survey and census before fitting models for any final small area estimation

Equivalence Table: ***Source of light***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Source of light	1	1
	2	2
	3	1 & 2
	4	4
	5	5
	6	3
	7	6, 7

Notes: NIS notes that CSES 2009 category 3 may be classifiable as category 1, since when a household has city power working, they will not use a generator.

Table CSES B9:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Housing conditions: Main cooking fuel – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Main cooking fuel	Page 6, Form B: Household questionnaire Part 4, column 3	1. Firewood 2. Charcoal 3. Kerosene 4. LPG 5. Electricity 6. None 7. Other (specify)	Section 04 Housing, Q22.	1. Firewood 2. Charcoal 3. Liquid petroleum gas LPG 4. Kerosene 5. Publicly provided electricity / City power 6. Household generator 7. None / don't cook 8. Other (specify)	Similar categories. Check whether percentages match after amalgamating categories for both survey and census before fitting models for any final small area estimation

Equivalence Table: ***Main cooking fuel***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Main cooking fuel	1	1
	2	2
	3	4
	4	3
	5	5, 6
	6	7
	7	8

Notes: NIS notes that CSES 2009 category 6 may be classifiable instead as category 7 in the census (rather than category 5

Table CSES B10:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Housing conditions: Toilet facility within premises – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Toilet facility within premises	Page 6, Form B: Household questionnaire Part 4, column 4	1. Not available 2. Connected to sewerage 3. Septic tank 4. Pit latrine 5. Other type (specify)	Section 04 Housing Q19a (on premises, i.e. close to dwelling) , Q19b (off premises)	1. Pour flush (or flush) connected to sewerage 2. Pour flush (or flush) to septic tank or pit 3. Pour flush (or flush) to elsewhere (i.e. not a sewer or pit/tank 4. Pit latrine with slab 5. Pit latrine without slab or open pit 6. Latrine overhanging field or water (drop in the field, pond, lake, river, sea) 7. None 8. Other	Similar categories except that census only asks what toilet facilities are on the premises. Off-premise categories for survey can be recoded as equivalent to census code 1. 'not available' (on premises). Check whether percentages match after amalgamating categories for both survey and census before fitting models for any final small area estimation

Equivalence Table: ***Toilet facility within premises***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Toilet facility within premises	1	7
	2	1
	3	2
	4	4,5
	5	3, 6, 8

Table CSES B11:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Housing conditions: Main source of drinking water – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Main source of drinking water	Page 6, Form B: Household questionnaire Part 4, column 5 "main source of drinking water supply"	1. Piped water 2. Tube / pipe well 3. Protected dug well 4. Unprotected dug well 5. Rain 6. Spring, river, stream, lake/pond 7. Bought 8. Other (specify)	Section 04 Housing, Q8 (wet season), Q12 (dry season) p10	1. Piped in dwelling or on premises 2. Public tap 3. Tubed / piped well or borehole 4. Protected dug well (including lining, headwall, platform, cover) 5. Unprotected dug well 6. Pond, river or stream (fetch water from pond, river, stream) 7. Pond, river or stream (pump to the house) 8. Improved rainwater collection 9. Unimproved rainwater collections 10. Water bought from tanker, truck, vendor (at home) 11. Water bought from tanker, truck, vendor (at distance) 12. Bottled water 13. Other (Specify)	Similar categories, but watch for seasonal changes in water sources given timing of survey and census. Check whether percentages match after amalgamating any categories for both survey and census before fitting models for any final small area estimation Note there are other questions on water in both survey and census. These, while different, may still be able to help with matching of categories.

Equivalence Table: ***Main source of drinking water – Household level***

Topic	***2008 Census Categories***	***2009 CSES Categories***
Main source of drinking water – Household level	1	1, 2
	2	3
	3	4
	4	5
	5	8, 9
	6	6, 7
	7	10, 11, 12
	8	13

Table CSES B12:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Housing conditions: Number of rooms occupied by household - Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Number of rooms occupied by household	Page 6, Form B: Household questionnaire Part 4, column 7	1. One room 2. Two rooms 3. Three rooms 4. Four rooms 5. Five rooms 6. Six rooms 7. Seven rooms 8. Eight rooms or more	Section 04 Housing, Q3	Coded as number of rooms	Census excludes kitchen, bathroom, toilet and storeroom. Survey excludes kitchen, toilet and bathrooms. NIS indicates storeroom also excluded in Khmer version of survey questionnaire Note: Floor area is not available in the census. Check whether percentages match after amalgamating categories for both survey and census before fitting models for any final small area estimation

Table CSES B13:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

House ownership – household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Home ownership	Form B: Part 4: Housing conditions and facilities, 1.	1. Owner occupied 2. Rent 3. Not owner, but rent free 4 Other (specify)	Section 04 Housing Q24	1. Owned by the household 2. Not owned, but no rent is paid 3. Rented 4. Other	Code : 1 to 1 2 to 3 3 to 2 4 to 4

Table CSES B14:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Radio / transistor	Form B Part 4: Household conditions and facilities, 8	Total number of items	Section 09 Durable Goods 01	Total number	Question is exact match, except that census asks for radio /transistor, and survey for radio only
Television	Form B Part 4: Household conditions and facilities, 9	Total number of items	Section 09 Durable Goods 02	Total number	Question is exact match
Telephone (fixed)	Form B Part 4: Household conditions and facilities, 10	Total number of items	Section 09 Durable Goods 03	Total number	Question is exact match
Cell phone	Form B Part 4: Household conditions and facilities, 11	Total number of items	Section 09 Durable Goods 04	Total number	Question is exact match
Personal computer	Form B Part 4: Household conditions and facilities, 12	Total number of items	Section 09 Durable Goods 30	Total number	Question is exact match

Table CSES B15:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Bicycle	Form B Part 4: Household conditions and facilities, 13	Total number of items	Section 09 Durable Goods 09	Total number	Question is exact match,
Motorcycle	Form B Part 4: Household conditions and facilities, 14	Total number of items	Section 09 Durable Goods 10	Total number	Question is exact match
Car / van	Form B Part 4: Household conditions and facilities, 15	Total number of items	Section 09 Durable Goods 11 (car) and 12 (jeep/van)	Total number	Question is exact match, except that census asks for car/ van, and survey for car and jeep/van in separate categories
Boat	Form B Part 4: Household conditions and facilities, 16	Total number of items	Section 09 Durable Goods 34 (rowing boat) and 35 (motor boat)	Total number	Question is exact match, except that survey asks separately for rowing boat and motor boat

Table CSES B16:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Socio-economic Survey (CSES) 2009

Topic	***2008 Census Question***	***2008 Census Categories***	***2009 CSES Question***	***2009 CSES Categories***	***Notes***
Tractor (big tractor)	Form B Part 4: Household conditions and facilities, 17(a)	Total number of items	Section 09 Durable Goods 37	Total number	Question is exact match,
Tractor (hand tractor)	Form B Part 4: Household conditions and facilities, 17(b)	Total number of items	Section 09 Durable Goods 42	Total number	Question is exact match

CAS 2008

Table CAS A1:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Roof material – Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Roof material	Form A: Houselist, p1, column 4	1. Bamboo / thatch / grass 2. Tiles 3. Wood / plywood 4. Concrete / brick / stone 5. Galvanised iron / aluminium / other metal sheets 6. Asbestos cement sheets 7. Plastic / synthetic material sheets 8. Other (specify)	Section 04 Housing: Q5	11. No roof 12. Palm/bamboo/thatch 21Plastic sheet 22.Wood planks 31 Metal 32 Calamine / cement fiber 33 Ceramic tiles 34 Clay tiles 35 Cement 98 Other (Specify)	Some but not all categories are identical; amalgamation of categories likely to be required to match census and survey information

Equivalence Table: ***Roof Material***

Topic	***2008 Census Categories***	***2008 CAS Categories***
Roof material	1	12
	2	33, 34
	3	22
	4	35
	5	31
	6	32
	7	21
	8	98

Notes There is a “no roof” category in CAS 2008, but in discussions with NIS it was not used, as to have a house there must be a roof.

Table CAS B1:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Household size, Head of household - Household level

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Household size	Page 3, Form B: Household questionnaire Part 2, derivable as sum of number of entries (i.e. names) in column 2.	Numeric: 1, 2, 3,…..	See Household listing information carried out as a preliminary in order to select houses that contain women with children under 5 years of age. This is not a formal part of the CAS questionnaire	Numeric: 1, 2, 3,…..	Usually need to watch during modelling for household sizes in census that are very much in excess of those in the survey, especially if a 'household size squared' term is included in the survey-based model.

Table CAS B2:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Fertility - Individual level

Note: Fertility measures in the 2008 census (see Form B Household questionnaire Part 3 Fertility information of females aged 15 and over) include: number of children born alive, number of children living, and how many children have died. This can be used to determine who in the census is a mother. Number of children for each mother is available in the same part of the census. This means that characteristics of mothers (and their children) can be derived from both census and CAS 2008, in the case of the census by subsetting, and in the case of CAS 2008 because information was only collected from mothers (and their children). See CAS 2008 Section 2.

Table CAS B3:

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Marital Status– Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Marital Status	Page 3, Form B: Household questionnaire Part 2, column 6	1. Never married 2. Married (i.e. currently married) 3. Widowed 4. Divorced 5. Separated	Section 01 "Initial Visit", A, "List of household members", column 9	1. Married / living together 2. Divorced / separated 3. Widowed 4. Never married / never lived with partner	"Living together" in CAS 2008 is "never married" in census. It is unclear how to resolve this, except by having only three categories: widowed, divorced / separated, other. However then, married and not married are in the *same* category, "other". (CAS 2008 uses the same categories as CSES 2009)

Note: Marital status available in CAS 2008 only for mothers, so need to subset census data – see also notes for CAS 2008 on fertility.

Table CAS B4

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Education – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Education	Page 4, Form B: Household questionnaire Part 2, column 14 (a) and (b)	(a) Attending school 1. Never 2. Now 3. Past (b) Highest grade - Never attended 00 No class completed 01 Class 1 completed …. 12 Class 12 completed 13 Lower 2^{ndary} diploma 14 2^{ndary} school baccalaureate holder 15 Technical / vocational pre-2^{ndary} diploma 16. Technical / vocational post-2^{ndary} diploma 17. Undergraduate 18. Graduate 19. Postgrad and above 20. Other (specify)	Section 2A Q218, Q219 & Q220. Note: Q218 is whether ever attended school (Y/N). Q219 is highest level *attended* in four categories Q220 is which class (1-12) was completed	Highest grade Q218 (2) None Q219 (8) Don't know Q220 01 Class 1 completed …. Q220 12 Class 12 completed Q219 (4) *Attended* Higher [education]	Completion of Grades 1-12 are used in both census and CAS 2008 Q220 *Note: Education available in CAS 2008 only for mothers, so need to subset census data – see also notes for CAS 2008 on fertility.*

Equivalence Table: ***Education***

Topic	***2008 Census Categories***	***2008 CAS Categories***
Education	-, 00	Q218 (2)
	01	Q220 01
	02	Q220 02
	03	Q220 03
	04	Q220 04
	05	Q220 05
	06	Q220 06
	07	Q220 07
	08	Q220 08
	09	Q220 09
	10	Q220 10
	11	Q220 11
	12	Q220 12
	13	Q219 (2)
	14	Q219 (3)
	15	N/A
	16	N/A
	17	Q219 (4)
	18	Q219 (4)
	19	Q219 (4)
	20	N/A

Note: CAS 2008 categories 15 & 16 related to technical education and 20 other, do not code to the education categories in CAS 2008

Table CAS B5

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Employment – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Employment	Page 4, Form B: Household questionnaire Part 2, column 16-23. Questions asked are: Main activity Employment period Occupation Employment status Industry, trade or service Sector 2ndary economic activity Place of work or school Proportion of year employed - B(17)	See census questionnaire page 4, Form B	Q 227 Mother's occupation	Same categories as used for occupation in census (Information from NIS 1 June 2010)	

Note: Employment information only available for mothers in CAS 2008

Table CAS B6

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Employment – Individual level

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Antenatal care / delivery	Page 4, Form B: Household questionnaire Part 3, Fertility (8)	1. Doctor 2. Nurse 3. Midwife 4 Traditional birth attendant (TBA) 5 Other 6 None	Q 231 Antenatal care during last pregnancy	1. Doctor / medical assistant 2. Nurse 3. Midwife 4 Traditional birth attendant (TBA) 5 Other 6 None)	

Note: The question in the census is about delivery. The question in CAS 2008 is about antenatal care. Whether these are sufficiently similar will need checking via percentages in each category when comparing survey and census. Joel Conkle at UNICEF Cambodia has commented that they are more so nor there s a USD15 incentive paid to health clinics if delivery is supervised there, but this incentive began in 2009 and until then many deliveries were not supervised in clinics, even for mothers who had received antenatal care.

Table CAS B7

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric Survey (CAS) 2008

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Electricity	Form B, Part 4, Household conditions and facilities: any of 9, 10, 12, 18, 19	Total number of items >=1 when summed over the household items	Section 1 Q107 – 1st subquestion		Using any of these items requires electricity
Radio / transistor	Form B Part 4: Household conditions and facilities, 8	Total number of items	Section 1 Q107 – 2nd subquestion	Total number	Question is exact match, except that census asks for radio /transistor, and survey for radio only
Television	Form B Part 4: Household conditions and facilities, 9	Total number of items	Section 1 Q107 – 3rd subquestion	Total number	Question is exact match
Cell phone	Form B Part 4: Household conditions and facilities, 11	Total number of items	Section 1 Q107 – 4th subquestion	Total number	Question is exact match

Note: All available at household level in both census and CAS 2008

Table CAS B8

Links between Cambodian Census of Population and Housing 2008 and Cambodian Anthropometric (CAS) 2008

Topic	***2008 Census Question***	***2008 Census Categories***	***2008 CAS Question***	***2008 CAS Categories***	***Notes***
Bicycle	Form B Part 4: Household conditions and facilities, 13	Total number of items	Section 1 Q109 – 1st subquestion	Total number	Question is exact match, except that CAS 2008 includes cyclo
Motorcycle	Form B Part 4: Household conditions and facilities, 14	Total number of items	Section 1 Q109 – 2nd subquestion	Total number	Question in CAS 2008 includes motorcycle, moped and scooter (which given types of vehicles in Cambodia are essentially the same)
Car / van	Form B Part 4: Household conditions and facilities, 15	Total number of items	Section 1 Q109 – 3rd subquestion	Total number	Question is exact match, except that census asks for car/ van, and survey for car /truck/van in English version of questionnaire, but identical in Khmer
Boat	Form B Part 4: Household conditions and facilities, 16	Total number of items	Section 1 Q109 – 3rd subquestion 4 & 5	Total number	Question is exact match, except that survey asks separately for boat with motor and boat without motor

Note: All available at household level in both census and CAS 2008

Appendix B

Table B1. Explanatory variables from 2008 CAS

Variable Name	Variable Label
ageyr1 *	Child's age in years=0
ageyr2	Child's age in years=1
ageyr3	Child's age in years=2
ageyr4	Child's age in years=3
ageyr5	Child's age in years=4
natcare_doc	1 if antenatal care(for last pregnancy) is a Doctor/Medical Assistant
natcare_nurs	1 if antenatal care(for last pregnancy) is a Nurse
natcare_midwif	1 if antenatal care(for last pregnancy) is a Midwife
natcare_trad	1 if antenatal care(for last pregnancy) is a Traditional Birth Attendant
natcare_other	1 if antenatal care(for last pregnancy) is Other
natcare_no*	1 if did not see anyone for antenatal care(for last pregnancy)
school	1 if mother attended school
educ_primary	1 if the mother's highest level of education is primary
educ_low_second	1 if the mother's highest level of education is lower secondary
educ_uper_second	1 if the mother's highest level of education is upper secondary
educ_higher	1 if the mother's highest level of education is higher
w_spouse	1 if the mother is married or has a partner
electric	1 if household has electricity
radio	1 if household has radio
tv	1 if household has television
celphon	1 if household has mobile telephone
bike	1 if one of the household members owns a bicycle
motor	1 if one of the household members owns a motorcycle/scooter
car	1 if one of the household members owns a car/truck/van
boat	1 if one of the household members owns a boat (with or without motor)
roof_natural	1 if the house has natural roofing (palm/bamboo/thatch/no roof)
roof_finished	1 if the house has finished roofing (metal/cement/ceramic/clay tiles)
roof_rudim*	1 if the house has rudimentary roofing (plastic sheet/wood planks)
sub_agri	1 if the mother is in subsistence agricultural, fishery and related work
rural	1 if the child is living in the rural area
girl	1 if the child is a girl

* indicates baseline category

Table B2. Explanatory variables from 2009 CSES

Variable	Variable Label
ageyr1*	Child's age in years=0
ageyr2	Child's age in years=1
ageyr3	Child's age in years=2
ageyr4	Child's age in years=3
ageyr5	Child's age in years=4
hhsize	household size
hhsq	(hhsize-5)^2
prop_emp	Proportion of family members who are actively employed
crop_farmer	1 if household head a subsistence crop farmer
head_male	1 if household head is male
employed	1 if household head is employed
read	1 if the household head can read a simple message in any language
write	1 if the household head can write a simple message in any language
school	1 if household head attended school
sch_yr	Number of years household head attended school
educ_level	Highest education attainment for household head
numroom	number of rooms in a dwelling unit
nroom2	(numroom-1)^2
head_migrant	1 if household head is a migrant
head_married	1 if household head is a married
wall_bamboo	1 if primary construction material for the wall is bamboo
wall_wood	1 if primary construction material for the wall is wood
wall_concrete	1 if primary construction material for the wall is concrete
wall_metal	1 if primary construction material for the wall is metal
wall_fibcement	1 if primary construction material for the wall is fibrous cement/Asbestos
wall_mixed	1 if primary construction material for the wall is a mixture of various materials
wall_clay	1 if primary construction material for the wall is clay/dung with straw
wall_oth*	1 if primary construction material for the wall is others
roof_bamboo	1 if primary construction material for the roof is bamboo
roof_tiles	1 if primary construction material for the roof is tiles
roof_fibcement	1 if primary construction material for the roof is fibrous cement
roof_metal	1 if primary construction material for the roof is metal
roof_oth*	1 if primary construction material for the roof is others
roof_concrete	1 if primary construction material for the roof is concrete
roof_plastic	1 if primary construction material for the roof is plastic
floor_clay	1 if primary construction material for the floor is clay
floor_wood	1 if primary construction material for the floor is wood
floor_cement	1 if primary construction material for the floor is cement/brick/stone
floor_parquet	1 if primary construction material for the floor is parquet/polished wood
floor_stone	1 if primary construction material for the floor is polished stone
floor_oth*	1 if primary construction material for the floor is others
floor_tiles	1 if primary construction material for the roof is ceramic tiles
own_acctwrkr	1 if employment status in main occupation is own account worker
toilet_none	1 if household does not have a toilet
toilet_sewerage	1 if household has a toilet that is connected to a sewerage system
toilet_septic	1 if household has a toilet connected to a septic tank

(Contd overleaf)

Table B2. Explanatory Variable from 2009 CSES continued…

Variable	Variable Label
toilet_pit	1 if household uses a pit with or without slab for toilet
toilet_oth*	1 if household uses other types of toilet
drink_public	1 if drinking water source is public tap
drink_tubed	1 if drinking water source is tubed/piped well or borehole
drink_protwell	1 if drinking water source is protected dug well
drink_unpropwel	1 if drinking water source is unprotected dug well
drink_rainwater	1 if drinking water source is rain water
drink_pond	1 if drinking water source is pond, river or stream
drink_tankr*	1 if drinking water source is tanker truck or vendor
drink_oth	1 if drinking water source is others.
firewood	1 if household uses firewood for cooking
charcoal	1 if household uses charcoal for cooking
LPG	1 if household uses LPG for cooking
kerosene	1 if household uses kerosene for cooking
fuel_electric	1 if household uses electricity for cooking
fuel_none*	1 if household don't cook
fuel_oth	1 if household uses other kinds of fuel for cooking
light_electric	1 if household's main source of lighting is electricity
generator	1 if household's main source of lighting is a generator
light_batry	1 if household's main source of lighting is a battery
light_kerosene	1 if household's main source of lighting is kerosene
light_candle*	1 if household's main source of lighting is candle
light_oth	1 if household's main source of lighting is others
bike_n	Number of bicycles
tv_n	Number of television sets
tract_s_n	Number of small tractors (hand tractor)
tract_b_n	Number of big tractors (tractor)
motor_n	Number of motorcycles
tel_n	Number of telephones
comp_n	Number of computers
cell_n	Number of cell phones
car_van_n	number of car/van
boat_n	number of boats
house_owned	1 if dwelling's legal status is owned
house_rented	1 if dwelling's legal status is rented
house_oth	1 if dwelling's legal status is others
girl_child	1 if child is a girl
rural	Rural area
p_primary	Proportion of household members with primary education
p_secondary	Proportion of household members with secondary education
p_elem	Proportion of household members with elementary education
p_postsecond	Proportion of family members with postsecondary education
pkids6	Proportion of household members age less than 7
pkids714	Proportion of household members age 7-14
pmem_1565	Proportion of household members age 15-65
pmem_senior	Proportion of household members age over 65

* indicates baseline category

Appendix C

Table C1. Model for *ZH* (standardized height-for-age) from CAS2008

Variable	Coef.	Std. Err.	t	P>t
girl	0.1381	0.0369	3.74	0.000
ageyr2	-0.9849	0.0590	-16.68	0.000
ageyr3	-1.1223	0.1414	-7.94	0.000
ageyr4	-1.1146	0.1285	-8.67	0.000
ageyr5	-1.2904	0.1338	-9.65	0.000
rural	-1.4517	0.4937	-2.94	0.003
sub_agri	-0.4940	0.3671	-1.35	0.178
w_spouse	-0.1636	0.0955	-1.71	0.087
natcare_doc	-0.1591	0.0617	-2.58	0.010
natcare_nurs	-0.1401	0.1013	-1.38	0.167
educ_primary	1.7548	0.0641	27.39	0.000
educ_low_second	1.9199	0.0838	22.92	0.000
educ_uper_ second	2.1231	0.1192	17.81	0.000
educ_higher	2.4632	0.3906	6.31	0.000
roof_natural	-1.0629	0.3785	-2.81	0.005
roof_finished	-0.7764	0.3481	-2.23	0.026
school	-1.6302	0.0740	-22.03	0.000
electric	-0.4047	0.1676	-2.42	0.016
radio	0.3055	0.1060	2.88	0.004
tv	0.3592	0.1545	2.32	0.020
celphon	0.1805	0.0467	3.87	0.000
motor	0.1375	0.0440	3.12	0.002
car	0.2748	0.1743	1.58	0.115
ageyr3XR	-0.2471	0.1466	-1.69	0.092
ageyr4XR	-0.4027	0.1346	-2.99	0.003
ageyr5XR	-0.2889	0.1418	-2.04	0.042
sub_agriXR	0.5160	0.3694	1.40	0.163
natcare_nursXR	0.1778	0.1099	1.62	0.106
roof_naturalXR	1.5119	0.5347	2.83	0.005
roof_finishedXR	1.3522	0.5130	2.64	0.008
electricXR	0.5580	0.1794	3.11	0.002
radioXR	-0.2000	0.1137	-1.76	0.079
tvXR	-0.2969	0.1605	-1.85	0.064
carXR	-0.6584	0.2441	-2.70	0.007
_cons	0.0665	0.3339	0.20	0.842

Table C2. Model for *ZW* (standardized weight-for-age) from CAS2008

Variable	Coef.	Std. Err.	t	P>t
girl	0.1518	0.0764	1.99	0.047
ageyr2	-0.6093	0.0449	-13.58	0.000
ageyr3	-0.8095	0.0447	-18.11	0.000
ageyr4	-0.9096	0.0457	-19.89	0.000
ageyr5	-1.0110	0.0468	-21.59	0.000
sub_agri	-0.7497	0.2135	-3.51	0.000
natcare_nurs	-0.1047	0.0767	-1.36	0.172
educ_primary	2.0568	0.0446	46.12	0.000
educ_low_second	2.1380	0.0593	36.03	0.000
educ_uper_second	2.4760	0.1471	16.83	0.000
educ_higher	2.1236	0.1872	11.35	0.000
school	-1.9787	0.0501	-39.50	0.000
electric	-0.2471	0.1636	-1.51	0.131
radio	0.0705	0.0323	2.18	0.029
tv	0.2492	0.1281	1.94	0.052
motor	0.3922	0.1009	3.88	0.000
car	0.4650	0.1122	4.15	0.000
girlXR	-0.1196	0.0823	-1.45	0.146
w_spouseXR	-0.1718	0.0753	-2.28	0.023
sub_agriXR	0.7790	0.2155	3.62	0.000
educ_uper_ second XR	-0.3348	0.1657	-2.02	0.043
natcare_nursXR	0.1491	0.0828	1.80	0.072
carXR	-0.5414	0.1601	-3.38	0.001
electricXR	0.3762	0.1713	2.20	0.028
motorXR	-0.3114	0.1064	-2.93	0.003
tvXR	-0.2119	0.1323	-1.60	0.109
celphonXR	0.1204	0.0411	2.93	0.003
_cons	-0.8624	0.0844	-10.22	0.000

Table C3. Model for *ZWH* (standardized weight-for-height) from CAS2008

Variable	Coef.	Std. Err.	t	P>t
girl	0.2034	0.0814	2.50	0.013
ageyr3	-0.2270	0.0485	-4.68	0.000
ageyr4	-0.2066	0.0484	-4.27	0.000
ageyr5	-0.2281	0.0494	-4.61	0.000
rural	0.3187	0.1041	3.06	0.002
sub_agri	-0.6065	0.1825	-3.32	0.001
natcare_doc	0.2639	0.1083	2.44	0.015
natcare_midwif	0.2055	0.0831	2.47	0.013
educ_primary	1.5323	0.0500	30.63	0.000
educ_low_second	1.5242	0.0648	23.53	0.000
educ_uper_second	1.6110	0.1132	14.24	0.000
educ_higher	1.0562	0.1809	5.84	0.000
school	-1.5189	0.0572	-26.57	0.000
motor	0.2938	0.0973	3.02	0.003
car	0.4600	0.1297	3.55	0.000
celphon	-0.1736	0.0917	-1.89	0.058
girlXR	-0.2096	0.0883	-2.37	0.018
ageyr2XR	-0.3653	0.0511	-7.15	0.000
sub_agriXR	0.6299	0.1861	3.39	0.001
natcare_docXR	-0.2771	0.1263	-2.19	0.028
natcare_midwifXR	-0.2900	0.0900	-3.22	0.001
roof_finishedR	-0.0597	0.0398	-1.50	0.133
motorXR	-0.2940	0.1044	-2.82	0.005
carXR	-0.2323	0.1794	-1.30	0.195
celphonXR	0.1917	0.0997	1.92	0.055
_cons	-0.7387	0.1003	-7.37	0.000

Table C4. Model for *ZH* (standardized height-for-age) from CSES2009

Variable	Coef.	Std. Err.	t	P>t
ageyr2	-0.2876	0.2165	-1.33	0.184
ageyr3	-0.5208	0.2074	-2.51	0.012
ageyr4	-0.9950	0.0885	-11.25	0.000
ageyr5	-0.9834	0.0858	-11.47	0.000
hhsize	0.0398	0.0260	1.53	0.126
hhsq	-0.0089	0.0052	-1.71	0.087
pkids714	-1.3827	0.5363	-2.58	0.010
p_elem	0.9826	0.3162	3.11	0.002
p_secondary	1.8014	0.8950	2.01	0.044
p_postsecond	2.2113	1.0813	2.04	0.041
numroom	-0.1609	0.0677	-2.38	0.017
floor_stone	-1.6007	0.2580	-6.20	0.000
wall_bamboo	-0.5304	0.2336	-2.27	0.023
wall_concr~e	0.5801	0.1777	3.26	0.001
wall_mixed	0.4379	0.3388	1.29	0.196
roof_plastic	1.0693	0.1345	7.95	0.000
toilet_none	0.5930	0.2387	2.48	0.013
toilet_pit	-0.7506	0.3378	-2.22	0.026
toilet_sewerage	-0.4996	0.1821	-2.74	0.006
drink_oth	-1.0612	0.7849	-1.35	0.176
drink_rainwater	-0.2508	0.0772	-3.25	0.001
light_batry	-0.7468	0.3102	-2.41	0.016
light_kerose	-0.8228	0.2847	-2.89	0.004
light_oth	-0.4891	0.2073	-2.36	0.018
generator	0.3163	0.2289	1.38	0.167
charcoal	0.5193	0.1739	2.99	0.003
tract_b_n	3.9444	1.2130	3.25	0.001
car_van_n	-0.3230	0.1755	-1.84	0.066
ageyr2XR	-0.3743	0.2250	-1.66	0.096
ageyr3XR	-0.3036	0.2142	-1.42	0.156
head_ageXR	0.0054	0.0029	1.86	0.063
head_migrantXR	-0.0840	0.0619	-1.36	0.175
crop_farmedR	0.1375	0.0811	1.70	0.090
house_rentedXR	1.5741	0.8611	1.83	0.068
own_acctwrkrXR	-0.1486	0.0995	-1.49	0.136
numroomXR	0.1332	0.0816	1.63	0.103
pkids714XR	1.1400	0.5678	2.01	0.045
p_elemXR	-0.9384	0.3485	-2.69	0.007

Variable	Coef.	Std. Err.	t	P>t
p_postsecondXR	-3.4324	1.5669	-2.19	0.029
p_secondaryXR	-1.9198	1.0628	-1.81	0.071
readXR	0.1484	0.0799	1.86	0.063
wall_bambooXR	0.5020	0.2413	2.08	0.038
wall_concreteXR	-0.7736	0.2331	-3.32	0.001
wall_metalXR	0.1991	0.1314	1.52	0.130
toilet_non~R	-0.6225	0.2501	-2.49	0.013
toilet_pitXR	0.7294	0.3834	1.90	0.057
toilet_sewerageXR	0.8381	0.3076	2.72	0.006
light_batryXR	0.9313	0.3243	2.87	0.004
light_keroseneXR	0.8936	0.3015	2.96	0.003
drink_tubedXR	-0.1599	0.0811	-1.97	0.049
drink_unpropwelXR	-0.1919	0.0950	-2.02	0.043
LPGXR	0.5191	0.2821	1.84	0.066
charcoalXR	-0.7443	0.2374	-3.13	0.002
comp_nXR	-0.6542	0.4027	-1.62	0.104
car_van_nXR	0.6619	0.2788	2.37	0.018
cell_nXR	0.0892	0.0528	1.69	0.091
tract_b_nXR	-3.9899	1.2659	-3.15	0.002
_cons	-1.6185	0.2002	-8.08	0.000

Table C5. Model for *ZW* (standardized weight-for-age) from CSES2009

Variable	Coef.	Std. Err.	t	P>t
girl_child	0.0708	0.0364	1.94	0.052
ageyr2	-0.4551	0.0642	-7.09	0.000
ageyr3	-0.7635	0.0614	-12.44	0.000
ageyr4	-1.0976	0.1096	-10.01	0.000
ageyr5	-0.9819	0.0604	-16.25	0.000
rural	-0.9311	0.4061	-2.29	0.022
crop_farmer	-0.0811	0.0519	-1.56	0.118
own_acctwrkr	-0.2374	0.0981	-2.42	0.016
pmem_1565	0.9518	0.3451	2.76	0.006
p_elem	0.3273	0.0926	3.54	0.000
p_postsecond	0.7695	0.5115	1.50	0.133
numroom	-0.1910	0.0766	-2.49	0.013
nroom2	0.0224	0.0120	1.86	0.062
floor_clay	-0.5727	0.3407	-1.68	0.093
floor_wood	-0.8306	0.3483	-2.38	0.017
floor_cement	-0.7263	0.3384	-2.15	0.032
floor_parquet	-0.9312	0.3860	-2.41	0.016
floor_stone	-0.7078	0.3617	-1.96	0.050
floor_tiles	-0.7132	0.3499	-2.04	0.042
wall_concrete	0.2464	0.1414	1.74	0.082
roof_tiles	-0.0472	0.0479	-0.98	0.325
roof_fibcecement	-0.1074	0.0724	-1.48	0.138
roof_plastic	0.6975	0.1002	6.96	0.000
toilet_none	0.2890	0.1526	1.89	0.058
drink_unpropwell	-0.0961	0.0603	-1.59	0.111
drink_rainwater	-0.0834	0.0512	-1.63	0.103
drink_tubed	-0.1015	0.0508	-2.00	0.046
drink_oth	-0.5628	0.4023	-1.40	0.162
light_kerosene	-0.6933	0.3939	-1.76	0.078
light_batry	-0.9691	0.3936	-2.46	0.014
light_electricc	-0.6407	0.3551	-1.80	0.071
light_oth	-1.3627	0.6463	-2.11	0.035
generator	0.3615	0.2884	1.25	0.210
tv_n	0.2345	0.0786	2.98	0.003
boat_n	-0.5100	0.2192	-2.33	0.020
ageyr4XR	0.2166	0.1119	1.94	0.053
own_acctwrkrXR	0.2083	0.1144	1.82	0.069
house_rentedXR	0.6876	0.2974	2.31	0.021

Variable	Coef.	Std. Err.	t	P>t
pmem_1565XR	-0.9071	0.3728	-2.43	0.015
numroomXR	0.1558	0.0895	1.74	0.082
nroom2XR	-0.0224	0.0131	-1.71	0.088
floor_woodXR	0.3098	0.1374	2.26	0.024
wall_concreteXR	-0.1976	0.1719	-1.15	0.251
toilet_noneXR	-0.3083	0.1610	-1.92	0.056
light_keroseneXR	1.0185	0.3932	2.59	0.010
light_electricXR	0.9862	0.3556	2.77	0.006
light_batryXR	1.3356	0.3908	3.42	0.001
light_othXR	1.3226	0.6533	2.02	0.043
charcoalXR	-0.1739	0.1084	-1.60	0.109
car_van_nXR	0.2809	0.1612	1.74	0.082
tv_nXR	-0.2195	0.0903	-2.43	0.015
cell_nXR	0.1040	0.0353	2.94	0.003
comp_nXR	-0.6970	0.2177	-3.20	0.001
boat_nXR	0.3777	0.2276	1.66	0.097
_cons	0.3788	0.5163	0.73	0.463

Table C6. Model for *ZHW* (standardized height-for-weight) from CSES2009

Variable	Coef.	Std. Err.	t	P>t
girl_child	0.0957	0.0503	1.90	0.057
ageyr2	-0.5802	0.0864	-6.71	0.000
ageyr3	-0.7414	0.0840	-8.83	0.000
ageyr4	-0.9442	0.1705	-5.54	0.000
ageyr5	-0.7450	0.0821	-9.07	0.000
rural	-1.9193	0.4835	-3.97	0.000
own_acctwrkr	-0.3205	0.1372	-2.34	0.020
p_elem	0.3462	0.1240	2.79	0.005
p_postsecond	1.0016	0.6679	1.50	0.134
house_owned	0.1931	0.1174	1.65	0.100
house_rented	0.3676	0.2267	1.62	0.105
floor_cement	-0.4445	0.2452	-1.81	0.070
floor_parq~t	-0.6230	0.2909	-2.14	0.032
floor_stone	0.8674	0.2690	3.23	0.001
floor_tiles	-0.3676	0.1963	-1.87	0.061
floor_wood	-0.4043	0.2240	-1.80	0.071
wall_bamboo	0.4922	0.2398	2.05	0.040
wall_mixed	-0.5188	0.3000	-1.73	0.084
roof_fibcement	-0.2608	0.0970	-2.69	0.007
roof_tiles	-0.1509	0.0694	-2.17	0.030
toilet_sewerage	0.3110	0.1721	1.81	0.071
drink_tubed	-0.3707	0.2206	-1.68	0.093
drink_unpropwel	-0.4590	0.2125	-2.16	0.031
light_batry	-1.5436	0.4622	-3.34	0.001
light_electricc	-1.7828	0.3956	-4.51	0.000
light_kerosee	-1.1096	0.4317	-2.57	0.010
light_oth	-2.5105	0.3960	-6.34	0.000
generator	-1.4370	0.8936	-1.61	0.108
charcoal	-0.3078	0.1405	-2.19	0.028
tv_n	0.3619	0.1122	3.23	0.001
boat_n	-0.6142	0.2927	-2.10	0.036
car_van_n	0.1802	0.1034	1.74	0.082
comp_n	-0.2321	0.1647	-1.41	0.159
tract_b_n	-3.3224	0.7708	-4.31	0.000
ageyr4XR	0.2872	0.1739	1.65	0.099
head_ageXR	-0.0036	0.0025	-1.43	0.154
head_migrantXR	0.1371	0.0572	2.40	0.017
crop_farmerXR	-0.2386	0.0721	-3.31	0.001

Variable	Coef.	Std. Err.	t	P>t
own_acctwrkrXR	0.4038	0.1651	2.45	0.014
floor_cementXR	0.3624	0.2612	1.39	0.165
floor_woodXR	0.5321	0.2293	2.32	0.020
wall_bambooXR	-0.8910	0.2777	-3.21	0.001
wall_woodXR	-0.2897	0.1408	-2.06	0.040
wall_metalXR	-0.4915	0.1701	-2.89	0.004
toilet_sewerageXR	-0.3547	0.2764	-1.28	0.199
drink_tubewelXR	0.3326	0.2304	1.44	0.149
drink_unpropwelXR	0.4511	0.2265	1.99	0.046
light_electricXR	2.3324	0.4921	4.74	0.000
light_batryXR	1.9785	0.5419	3.65	0.000
light_keroseneXR	1.5839	0.5168	3.06	0.002
light_othXR	2.8953	0.5148	5.62	0.000
firewoodXR	-0.2361	0.1799	-1.31	0.189
generatorXR	1.9075	0.9502	2.01	0.045
LPGXR	-0.6516	0.2856	-2.28	0.023
tract_b_nXR	3.4654	0.8316	4.17	0.000
boat_nXR	0.5305	0.3007	1.76	0.078
tv_nXR	-0.3671	0.1261	-2.91	0.004
_cons	2.0395	0.3848	5.30	0.000