

Using Poverty Maps to Improve the Design of Household Surveys

The Evidence from Tunisia

Gianni Betti
Vasco Molini
Dan Pavelesku



WORLD BANK GROUP

Poverty and Equity Global Practice

May 2021

Abstract

This paper proposes a new method for improving the design effect of household surveys based on a two-stage design in which the first stage clusters, or primary selection units, are stratified along administrative boundaries. Improvement of the design effect can result in more precise survey estimates (smaller standard errors and confidence intervals) or reduction of the necessary sample size, that is, a reduction in the budget needed for a survey. The proposed method is based on the availability of a previously conducted poverty mapping, that is, spatial descriptions of the distribution of poverty, which are finely disaggregated in small geographic units, such as cities, municipalities, districts, or

other administrative partitions of a country that are linked to primary selection units. Such information is then used to select primary selection units with systematic sampling by introducing further implicit stratification in the survey design, to maximize the improvement of the design effect. The proposed methodology has been implemented for the new 2021 Household Budget Survey in Tunisia, conducted under a cooperation project funded by the World Bank. The underlying poverty mapping is based on the 2015 Household Budget Survey and the 2014 Population and Housing Census.

This paper is a product of the Poverty and Equity Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/prwp>. The authors may be contacted at vmolini@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Using Poverty Maps to Improve the Design of Household Surveys: The Evidence from Tunisia

Gianni Betti

University of Siena
Siena, Italy
gianni.betti@unisi.it

Vasco Molini¹

The World Bank
Rabat, Morocco
vmolini@worldbank.org

Dan Pavelesku

The World Bank
Algiers, Algeria
dpavelesku@worldbank.org

JEL: C83, D63, C81.

Keywords: Survey design, implicit stratification, poverty map

¹ Corresponding author

1. Introduction

To establish trust in the measurement of poverty and inequality, statistical offices must regularly assess and continuously improve the quality of their processes as well as the accuracy of their data. This should be one of the prerequisites to prevent misguided policies, especially when such policies are based on data from household budget surveys.

Over the past decades, national statistical offices worldwide have started to develop reports describing the quality criteria adopted in the surveys and explaining any instances in which these criteria could not be met. One common method for reporting the precision of poverty measures is to include sampling variability or standard errors and, consequently, the confidence intervals of such estimates.

This practice is needed much more when poverty measures are disaggregated at smaller territorial levels or for some sub-groups of the population. Recently, the United Nations Economic Commission for Europe (UNECE) has addressed these issues in a series of conferences resulting in a useful guide for data disaggregation (UNECE, 2020).

While in the past many national sample surveys were conducted with the aim of getting estimates that were significant at national level, only nowadays it is becoming more and more important to first define the exact type of *domains* for which national statistical offices intend to calculate, and then publish poverty measures. This is necessary because when such measures were disaggregated at lower territorial levels or for other *domains*, often such estimates ended up being non-significant according to standard errors calculated *ex-post*. The *ex-ante* definition of sub-sample sizes in particular and sample designs in general, therefore, becomes crucial to have statistically significant estimates.

There is obviously a trade-off between more precise *domain* estimates and the total cost of a sample survey. For this reason, several statistical offices have a strong incentive to plan and implement highly efficient sampling designs. Indeed, an efficient sample design is a prerequisite for any sample survey, and it plays a fundamental role in household budget surveys. As a matter of fact, improvements

in the design (called ‘design effect’ or simply *deft*) can result in more precise survey estimates (smaller standard errors and confidence intervals) or in the reduction of the necessary sample size, therefore, a reduction in the budget needed for the surveys.

This paper aims to propose a new method for increasing the efficiency of two-stage household sample surveys by incorporating information from previous small area estimates of poverty in the selection process of Primary Selection Units (PSUs). Poverty estimates in small areas are spatial descriptions of the distribution of poverty that is finely disaggregated in small geographic units, such as cities, municipalities, districts, or other administrative partitions within a country.

In the context of improving the sample design in a multi-stage scheme, these poverty maps should satisfy at least two key properties: i) the small area should be directly linked to the PSUs of the design; ii) standard errors (precision) of such small area estimates should be relatively small and they should be easily calculated by the National Statistical Office conducting the household survey.

For the above reasons, the methodological originality of this contribution is twofold. First, the new method takes into account the theory of small area estimation (SAE, also referred as poverty mapping) methods developed in the context of the World Bank by Elbers, Lanjouw and Lanjouw (ELL, 2002 & 2003), later refined by Tarozzi and Deaton (2009) and Van der Weide (2014); this for a notable improvement in the design effect of household surveys and specifically for Household Budget Surveys (HBSs). The proposed method, however, is very versatile and can be implemented also using alternative approaches to poverty estimation at the ‘small area level’ (as defined in Rao, 2003): in particular, the Empirical Best (EB) of Molina and Rao (2010), and the more recent proposal H3-CensusEB of Molina (2019). The present paper is based on Molina’s H3-CensusEB estimator found to be superior in terms of bias and mean square error (MSE) to ELL and Molina and Rao’s Empirical Bayes prediction estimates (Corral, Molina and Nguyen, 2020; Corral et al., 2021).

Second, the new method described in this paper implements, for the first time in a developing country, the recent JRR method, from Verma and Betti (2011), for

estimating the *ex-post* standard errors of poverty and inequality measures from HBSs, that is, *complex measures* from *complex surveys*. In this way, it allows the evaluation of the impact of the so-called *implicit stratification* (or *hidden*; Wolter, 1985), which is present when PSUs are selected with systematic sampling from one or more lists that are ordered according to some variables correlated with the phenomenon under investigation.

The present paper is composed of five sections. Following this introduction, Section 2 presents the theory behind the estimation of standard errors in complex surveys. Attention is paid to the two-stage stratified sampling design, where, in the first stage, selection of PSUs is performed with systematic sampling, i.e., where implicit stratification is present. Since systematic sampling can be seen as a random selection of a single PSU from each implicit stratum, it is not possible to derive a consistent estimator of the sampling variance (Wolter, 1985). Therefore, recent replication methods have been developed and implemented for estimating the variance of complex measures from complex surveys. Among the most adopted methodologies are bootstrap, Taylor Linearisation, Jack-knife Repeated Replications (JRR) and Balanced Repeated Replications (BRR). On this basis, the second part of Section 2 briefly presents the recent methodology of Verma and Betti (2011) for the implementation of a variant of JRR in the case of systematic selection of PSUs and the possibility of decomposition of the design effect in several components (Kish, 1965; Verma, 1991; Liu *et al.*, 2002).

The definition of these components is useful for highlighting the proposed implicit stratification. As previously mentioned, we aim to introduce a further implicit stratification by selecting the PSUs with a systematic sampling within each explicit stratum. PSUs are ordered by their relative ranking in terms of the per capita consumption expenditure obtained from the poverty mapping approach. Therefore, Section 3 is devoted to a brief presentation of the World Bank methodology.

Since, in general, we believe that any new methodological proposal is useful if it can be put into practice, the second part of Section 3 presents the implementation of poverty mapping in Tunisia, based on the 2014 Census and the 2015 Household

Budget Survey (EBCNV). Moreover, Section 4 is devoted to the presentation of the new sample design in the case of the 2021 Household Budget Survey in Tunisia. Three scenarios are taken into account: i) PSUs are selected with Simple Random Sampling (SRS) within the 24 governorates (explicit strata) in Tunisia; ii) PSUs are, first, stratified among the 24 governorates and then further stratified (implicitly) with systematic sampling by ordering the lists according to some economic and geographical variables, which consequently make the list a **geographical serpentine**; iii) PSUs are, first, stratified among the 24 governorates, and then further stratified (implicitly) with systematic sampling by ordering the lists of PSUs according to the average per-capita consumption expenditure estimated by the poverty mapping method presented in Section 3.

This third scenario represents the primary, original contribution of this paper, and the design effect gain is compared with other scenarios. Such gain in the design effect has an impact on both the estimate variance (for higher precision in the estimates), and in the optimal sample size (i.e., for a smaller budget).

Finally, Section 5 concludes the paper. After summarizing the main results of the paper, the section aims to briefly present some variants of the proposed methodology when Census and/or Household Budget Surveys are not available for recent years in order to perform the poverty mapping exercise.

2. Theory of design effects, variance estimation, and resampling techniques

This section aims to present the theory behind the estimation of standard errors in complex surveys such as HBSs. Since the *implicit stratification* proposed is based on poverty estimates and per capita consumption expenditure performed with the World Bank ELL method (described in further detail in Section 3), here the notation adopted is a mixture of the notation for sampling techniques in Wolter (1985) and the notation for poverty mapping in ELL (2003), as follows.

We have a *population* divided in L strata ($l: 1, 2, \dots, L$) and clustered in M_l clusters ($m: 1, 2, \dots, M_l$); in each cluster m there are K_m households ($k: 1, 2, \dots, K_m$) where

HS_k individuals live (j: 1, 2, ..., HS_k), and where N_l is the number of households in stratum l. Population size N can therefore be defined as:

$$N = \sum_{l:1}^L N_l = \sum_{m:1}^M K_m$$

From this population, and for each stratum l (l: 1, 2, ...L) we select a *sample* of C_l clusters (Primary Selection Units, PSU; c: 1, 2, ..., C_l); within each PSU c we then select h_c households (h: 1, 2, ..., h_c), where the number of households in each stratum is denoted by n_l. Then, sample size n can, therefore, be defined as:

$$n = \sum_{l:1}^L n_l = \sum_{c:1}^M h_c$$

Taking into account a general two-stage sampling design where, in the first stage, clusters (PSUs) are stratified in the L strata and are drawn with equal probability; and where, in the second stage, households are drawn with equal probability, then we can define the estimator of the total (Wolter, 1985 p. 13): $\hat{Y} = \sum_{l:1}^L \hat{Y}_l$, where

$$\hat{Y}_l = (N_l/n_l) \sum_{c:1}^{C_l} K_c \sum_{h:1}^{H_c} y_{ch} / h_c \quad (1)$$

Since the L strata are disjoint sub-sets of the population (they constitute a *partition*), then they are independent of each other, and, therefore, the (true) variance of estimator (1) is:

$$\begin{aligned} \text{Var}\{\hat{Y}\} &= \sum_{l:1}^L \text{Var}\{\hat{Y}_l\}, \text{ where} \\ \text{Var}\{\hat{Y}_l\} &= N_l^2 (1 - f_{1l}) \left(\frac{1}{n_l} \right) \left[\sum_{m:1}^{M_l} \frac{(Y_m - Y_l/N_l)^2}{(N_l - 1)} \right] + \\ &(N_l^2/n_l) \left[\sum_{m:1}^{M_l} K_m^2 (1 - f_{2m}) (S_m^2/h_m) / N_l \right] \quad (2) \end{aligned}$$

and where

$$\begin{aligned} Y_m &= \sum_{k:1}^{K_m} Y_{mk}, \quad S_m^2 = \sum_{k:1}^{K_m} (Y_{mk} - \hat{Y}_m)^2 / (K_m - 1), \quad \hat{Y}_m = \sum_{k:1}^{K_m} Y_{mk} / K_m \\ f_{1l} &= n_l / N_l, \quad f_{2m} = h_m / K_m. \end{aligned}$$

Equation (2) could be rewritten in terms of variance *within* clusters S_{2l}^2 and *between* clusters S_{1l}^2 in a certain stratum l, as follows:

$$Var\{\hat{Y}_l\} = N_l^2(1 - f_{1l})(1/n_l)S_{1l}^2 + (N_l^2/n_l)S_{2l}^2 \quad (3).$$

Equation (3) will be the starting point for attempting to improve a household survey design.

In many surveys, when samplers have the possibility to group elements into clusters, they try to reduce the variance (3) by minimizing the components *within* clusters S_{2l}^2 ; rewriting (3) in terms of relative variance and assuming that the same number of sample elements or households is selected from each PSU (i.e. $h_c = h$). Hansen *et al.* (1953) and Cochran (1977) have shown that such relative variance could be expressed in terms of intra-class correlation δ , which is related to the degree of homogeneity of elements/households within clusters. So, the aim is to reduce such intra-cluster correlation and consequently the variance (3).

However, in general, the construction of clusters cannot be modified easily when planning a household survey. In many countries, clusters represent the Census Enumeration Areas (EAs), in others they correspond to some geographical or administrative units, such as municipalities or counties. The *vicinity* of households clustered in a PSU, therefore, is an essential condition for its implementation, since the corresponding survey costs are reduced and, consequently, clustering is preferred to simple random sampling. For all these reasons, the *within* clusters component S_{2l}^2 of variance cannot be modified and, therefore, cannot be reduced.

In such cases, samplers are able to do little in terms of δ or variance *within* clusters reduction. Alternatively, efforts have been made to try to reduce the effect of variance *between* clusters S_{1l}^2 in each stratum l . In practice, clusters or PSUs are ordered according to some external variables available in the frame list, which are in some way correlated with the phenomenon under investigation Y: this permits getting more homogenous clusters and then it permits a reduction in variance *between* clusters S_{1l}^2 .

To make the sampling design more practical, the selection of clusters is performed with systematic sampling of such ordered frame lists. According to Kish (1965) this type of design can be seen as an (*implicit*) stratified sampling, where one single cluster or PSU is selected from each (*implicit*) stratum.

In developing countries, implicit stratification is often introduced in household surveys by sorting clusters in a certain geographical order within each explicit stratum, since frame lists of EAs may not contain other variables correlated with the phenomenon under investigation. In this way, the classical construction of a *serpentine* of clusters, is the most effective method that has been implemented so far for the preparation of the list frame of clusters from which the systematic sampling is conducted. In the present paper we aim to further improve the implicit stratification of a geographic serpentine, by ordering the clusters according to another socio-economic characteristic: a specific description of the living conditions of the clusters included in the so-called *poverty mapping*. This approach will be further described in Section 4.

However, it is well-known in the literature that this design of systematic selection of clusters in the first stage makes it impossible to find the exact formula for an unbiased estimator of the variance in (2), which could be renamed as $Var\{\hat{Y}_{sys}\}$, referring to systematic sampling. Wolter (1985) and Cochran (1977) proposed several methods of approximate estimators of $Var\{\hat{Y}_{sys}\}$, while in the last two decades, replication methods have reached more precise and unbiased estimators, which include Taylor Linearization (TL), Jack-knife Repeated Replications (JRR), Balanced Repeated Replications (BRR), and Bootstrap and Random Groups (RG). Recently, Verma and Betti (2011) have proposed a developed version of JRR, while Arnab *et al.* (2015) have proposed some more refined methods of variance estimation in the case of systematic sampling in the first stage. They show that such proposed methods and JRR perform better than other methods that are usually adopted by National Statistical Offices in developing countries.

In light of the above, the variance formulas (2) and (3) are not applicable when estimating complex measures from complex sample designs, but they are very useful as benchmarks for comparison (Kalton *et al.*, 2015). Such comparisons could also be done with reference to Simple Random Sampling (SRS). In fact, Kish (1965) firstly defined the so-called *design effect* as “...the ratio of the variance of any estimate, say, y , obtained from a complex design to the variance of y that would

apply with a SRS or unrestricted sample of the same size”. In formulas, $Deft^2$ ($Deft$ in terms of standard errors), could be written as:

$$Deft^2 = \frac{var\{\hat{Y}_{complex}\}}{var\{\hat{Y}_{SRS}\}} \quad (4).$$

It could be very useful to decompose $Deft^2$ in several multiplicative parts; the most important ones are: i) effect of stratification (usually, but not always,² smaller than 1); ii) effect of clustering (greater than 1); iii) effect of sample weighting adjustment (greater than 1; known as the *Kish effect*).

In Section 3, we provide a brief description of the theoretical background of poverty mapping and its implementation in Tunisia from 2014 to 2015.

3. Theory of poverty mapping and its application in Tunisia

There are several methodologies for estimating poverty at a local level; one of the first was developed by Elbers *et al.* (2002, 2003), and it is often known as the ELL approach. In general, such model-based techniques are defined as Small Area Estimation (SAE) methods (Rao, 2003), and may fall in one of the broad categories: i) area based models (for examples Fay and Herriot, 1979) in which the model is estimated on the basis of areas as statistical units; and ii) unit-level models, where instead the unit of analysis in the model is the individual or the household. Within this second category, Molina and Rao (2010) have innovated with the Empirical Best (EB): in the general literature of SAE, both ELL and EB are nested error model used for small area estimation, which were originally proposed by Battese *et al.* (1988). Later, the World Bank improved the original ELL method by including EB predictors and incorporating heteroskedasticity and survey weights in order to take into account complex sampling designs (Van der Weide, 2014).

The present paper is based on Molina’s H3-CensusEB estimator found to be superior in terms of bias and mean square error (MSE) to ELL and Molina and

² Not always in the case of designs for multi-topic household surveys. In these cases, in fact, generally rural/poor areas are oversampled to get enough observations for domain level analysis on these populations – which has the opposite effect (design greater than 1).

Rao's Empirical Bayes prediction estimates (Corral, Molina and Nguyen, 2020; Corral et al., 2021).

The current poverty mapping method imputes consumption expenditure for census households based on a model estimated from the household survey by applying the estimated coefficients of the model to the same variables from the census data. Poverty and inequality statistics for each small area are then calculated with the imputed consumption estimates of census households. The method does not only estimate poverty indices but also their corresponding standard errors. This methodology involves three major steps. The first step is to select a set of variables that are common to a census and a household expenditure survey.

The subset of variables that are found both in the census and the survey is used to estimate a regression model of per capita consumption using the survey data. In the second step, the set of parameter estimates obtained from the regression model is applied to the same set of variables identified in the census data to obtain predicted per capita consumption for each census household. Finally, based on the estimated level of per capita consumption, estimates of poverty, inequality and other welfare measures, as well as their standard errors, are calculated for any geographical unit with enough households to obtain reliable estimates. The point estimates and standard errors of the welfare indicators are calculated by Monte-Carlo simulations. In each simulation, a set of values for the model parameters are drawn from their estimated distributions, and an estimate of consumption expenditure and poverty rates are obtained.

3.1 Poverty mapping in Tunisia

This section describes the implementation of the method and key findings of small-area poverty estimation in Tunisia. It provides poverty rates and per capita consumption expenditure at the national level, and disaggregated at regional, governorate, and delegation levels – which constitute official poverty measures – and also disaggregated at the PSU level – which are measures used for implementing the new and original design of the EBCNV 2021.

The results in this section are derived from micro-data in the Population and Housing Census (RGPH) 2014 and the National Survey of Budget, Consumption and Standard of Living of Households (EBCNV) of 2015, which has the typical structure of HBSs. Details on the regressors included in the models are reported in INS (2020).

Mapping the poverty rates in Tunisia demonstrates that there is a high concentration of poverty in the North-West and Center-West landlocked regions. Poverty rates decrease moving towards the coastal North-East, Center-East and the Greater Tunis regions, although there are pockets of relatively high poverty rates there as well. Southern regions (South West and South East) are characterized by diverse levels of poverty.

The performance of models is tested by comparing the poverty rates obtained with the poverty mapping method with the EBCNV 2015 survey estimates at the regional and governorate levels. This comparison is feasible because the EBCNV 2015 is representative at both regional and governorate levels.

As shown in Table 1, poverty rates estimated on the basis of a household survey are more precise at the national level only. In fact, at regional levels, estimates from poverty mapping (Census 2014) show coefficients of variation (in the range of 8%-19%) that are lower compared to survey estimates (EBCNV 2015), with the coefficient of variation in the range of 24%-69%.

Table 1. Poverty rates by region

Region	EBCNV 2015			Poverty mapping - Census 2014			
	Poverty rate	Standard error	cv	Poverty rate	Standard error	cv	Absolute difference
TUNISIA	15.2	0.227	0.015	15.1	0.356	0.024	0.1
Greater Tunis	5.3	0.365	0.069	6.1	0.050	0.008	0.8
North-East	11.6	0.600	0.052	11.9	0.223	0.019	0.3
North-West	28.4	0.762	0.027	25.8	0.353	0.014	2.6
Centre-East	11.4	0.535	0.047	11.7	0.197	0.017	0.3
Centre-West	30.8	0.742	0.024	29.3	0.392	0.013	1.5
South-East	18.5	0.635	0.034	17.8	0.296	0.017	0.7
South-West	17.5	0.614	0.035	18.2	0.251	0.014	0.6

Source: our elaborations from INS (2020).

The greater reduction in variance is more clearly evident when disaggregating poverty measures at lower territorial levels. For example, Table 2 shows standard errors and the coefficient of variation of poverty measures estimated at the governorate level in Tunisia. Since sub-sample sizes are decreasing from regions to governorates, direct standard errors increase exponentially, that is, the coefficients of variation are higher than 10% in one-third of the governorates. On the other hand, standard errors of poverty rates calculated according to the poverty mapping show a very low increase. In fact, the coefficients of variation are all less than 10% except in the case of the Governorate of Ariana.

In the analysis of the poverty mapping measure at lower levels, in particular at the PSU level, direct estimates are affected by too high sampling errors and are therefore no longer significant. On the other hand, estimates from poverty mapping still show very limited standard errors. In this context, we shift our presentation from poverty rate to per capita consumption expenditure. This is done for two main reasons: i) per capita consumption expenditure estimates are usually more precise (see Betti *et al.*, 2018 in general, and INS, 2020 in this context); ii) due to their higher precision, per capita consumption expenditure is the variable used in our proposal for implementing implicit stratification in Section 4.

As previously mentioned, in Tunisia there are about 39,000 clusters used as PSUs. Before presenting results from the poverty mapping at cluster level, it is fundamental to bear in mind that although the PSU-level estimates are not suitable for dissemination and for policy making, they can still be very useful in informing sample design. In order to present such results, we synthesized measures at the governorate level in Table 3.

Table 2. Poverty rates by Governorate

Governorate	EBCNV 2015			Poverty mapping - Census 2014		
	Poverty rate	Standard error	cv	Poverty rate	Standard error	cv
TUNIS	3.5	0.575	0.164	4.6	0.360	0.078
ARIANA	5.4	0.698	0.129	7.0	1.024	0.146
BEN AROUS	4.3	0.595	0.138	5.6	0.288	0.051
MANOUBA	12.1	1.379	0.114	9.8	0.697	0.071

NABEUL	7.4	0.800	0.108	8.1	0.517	0.064
ZAGHOUAN	12.1	1.311	0.108	13.7	0.605	0.044
BIZERTE	17.5	1.120	0.064	16.9	0.535	0.032
BEJA	32.0	1.848	0.058	26.4	0.641	0.024
JENDOUBA	22.4	1.215	0.054	21.5	0.858	0.040
LE KEF	34.2	1.459	0.043	33.1	0.843	0.025
SILIANA	27.7	1.787	0.064	24.7	1.007	0.041
SOUSSE	16.2	1.095	0.068	14.3	1.108	0.078
MONASTIR	8.3	1.149	0.138	7.7	0.426	0.055
MAHDIA	21.1	1.682	0.080	25.0	0.685	0.027
SFAX	5.8	0.661	0.114	6.3	0.400	0.063
KAIROUAN	34.9	1.330	0.038	29.3	0.354	0.012
KASSERINE	32.8	1.305	0.040	33.6	0.297	0.009
SIDI BOUZIDE	23.1	1.175	0.051	25.0	0.579	0.023
GABES	15.8	1.038	0.066	16.9	0.671	0.040
MEDNINE	21.6	1.179	0.055	18.5	0.733	0.040
TATAOUINE	15.0	0.995	0.066	17.8	0.815	0.046
GAFSA	18.0	1.085	0.060	19.0	0.862	0.045
TOZEUR	14.6	0.982	0.067	14.6	0.746	0.051
KEBILI	18.5	1.081	0.058	19.2	0.300	0.016

Source: our elaborations from INS (2020).

For each governorate, Table 3 reports the minimum estimated value of the coefficient of variation (i.e. standard error over per capita consumption expenditure), the maximum value of the coefficient of variation, and the median value. From Table 3 we can see that in many cases per capita consumption estimates at cluster level are very precise, with an error of even less than 2% of the estimated mean. In the column “Median” we may conclude that in all the governorates at least 50% of cluster estimates have an error smaller than 10% of the estimated mean. Finally, in the column “Maximum” we can observe that in most of the cases the error is smaller than 50% of the estimated mean. In other words, there are very few cases (clusters) in which the estimated precision is not high and the estimates are not significant.

In order to reinforce the findings of Table 3, we have examined the 95% confidence intervals of per capita consumption at cluster level, reported in Figure 1. The graphs shows that, once PSUs have been ranked by per capita consumption expenditure (or poverty rate), confidence intervals among the richer clusters are disjoint with confidence intervals among clusters whose per capita consumption

are close to the national average value, and disjoint with confidence intervals among the poorer clusters.

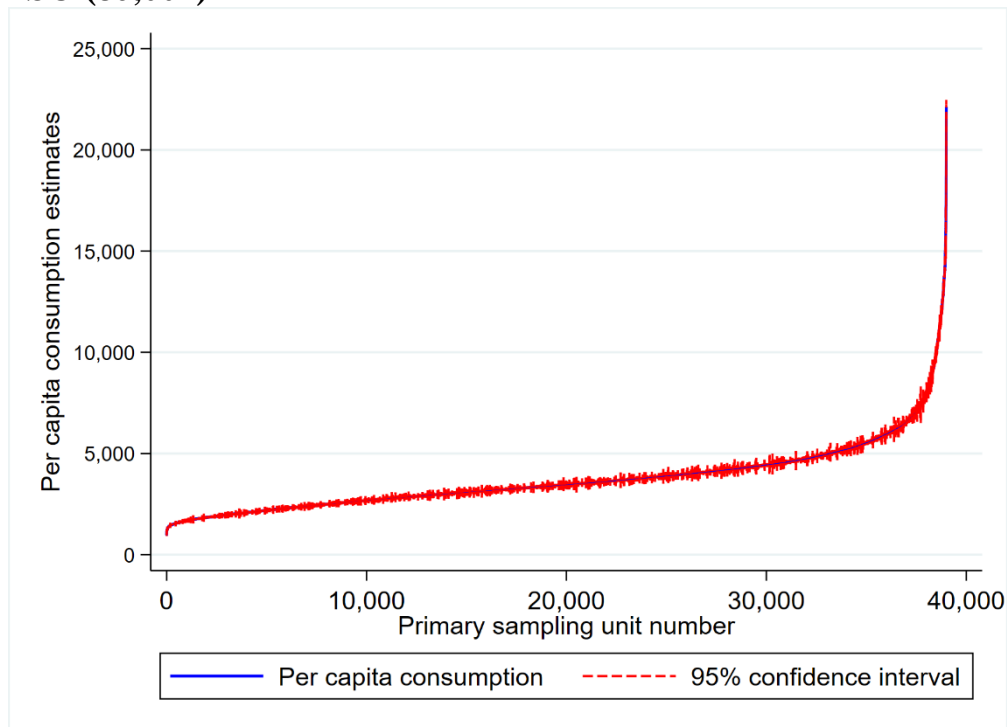
Table 3. Summary statistics on CVs at PSU level by Governorate

Governorate	Minimum	Median	Maximum
TUNIS	0.021	0.054	0.396
ARIANA	0.021	0.046	0.448
BEN AROUS	0.017	0.055	0.526
MANOUBA	0.021	0.075	0.407
NABEUL	0.022	0.060	0.459
ZAGHOUAN	0.026	0.073	0.579
BIZERTE	0.027	0.073	0.539
BEJA	0.035	0.082	0.500
JENDOUBA	0.023	0.065	0.610
LE KEF	0.024	0.078	0.453
SILIANA	0.027	0.081	0.455
SOUSSE	0.027	0.070	0.566
MONASTIR	0.024	0.055	0.379
MAHDIA	0.029	0.075	0.591
SFAX	0.022	0.063	0.513
KAIROUAN	0.026	0.062	0.481
KASSERINE	0.021	0.074	0.583
SIDI BOUZIDE	0.031	0.069	0.513
GABES	0.021	0.077	0.499
MEDNINE	0.023	0.071	0.540
TATAOUINE	0.030	0.087	0.649
GAFSA	0.024	0.085	0.570
TOZEUR	0.028	0.084	0.600
KEBILI	0.037	0.079	0.428

Source: our elaborations from INS (2020).

In order to better highlight the findings, in the Annex we have reported four graphs corresponding to the 1,800 clusters selected for the new EBCNV 2021 survey (see Section 4). In particular, Figure A1 reports confidence intervals for all 1,800 clusters, whereas Figures A2, A3 and A4 report confidence intervals for the bottom of the distribution, the distribution near the national mean consumption level, and the top of the distribution, respectively.

Figure 1. Estimated per capita consumption (TND per annum) for each PSU (39,001)



4. Implementation of the new design in Tunisia: EBCNV 2021

In this section, we discuss the implementation of the proposed design in the case of the new 2021 EBCNV in Tunisia. This could easily be extended to any HBS in any country and – with some specific adjustments – to any household survey.

Tunisia has conducted 10 EBCNVs so far, starting in 1966, then in 1974, and from 1980 one has been completed every five years. In the ninth edition, in 2010, the sample size of about 13,400 households allowed the survey results to be representative at regional levels, such as the seven regions reported in Table 1. For the tenth edition, in 2015, INS (2015) aimed to increase the survey sample size to let the survey estimates to be representative at the level of the 24 governorates. In order to reach this goal, 24 explicit strata were defined, one for each governorate, and the new sample size was fixed at more than 25,000 households. This sample size was defined on the basis of: i) the information of the population at that time; ii) the *explicit stratification* of the PSU in the design, and iii) the expected maximum error of 10% at the governorate level.

However, the EBCNV 2015 had a more efficient design than that taken into account for selecting the sample size. PSUs were selected with systematic

sampling, ordering PSUs by the average household size and prevalent economic status of households, as described in a document from the previous survey (INS, 2010). Moreover, it has been verified *ex-post* that further implicit stratification was also introduced by ordering PSUs, at the minimum, by delegation, thus forming a sort of geographic serpentine. These three factors contributed to better implement the design and the resulting errors became smaller than those expected when planning the design. Taking into account such implicit stratification, Betti and Pavelesku (2019) have estimated the real design effect of the EBCNV 2015 and the relative standard errors for estimates of mean per capita consumption expenditures at the governorate level. These results are reported in Table 4.

Since the aim of the EBCNV 2015 was to reach significative estimates at the governorate level, sub-sample sizes were planned to be as proportional as possible to sub-population sizes. This fact has resulted in homogenous post-stratification or calibration weights and small Kish effects, mostly well below 1.10. Moreover, the maximum errors estimated *ex-post* have been very small; even for the Governorates of Manouba and Mahdia have not been higher than 7%, well below the planned maximum percentage error of 10%.

Table 4. Per capita consumption expenditure, standard error and its components

Gov	Domain	Expenditure per capita	Standard error	Cv %	Kish	Deft
1	Tunis	5,810.12	296.78	5.11%	1.03	1.73
2	Ariana	5,461.10	274.30	5.02%	1.08	1.82
3	Ben Arous	4,878.05	177.31	3.63%	1.03	1.73
4	Manouba	4,377.36	296.80	6.78%	1.07	1.79
5	Nabeul	3,919.35	113.59	2.90%	1.00	1.68
6	Zaghouan	3,052.25	115.37	3.78%	1.00	1.68
7	Bizerte	2,867.93	88.08	3.07%	1.00	1.68
8	Béja	2,471.60	101.92	4.12%	1.02	1.71
9	Jandouba	2,943.07	96.03	3.26%	1.03	1.74
10	Kef	2,362.78	69.36	2.94%	1.05	1.77
11	Siliana	2,934.32	135.72	4.63%	1.06	1.78
12	Sousse	3,777.08	119.08	3.15%	1.02	1.72
13	Monastir	5,115.06	208.10	4.07%	1.02	1.72
14	Mahdia	3,195.92	210.05	6.57%	1.28	2.16

15	Sfax	4,698.13	133.87	2.85%	1.03	1.72
16	Kairouan	2,269.21	62.51	2.75%	1.01	1.70
17	Kasserine	2,542.98	92.31	3.63%	1.01	1.69
18	Sidi Bouzid	2,663.65	75.74	2.84%	1.04	1.74
19	Gabes	3,043.21	83.40	2.74%	1.01	1.69
20	Mednine	3,318.50	91.06	2.74%	1.05	1.76
21	Tataouin	3,538.70	106.89	3.02%	1.08	1.81
22	Gafsa	3,154.77	106.11	3.36%	1.01	1.70
23	Tozeur	3,191.66	106.38	3.33%	1.01	1.70
24	Gbeli	2,834.20	72.02	2.54%	1.03	1.74
	Tunisia	3,871.94	53.00	1.37%	1.30	2.19

Source: our elaborations of the EBCNV 2015.

During the first planning phases of the new EBCNV 2021, the INS initially took into account the same three aspects for determining the sample size: i) the information of the population at that time; ii) the *explicit stratification* of the PSU in the design, and iii) the expected maximum error of 10% at the governorate level. These would have resulted in a sample size of about 29,000 households and a total cost of about 6.9 million dinars.

Given the comparative benchmark of 6.9 million dinars and the benchmark of expected maximum error of 10% at the governorate level, the aim of the new sampling design has been to reduce both costs and maximum errors. Table 5 reports results from three scenarios, with different survey designs in terms of design effect and sample sizes needed to reach specified maximum percentage errors (left column). The three scenarios are as follows:

- a) Explicit stratification only: PSUs are stratified in the 24 governorates, and then selected with simple random sampling; this is an unrealistic scenario nowadays, even in household surveys in Africa. However, the expected standard errors are usually calculated hypothesizing this scenario so that the implemented sample sizes are overestimated for the purposes of the survey with very high costs.
- b) Serpentine: PSUs are first stratified in the 24 governorates, and then further stratified (implicitly) with systematic sampling by ordering the lists

according to some economic and **geographical** variables, which consequently form the list as a **geographical serpentine**.

- c) Poverty map: this is the design proposed in the present paper; PSUs are firstly stratified in the 24 governorates, and then they are further stratified (implicitly) with systematic sampling by ordering the lists of PSUs by average per capita consumption expenditure estimated with the poverty mapping method presented in Section 3.

On the basis of the new information from the current population in Tunisia, and the design effect of 2.19 estimated on the basis of the EBCNV 2015, Table 5 reports the expected sample sizes to obtain maximum percentage errors reported in the left column. The expected design effects corresponding to scenarios a) and c) have been estimated using equation (3) in Section 2, where Y is the mean per capita consumption expenditure from the poverty mapping in equation (6) in Section 3.

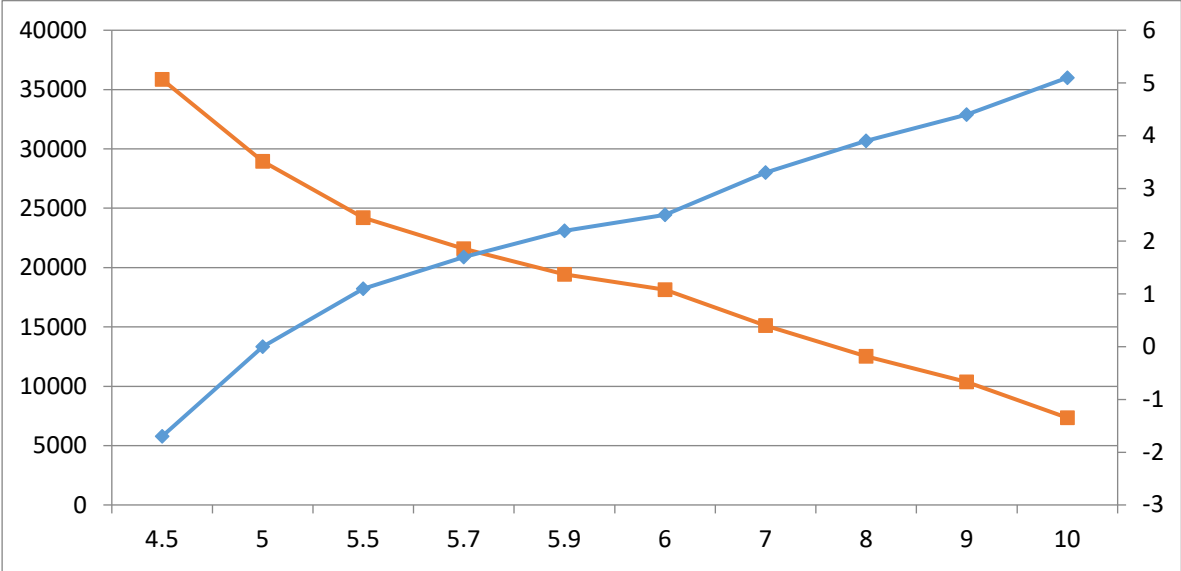
First of all, we can observe that in order to obtain a percentage error of about 10% a very high sample size is not needed when good designs with implicit stratification are taken into account (such as the classical “Serpentine”). However, this can be verified when very complex methods of variance estimation, usually based on replication techniques, are implemented. The recent JRR version, proposed by Verma and Betti (2011), and implemented to estimate the design effect 2.19 of the “Serpentine” design, is much more easily implemented by National Statistical Offices.

Table 5. Sample sizes in function of expected errors and type of design

	Explicit stratification	Serpentine	Poverty map
deft	3.71	2.19	1.87
error	4.5%	141315	35856
	5.0%	114073	28944
	5.5%	95345	24192
	5.7%	85129	21600
	5.9%	76616	19440
	6.0%	71509	18144
	7.0%	59591	15120
	8.0%	49375	12528
	9.0%	40862	10368
	10.0%	28944	10099

The further use of implicit stratification proposed by the “Poverty map” design is able to further reduce the design effect to 1.87. This could lead to a reduction in the needed sample size (and costs) and the desired maximum percentage error.

Figure 2. Sample sizes (left) and gains in cost (right) as function of expected maximum error



Based on the right column of Table 5, the INS has decided to adopt a final size of 21,600 households selected from 1,800 PSUs. This choice, also seen in Figure 2, is based on two main considerations: i) this size not only permits getting maximum percentage errors less than 10%, which can also reach 5.7%, which is even below the estimated maximum value in the EBCNV 2015; ii) thus, the total cost of the survey will be about 5.2 million Tunisian dinars, with a 1.7 million dinar increase in cost when compared with the initial estimated budget.

5. Discussion, further research, and concluding remarks

In this paper we have proposed an original methodology for improving the design effect of household surveys in general and Household Budget Surveys in particular. This is carried out by using the results of a poverty mapping exercise implemented by the method developed by the World Bank. Moreover, we have described how the new JRR method, proposed by Verma and Betti (2011), has been implemented for the first time in a developing country by a National Statistical Office for estimating standard errors for *complex measures* from *complex surveys*.

This discussion should take into account three main issues concerning the implementation of a poverty mapping for improving the design of household surveys: i) the capacities of National Statistical Offices to perform poverty mappings; ii) the problem of performing a poverty mapping when the census has been conducted too long ago; and iii) the variability affecting the mean per capita consumption per PSU in the poverty mapping exercise.

Concerning issue i), the poverty mapping exercise was initially quite an intensive procedure, which needed high econometric capacity and, therefore, only highly skilled World Bank consultants were able to implement it. In order to apply the ELL approach, the World Bank research team has developed a user-friendly software that can easily be implemented inside National Statistical Offices. PovMap version 2.1, the latest version as of 2015,³ provides computational solutions to all stages of poverty mapping activities. However, the new methods mentioned earlier are developed only in a Stata module produced by a team of researchers at the World Bank.⁴ This software is valid for use when there is a recent census available and a recent Household Budget Survey (or World Bank format of Living Standard Measurement Survey, LSMS) also available.

Concerning issue ii), often the question may be: “What should be done when the poverty map is too old?” In some countries or cases, the poverty map may be too old or not updated enough in order to properly implement the methodology proposed in this paper. This may occur primarily because a census is usually conducted every 10 years. There are two possible solutions depending on the type of recent household survey that is available. In the case of a recent HBS or LSMS, the proposed methodology could take into account the so-called “updated” or “further-updated” poverty maps that are prepared with recent methodologies developed under World Bank projects. In particular, Dabalen and Ferrè (2008) have shown how to implement the Lemieux (2002) approach to link the new HBS with the old HBS, and then update the poverty mapping through the use of the old Census. Moreover, Betti *et al.* (2013) have demonstrated that this updating is still

³ <https://www.worldbank.org/en/research/brief/software-for-poverty-mapping>.

⁴ The Stata package is available here <https://github.com/pcorralrodas/SAE-Stata-Package>.

valid also after several years, that is, when the poverty mapping is further updated with a third, more recently updated HBS. Otherwise, if a more recent HBS or LSMS is not available but, instead, there is a recent survey not covering consumption expenditure information, the proposed methodology could be complemented by the two methods proposed by Azzarri *et al.* (2006), or Dang *et al.* (2014, 2017).

The third issue covered in this discussion consists in the fact that the per capita consumption expenditure estimated at PSU level from the poverty mapping exercise is affected by (small) standard errors. Such variability should be taken into account in order to refine the original contribution proposed in this paper. We aim to develop this piece of new theoretical research and present it in a future paper.

References

- Azzarri, C., Carletto, G., Davis, B., Zezza, A. (2006), Monitoring poverty without consumption data: An application using the Albania panel survey, *Eastern European Economics*, **44(1)**, pp. 59-82.
- Battese, G.E., Harter, R.M., Fuller, W.A. (1988), An error-components model for prediction of county crop areas using survey and satellite data, *Journal of the American Statistical Association*, **83(401)**, pp. 28–36.
- Betti, G., Bici, R., Neri, L., Thomo, L., Sohnesen, T.P. (2018), Local poverty and inequality in Albania, *Eastern European Economics*, **56(3)**, pp. 223-245.
- Betti, G., Dabalen, A., Ferré, C., Neri, L. (2013), Updating poverty maps between Censuses: a case study of Albania, in Laderchi C.R., Savastano S. (eds.), *Poverty and Exclusion in the Western Balkans*, Economic Studies in Inequality, Social Exclusion and Well-Being 8, Springer Science+Business Media, New York.
- Betti, G., Pavelesku, D. (2019), Calculation of the design effect and JRR standard errors in the 2015 Household Budget Survey in Tunisia, Report to The World Bank.
- Corral, P., Molina, I., Nguyen, M.C. (2020), Pull your small area estimates up by the bootstraps. World Bank Policy Research Working Paper No 9256. Washington DC: The World Bank.

- Corral, P., Kastelic Himlein, K., Mcgee, K.R., Molina, I. (2021), A Map of the Poor or a Poor Map?, World Bank Policy Research Working Paper Series No. 9620. Washington DC: The World Bank.
- Dabalen, A., Ferrè, C. (2008), Updating Poverty Maps: A Case Study of Albania. Unpublished working paper. Washington DC: The World Bank,
- Dang, H.A., Lanjouw, P., Serajuddin, U. (2014), Updating Poverty Estimates at Frequent Intervals in the Absence of Consumption Data: Methods and Illustration with Reference to a Middle-Income Country, World Bank Policy Research Paper No. 7043. Washington DC: The World Bank.
- Dang, H.A., Lanjouw, P., Serajuddin, U. (2017), Updating Poverty Estimates at Frequent Intervals in the Absence of Consumption Data: Methods and Illustration with Reference to a Middle-Income Country, *Oxford Economic Papers*, **69(4)**, pp. 939-962.
- Elbers, C., Lanjouw, J.O., Lanjouw, P. (2002), Micro-level estimation of welfare. World Bank Policy Research Working Paper No 2911. Washington DC: The World Bank.
- Elbers, C., Lanjouw, J.O., Lanjouw, P. (2003), Micro-Level Estimation of Poverty and Inequality. *Econometrica*, **71(1)**, pp. 355–364.
- Institut National de la Statistique (INS) (2010), Enquête nationale sur le Budget, la Consommation et le Niveau de vie des ménages pour l'année 2010, Tunis.
- Institut National de la Statistique (INS) (2015), Enquête nationale sur le Budget, la Consommation et le Niveau de vie des ménages pour l'année 2015, Tunis.
- Institut National de la Statistique (INS) in collaboration with the World Bank (2020), Carte de la pauvreté en Tunisie, Tunis.
- Kish, L. (1965), *Survey Sampling*. New York: John Wiley & Sons.
- Lemieux, T. (2002), Decomposing changes in wage distributions: a unified approach, *Canadian Journal of Economics*, **35(4)**, pp. 646-688.
- Liu, J., Iannacchione, V., Byron, M. (2002), Decomposing design effects for stratified sampling. *Proceedings of the survey research methods section, American statistical association*, pp. 2124-2126.

- Molina, I. (2019), Desagregación de datos en encuestas de hogares: Metodologías de estimación en áreas pequeñas. CEPAL: Comisión Económica para América Latina y el Caribe.
- Molina, I., Rao, J.N.K. (2010), Small area estimation of poverty indicators. *Canadian Journal of Statistics*, **38(3)**, pp. 369–385.
- Rao J.N.K. (2003), *Small Area Estimation*. New York: Wiley.
- Tarozzi, A., Deaton, A. (2009), Using census and survey data to estimate poverty and inequality for small areas. *The Review of Economics and Statistics*, **91(4)**, pp. 773-792.
- UNECE (2020), *Poverty Measurement. Guide to Data Disaggregation*, Geneva: United Nations Economic Commission for Europe.
- Van der Weide, R. (2014), GLS estimation and empirical bayes prediction for linear mixed models with heteroskedasticity and sampling weights: A background study for the povmap project. World Bank Policy Research Working Paper No 7028. Washington DC: The World Bank.
- Verma, V. (1991), *Sampling Methods*, Statistical Institute for Asia and Pacific, Tokyo.
- Verma, V., Betti, G. (2011), Taylor linearization sampling errors and design effects for poverty measures and other complex statistics, *Journal of Applied Statistics*, **38(8)**, pp. 1549-1576.
- Wolter, K. (1985), *Introduction to Variance Estimation*. Springer-Verlag.

Annex

Figure A1. Estimated per capita consumption (TND per annum) for each 21st PSU (result in selected 1,800 PSUs)

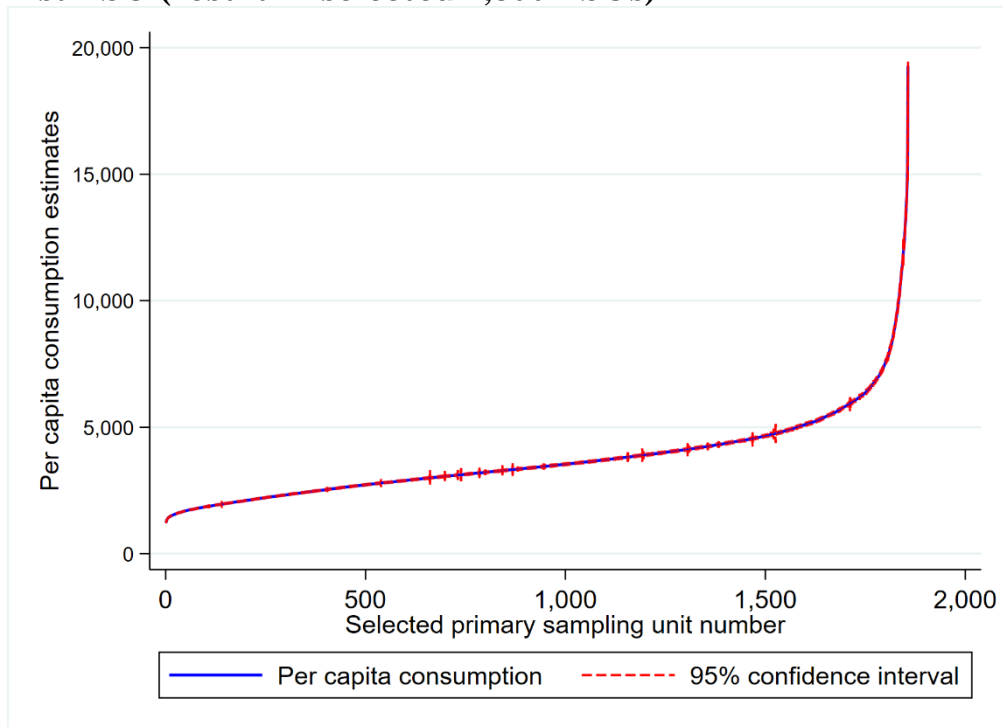


Figure A2. Estimated per capita consumption for each of selected 1,800 PSU, bottom distribution

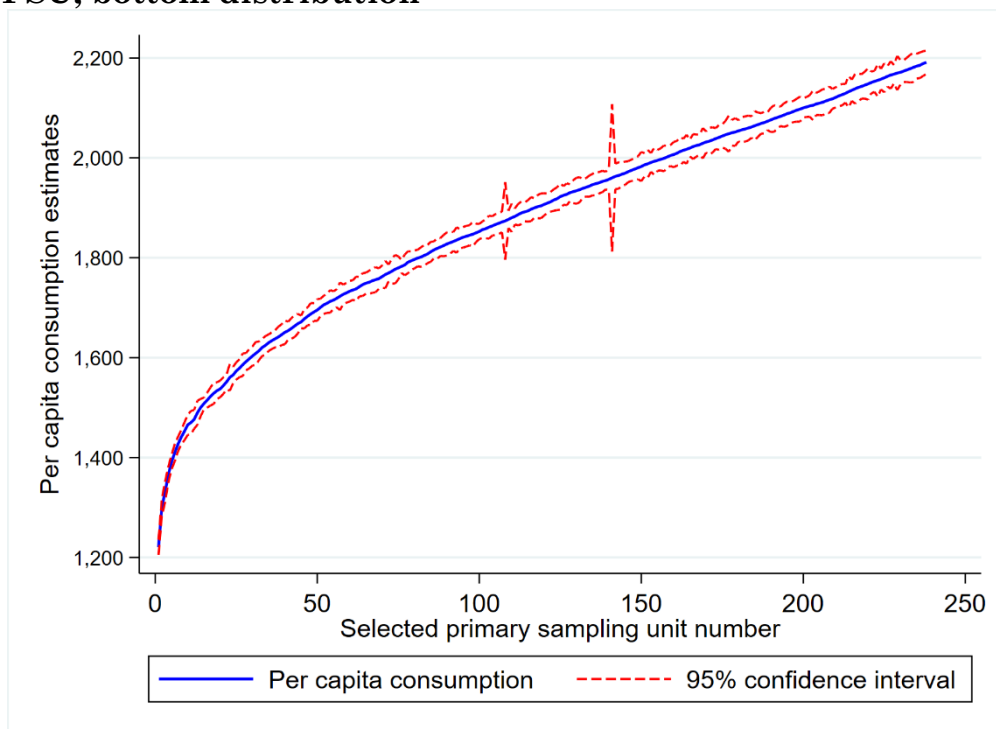


Figure A3. Estimated per capita consumption for each of selected 1,800 PSU, distribution near national mean per capita consumption of TND 3,871

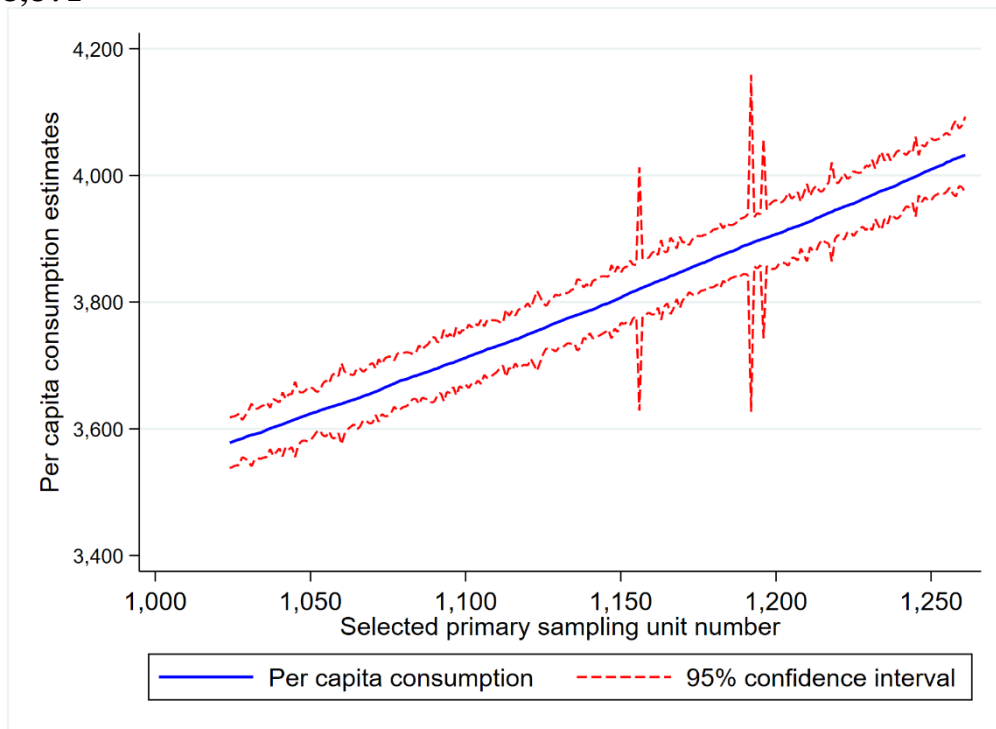


Figure A4. Estimated per capita consumption for each of selected 1,800 PSUs, top distribution

