DEVELOPMENT OF A SAMPLING METHOD FOR HOUSEHOLD SURVEYS IN POST-WAR BOSNIA AND HERZEGOVINA

Peter Lynn

Institute for Social and Economic Research

The Institute for Social and Economic Research (ISER) specialises in the production and analysis of longitudinal data. ISER incorporates the following centres:

- ESRC Research Centre on Micro-social Change. Established in 1989 to identify, explain, model and forecast social change in Britain at the individual and household level, the Centre specialises in research using longitudinal data.

- ESRC UK Longitudinal Studies Centre. This national resource centre was established in October 1999 to promote the use of longitudinal data and to develop a strategy for the future of large-scale longitudinal surveys. It is responsible for the British Household Panel Survey (BHPS) and for the ESRC's interest in the National Child Development Study and the 1970 British Cohort Study

- European Centre for Analysis in the Social Sciences. ECASS is an interdisciplinary research centre which hosts major research programmes and helps researchers from the EU gain access to longitudinal data and cross-national data sets from all over Europe.

The British Household Panel Survey is one of the main instruments for measuring social change in Britain. The BHPS comprises a nationally representative sample of around 5,500 households and over 10,000 individuals who are reinterviewed each year. The questionnaire includes a constant core of items accompanied by a variable component in order to provide for the collection of initial conditions data and to allow for the subsequent inclusion of emerging research and policy concerns.

Among the main projects in ISER's research programme are: the labour market and the division of domestic responsibilities; changes in families and households; modelling households' labour force behaviour; wealth, well-being and socio-economic structure; resource distribution in the household; and modelling techniques and survey methodology.

BHPS data provide the academic community, policymakers and private sector with a unique national resource and allow for comparative research with similar studies in Europe, the United States and Canada.

BHPS data are available from the Data Archive at the University of Essex
http://www.data-archive.ac.uk

Further information about the BHPS and other longitudinal surveys can be obtained by telephoning +44 (0) 1206 873543.

Readers wishing to cite this document are asked to use the following form of words:

> **Lynn, Peter (October 2003) 'Development of a Sampling Method for Household Surveys in Post-War Bosnia and Herzegovina',** *Working Papers of the Institute for Social and Economic Research*, **paper 2003-26. Colchester: University of Essex.**

For an on-line version of this working paper and others in the series, please visit the Institute's website at: http://www.iser.essex.ac.uk/pubs/workpaps/

Institute for Social and Economic Research
University of Essex
Wivenhoe Park
Colchester
Essex
CO4 3SQ UK
Telephone: +44 (0) 1206 872957
Fax: +44 (0) 1206 873151
E-mail: iser@essex.ac.uk
Website: http://www.iser.essex.ac.uk

ABSTRACT

This paper describes a method of sampling that has been developed for use on national household surveys in Bosnia and Herzegovina (BiH). The design and implementation of household surveys in BiH is particularly challenging for a number of reasons, including the absence of population registers or other sampling frames, the absence of relevant Census data or other population statistics, and continuing rapid change in the population of occupied dwellings (building, renovation, and associated population movement). The paper sets out the components of a sampling method that is both practical and statistically efficient. Both the practical and statistical arguments for each proposed component are presented. The results of a field test of the method are summarised. The sample design adopted for the BiH Household Budget Survey 2003-04, the first survey to use the new methodology, is described and the paper also sets out how the design can be adapted to the needs of other surveys. The sampling solutions adopted and described in this paper are likely to be applicable also to other situations in which analogous challenges exist.

Key words: Area sampling, Bosnia and Herzegovina, multi-stage sampling, sample design,

1. Background: Some History

Several factors make the design and implementation of household surveys in Bosnia and Herzegovina (BiH) particularly challenging. In one way or another, all of these factors are the result of recent history. It is therefore necessary to summarise some key aspects of the recent history of BiH in order to put the work described in this paper into context.

From 1945 until 1992, BiH was one of eight entities that collectively constituted Yugoslavia, a communist federal state that was ruled for three-quarters of this period by Marshall Tito. The eight entities included six republics - the other five being Slovenia, Croatia, Serbia, Montenegro and Macedonia - and two autonomous provinces, Vojvodina and Kosovo. During 1988 and 1989, the new leader of the Serbian Communist Party, Slobodan Milošević, organised a series of strikes and mass demonstrations against the ruling factions in Serbia, Vojvodina and Montenegro. This led to the resignation of the entire Politburos in both Vojvodina and Montenegro, both of which were replaced with Milošević's supporters. In March 1989, at Milošević's request, the Serbian Assembly passed constitutional amendments abolishing the political autonomy of Vojvodina and Montenegro. This move was met by mass demonstrations and a general strike in Kosovo. These were crushed by Serbian security forces and in June 1990 Milošević unilaterally abolished the Kosovan Assembly, leaving him in control of 4 out of the 8 votes in the Yugoslav federal government.

Perhaps fearing imminent Serbian dominance of the federation, Slovenia, Croatia and Macedonia began to assert their own identities. In October 1989, Slovenia passed a new state constitution, giving its state laws precedence over federal laws and declaring its right to secede. In January 1990, the Slovenian Communists walked out of the Yugoslav Communist Party Congress, renamed themselves the Party of Democratic Renewal and began making arrangements for multi-party elections in the Spring. Croatia too held multi-party elections in the Spring of 1990 and these were won by the new Croatian Nationalist Party led by Franjo Tudjman. During this election, the Serbs who formed the majority population in the Knin

region, adjacent to the north-western border of BiH, had organised themselves into the "Serbian Democratic Party" (SDS), primarily as a means of asserting their cultural identity. By the Summer, this party had been taken over by an extremist leader with close connections to Milošević and had begun causing civic unrest. Later in the year, the Serbs were referring to the Knin region as the "Serb Autonomous Region of the Krajina" and had formed their own parliament.

In BiH, as in most of the other republics, the Communist Party had disintegrated in early 1990. The neighbouring nationalisms of Croatia'a Tudjman and Serbia's Milošević were threatening presences (80% of BiH's borders are with Croatia or Serbia) and a number of nationalist or national parties took shape within BiH. In December 1990, a referendum in Slovenia resulted in an overwhelming vote in favour of becoming an independent state. The same happened in Croatia in May 1991 and the following month both republics formerly declared independence. Milošević initiated military action against both new states, but found well-organised resistance in Slovenia and soon withdrew. In Croatia, however, the Serb actions had escalated into a full-scale war by August 1991.

It has been estimated (Poulton, 1991) that in 1990 the population of BiH consisted of approximately 44% Muslims, 31% Serb, 17% Croat and 8% other. It is worth also noting that many commentators point to the fact that ethnic and religious origins in BiH had not played an important part in many people's lives prior to 1990. Inter-marriage was common and many people had never considered their origin until events forced them to do so (Malcolm, 1994). In fact, the three main groups are not ethnically distinct. Elections held in December 1990 returned a government with seats distributed between parties representing the three main national groups approximately in proportion to their population sizes. The Bosnian government supported Slovenia and Croatia in resisting Serbian oppression and any plans to re-draw the boundaries of the republics, but also opposed the intentions of Slovenia and Croatia to become completely independent as this would leave BiH in a particularly vulnerable position within a rump Yugoslavia completely dominated by Serbia. However, the EEC agreed in December 1991 formally to recognise Slovenia and Croatia as independent states. Simultaneously, they invited BiH to apply for independence and to first hold a referendum on the subject.

The referendum was held on 29 February and 1 March 1992. Despite a boycott by the SDS, which prevented ballot boxes entering the areas it controlled and forbade Serbs from voting, roughly 64% of the electorate did vote and voted almost unanimously for independence. Bosnia was recognised by the EEC as an independent state on 6 April 1992. Almost immediately, Serb military action in BiH escalated dramatically, with the explicitly-stated intention of annexing large swathes of northern and eastern BiH to Serbia. Thus began a long, complicated and vicious civil war, the details of which can be found elsewhere (Cigar 1995, Glenny 1996, Malcolm 2002, Rieff 1995).

Between 1992 and 1995, the Bosnian Serbs, supported by the Serbian army, attempted to "ethnically cleanse" the target areas in the north and east of BiH, forcing Muslims and Croats to move to other areas or simply murdering them. Many settlements were left completely empty of inhabitants while others saw dramatic reductions in population. Some areas, mainly in urban centres, saw rapid population growth caused by the arrival of large numbers of refugees. It has been estimated by several agencies (Papić, 2001) that over 250,000 persons died during the war and around 1.3 Million persons left the country during or soon after the war, of whom a little over half are believed to have found permanent solutions abroad and to have no intention of returning (UNHCR, 1998). Overall, it is estimated that around 50% of the surviving 1991 population of BiH have changed their place of residence.

The war came to an end with the Dayton Accord of 21 November 1995, which was later formalised as a treaty and signed by the Presidents of BiH, Croatia and Serbia in Paris on 14 December 1995 (Holbrooke, 1998). The treaty created two entities within the overall state of BiH. The Muslim-Croat Federation of Bosnia and Herzegovina (FBiH) was to account for 51% of the territory, with the Republika Srpska (RS) accounting for the remaining 49%. Each entity would have its own parliament with considerable autonomy. Importantly, confederation between Serbia and Republika Srpska was ruled out and the integrity of the state borders was emphasised. All citizens of BiH were to be given the right to return to their previous homes, in whichever entity they may be. Policies to encourage displaced persons to return home have been implemented with increased success in recent years, and in total almost 1 million persons have returned to their pre-war homes since 1995

according to the United Nations High Commission for Refugees (UNHCR 2003). Of these, about 530,000 were returning from elsewhere within BiH and about 430,000 returned from abroad.

2. Surveys and Population Statistics

Population censuses were carried out in Yugoslavia every ten years between 1951 and 1991. The last census carried out in BiH was therefore in 1991. Prior to the break-up of Yugoslavia, censuses and surveys were designed and co-ordinated by the Federal Statistical Office in Belgrade. Data collection was the responsibility of the Statistical Offices of each republic. Consequently, most of the expertise in survey design issues on the one hand, and analysis and reporting on the other, was concentrated in Belgrade.

No official surveys were carried out in BiH during the war or in the years immediately after. In 1996, a structure of official statistics institutions was developed. This primarily consists of three organisations: a State Agency - responsible for the production and dissemination of state-wide statistics – and two entity institutes, responsible for data collection and processing and the production and dissemination of entity-level statistics. Additionally, the area around and including the strategically-important town of Brčko in the north-east of BiH was constituted as a semi-autonomous district that does not belong to either entity. Brčko therefore has its own statistical office responsible for data collection and processing. Thus, any state-wide survey must involve four statistical institutions:

- The Agency for Statistics of the State of BiH (ASBiH: http://www.bhas.ba);

- The Federation of BiH Institute of Statistics (FBiHIS: http://www.fzs.ba);

- The Republika Srpska Institute for Statistics (RSIS: http://www.rzs.rs.ba);

- The Statistical Office of Brčko District (SODB: http://www.brcko.org).

In fact, a decision of the High Representative of BiH in October 2002 (OHR, 2002) obliges all parties to implement a proposed law on statistics that was endorsed by the Council of Ministers in July 2001 but not adopted. This involves eventually combining all four statistical institutions into a single organisation, but as of autumn 2003 there is no apparent progress towards this goal.

The first survey carried out by the statistical institutions in BiH was a Living Standards Measurement Survey in 2001. This is described in section 3 below.

3. The Living Standards Measurement Survey of 2001

In addition to loss of life and destruction of physical infrastructure, the war caused considerable social disruption and a decline in living standards for a large section of the population. Alongside these events, a period of economic transition to a market economy was occurring. The distributive impacts of this transition, both positive and negative, were unknown. While it was clear that welfare levels had changed, there was very little information on poverty and social indicators on which to base policies and programs. To provide such information, the three statistical institutions in BiH (the state agency and the two entity institutes) began developing the Living Standards Measurement Survey (LSMS) in 1999. The purpose of the survey was to collect data needed for assessing the living standards of the population and for providing key indicators for social and economic policy formulation. The survey was to provide data at the country and the entity level and to allow valid comparisons between entities to be made. A wide range of topics were covered: housing, education, health, employment, migration, credit, vouchers, social assistance, consumption, agricultural and non-agricultural activities.

The LSMS field work was carried out in a 2-month period between September and November 2001 by the three statistical institutions with financial and technical support from the UK Department for International Development (DfID), United Nations Development Program (UNDP), the Japanese Government, and the World Bank. The creation of a Master Sample for the survey was supported by the

Swedish Government through SIDA, the European Commission, DfID and the World Bank. The sample design is pertinent to the focus of this paper and will be described in some detail. Other aspects of the survey, including questionnaire content, are documented in World Bank (2003).

3.1 SAMPLE DESIGN

Advisers from Statistics Sweden developed the sample design for the LSMS and provided the statistical institutions in BiH with instructions for its implementation. The basic approach was to carry out a complete enumeration of households in a set of areas and then to treat this as a "master sample" from which a sample of households for the LSMS could be selected. In summary, the stages in the sample design and selection were as follows:

1. Selection of a sample of 25 municipalities, selected with probability proportional to estimated size within each of 6 explicit strata, but variable average sampling fractions between strata;

2. Complete field enumeration of all households in each of those municipalities;

3. Selection of 450 census enumeration areas (EAs) from within the 25 sampled municipalities. These were selected with probability proportional to the number of households enumerated at stage 2 within each of the 2 entities, but with variable average sampling fractions between the two entities;

4. Selection of 12 households, with equal probabilities, within each sampled EA, resulting in a total sample of 5,400 households: 3,000 in FBiH and 2,400 in RS.

Each of these stages are now described in more detail. Some conclusions will then be drawn regarding the statistical efficiency of the sample design.

Stage 1: Selection of municipalities

There were 146 municipalities in BiH at the time and for each an estimate of the number of households was available. Municipalities were sorted into six strata

6

formed by the cross-classification of entity (FBiH/RS) with an urban/rural indicator based on 1991 census data (mostly urban settlements/mostly mixed urban and rural settlements /mostly rural settlements). The number of municipalities to select in stratum i, m_i, was predetermined based on a decision (reason unknown to me and to the statistical institutions in BiH) to include approximately 50% of municipalities in the "mostly urban" strata, 20% of those in the "mixed" strata and 10% of those in the "mostly rural" strata. Selection was made with probability proportional to size within strata, where the size measure was a prior estimate of the number of households in the municipality (source unknown). Thus, for each municipality, the probability of selection was:

$$P_j = m_i \times \frac{N_{ij}}{N_{i*}}$$

- (1)

where

N_{ij} is estimated number of households in municipality j in stratum i (j = 1, …, M_i);

N_{i*} is the estimated total number of households in stratum i.

M_i is the total number of municipalities in stratum i (i=1, … , 6).

These probabilities are documented in Table 1, along with the values of M_i. The actual probabilities for the selected municipalities are listed in Table 2.

Stage 2: Field enumeration

In each sampled municipality, updated maps were created by the two entities' geodesic institutes. These were then used by interviewers as the basis for creating a listing of all dwellings in the municipality. At each dwelling, some simple demographic data was collected relating to every resident household, using face-to-face interviewing methods. If no contact was made after three attempts, the

Table 1: Selection probabilities of municipalities

i	Stratum	Total municipalities M_i	Sampled municipalities m_i	Probability for municipalities P_j
1	FBiH, mostly urban	10	5	$5 \times \dfrac{N_{1j}}{N_{1*}}$
2	FBiH, mostly mixed urban/rural	26	4	$4 \times \dfrac{N_{2j}}{N_{2*}}$
3	FBiH, mostly rural	48	5	$5 \times \dfrac{N_{3j}}{N_{3*}}$
4	RS, mostly urban	4	2	$2 \times \dfrac{N_{4j}}{N_{4*}}$
5	RS, mostly mixed urban/rural	29	5	$5 \times \dfrac{N_{5j}}{N_{5*}}$
6	RS, mostly rural	29	4	$4 \times \dfrac{N_{6j}}{N_{6*}}$

information was collected by proxy from neighbours. The data collected included the name, age, and sex of each household member.

Stage 3: Selection of enumeration areas

After completion of the field enumeration, many census enumeration areas (EAs) were found to now contain few, or even zero, households. Such small EAs were combined with geographically adjacent EAs to form groups of EAs, so that each group contained a minimum of 50 households. References to EAs hereafter refer to these grouped EAs.

Table 2: Selection probabilities of municipalities (selected municipalities)

Municipality j	Number of municipalities sampled in stratum m_i	Number of households in municipality N_{ij}	Total number of households in stratum N_{i*}	P_j
1 Novi Grad	5	31,453	202,307	0.77736
2 Centar Sarajevo	5	18,870	202,307	0.46637
3 Novo Sarajevo	5	19,839	202,307	0.49032
4 Zenica	5	36,447	202,307	0.90078
5 Tuzla	5	38,537	202,307	0.95244
6 Vogošća	4	5,371	146,688	0.14646
7 Travnik	4	14,375	146,688	0.39199
8 Visoko	4	11,312	146,688	0.30846
9 Breza	4	3,900	146,688	0.10635
10 Zavidovići	5	10,758	296,691	0.18130
11 Gradačac	5	13,182	296,691	0.22215
12 Posušje	5	4,564	296,691	0.07692
13 Kakanj	5	12,365	296,691	0.20838
14 Grude	5	3,823	296,691	0.06443
15 Srpska Ilidža	2	4,888	82,071	0.11912
16 Banja Luka	2	65,420	82,071	1.59423
17 Čajniče	5	1,487	182,543	0.04073
18 Novi Grad	5	8,961	182,543	0.24545
19 Prijedor	5	28,339	182,543	0.77623
20 Modriča	5	8,266	182,543	0.22641
21 Višegrad	5	5,581	182,543	0.15287
22 Kneževo	4	3,564	154,170	0.09247
23 Šamac	4	6,746	154,170	0.17503
24 Zvornik	4	14,623	154,170	0.37940
25 Srbac	4	7,215	154,170	0.18720

Note: The expected number of times that Banja Luka would be sampled under this design is 1.59. In other words, it would be certain to be sampled at least once. There is a 0.41 probability that it would be selected once and a 0.59 probability that it would be selected twice. Normal practice might be to treat such units as a separate stratum with P=1.0. It is not clear what practice was adopted in this case. I have assumed that Banja Luka was left on the list to be sampled PPS, and that if it had happened to be selected twice, that fact would be ignored (and the number of EAs to select would have be calculated in the same way as if it had only been sampled once). This is equivalent to just giving a selection probability of 1.0, so this is what has been assumed subsequently.

The number of EAs to select in municipality j was determined in the two entities as:

FBiH:
$$e_{ij} = 250 \times \frac{O_{ij}}{\sum\limits_{j \in s=1}^{14} O_{ij}}$$
- (2a)

RS:
$$e_{ij} = 200 \times \frac{O_{ij}}{\sum\limits_{j \in s=15}^{25} O_{ij}}$$
- (2b)

where O_{ij} is the enumerated number of households in municipality j in stratum i, not to be confused with the <u>prior</u> estimate of the number of households in municipality j in stratum i, N_{ij}. Note that the sums in the denominators of (2a) and (2b) are over the <u>sampled</u> municipalities in the respective entities. In other words, the sample EAs were allocated to sampled municipalities in proportion to the number of households in the municipality (but ignoring the numbers of households in non-sampled municipalities),

Within each municipality, the e_{ij} EAs were selected with probability proportional to size (number of households), so the probability of selecting EA k in municipality j (conditional upon having selected municipality j) was:

$$P_{k/j} = e_{ij} \frac{O_{ijk}}{O_{ij}}$$
- (3a)

So, substituting (2a) and (2b) respectively in (3a), we have:

FBiH:
$$P_{k/j} = \frac{250 \times O_{ijk}}{\sum\limits_{j=1}^{14} O_{ij}}$$
- (3b)

RS:
$$P_{k/j} = \frac{200 \times O_{ijk}}{\sum\limits_{j=15}^{25} O_{ij}}$$
- (3c)

The results of the enumeration showed that $\sum\limits_{j=1}^{14} O_{ij} = 224{,}796$ and $\sum\limits_{j=15}^{25} O_{ij} = 155{,}090$,

so FBiH:
$$P_{k/j} = \frac{250 \times O_{ijk}}{224{,}796}$$
- (3d)

and RS:
$$P_{k/j} = \frac{200 \times O_{ijk}}{155{,}090}$$
- (3e)

Stage 4: Selection of households

Twelve households were selected at random from each sampled EA, so the probability of selecting household l in EA k in municipality j (conditional upon having selected EA k in municipality j) was:

$$P_{l/jk} = \frac{12}{O_{ijk}}$$
- (4)

3.2 SELECTION PROBABILITIES

The overall probability of selection for household l in EA k in municipality j in stratum i is the product of the three conditional probabilities:

$$P_l = P_j \times P_{k/j} \times P_{l/jk}$$
- (5)

We can obtain this probability for FBiH by substituting from (1), (3d) and (4) and for RS by substituting from (1), (3e) and (4), viz:

FBiH:
$$P_l = \frac{3000 \times m_i \times N_{ij}}{224{,}796 \, N_{i*}}$$
- (6a)

RS:
$$P_l = \frac{2400 \times m_i \times N_{ij}}{155{,}090 \, N_{i*}}$$
- (6b)

11

The selection probabilities of households therefore vary between municipalities (but not between households within municipalities). The variation arises because different sampling fractions were used in each of the three strata within each entity (reflected in the term $m_i \big/ N_{i*}$) and because municipalities were selected within strata PPS (reflected in the term N_{ij}) – an imbalance that was not corrected at the subsequent stage, as it often is in PPS designs.

P_l is shown in Table 3 for each municipality. This indicates the variation in probabilities due to the design. It can be seen that there is a very large range of selection probabilities, from around 0.0006 in Čajniče to 0.0155 in Banja Luka. The consequence of this is a considerable design effect (Kish, 1965) which reduces the precision of estimates based upon the survey data. The actual design effect will vary from estimate to estimate, depending on the relationship between the statistic in question and the selection probabilities, in other words the relationship between the statistic in question and municipality, given that it is at the municipality level that selection probabilities vary. However, the design effect for a hypothetical statistic which has equal variance in each weighting class (municipality) can be a useful summary of the impact of the sample design on precision of estimates. Under this assumption of equal variance in each municipality, the design effect (due to variable selection probabilities) of a mean can be calculated as:

$$D\hat{e}ff_{VSP} = \frac{n \sum_{j=1}^{25} n_j \left(w_j^2 \right)}{\sum_{j=1}^{25} \left(n_j w_j \right)^2} \qquad - (7)$$

where n_j is the sample size (households) in municipality j and w_j is the design weight to be applied in analysis to each sample household in municipality j, $w_j = 1 \big/ P_l$ (Elliot, 1991).

12

Table 3: Overall selection probabilities of households

Municipality		Probability for municipality (from Table 2)	Probability for EAs	Probability for households	Overall probability
j		P_j	$P_{k/j}$	$P_{l/jk}$	P_l
1	Novi Grad	0.77736	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.01037
2	Centar	0.46637	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00622
3	Novo Sarajevo	0.49032	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00654
4	Zenica	0.90078	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.01202
5	Tuzla	0.95244	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.01271
6	Vogošća	0.14646	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00195
7	Travnik	0.39199	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00523
8	Visoko	0.30846	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00412
9	Breza	0.10635	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00142
10	Zavidovići	0.18130	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00242
11	Gradačac	0.22215	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00296
12	Posušje	0.07692	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00103
13	Kakanj	0.20838	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00278
14	Grude	0.06443	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00086
15	Srpska Ilidža	0.11912	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00184
16	Banja Luka	1.00000	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.01547
17	Čajniče	0.04073	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00063
18	Novi Grad	0.24545	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00380
19	Prijedor	0.77623	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.01201
20	Modriča	0.22641	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00350
21	Višegrad	0.15287	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00237
22	Kneževo	0.09247	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00143
23	Šamac	0.17503	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00271
24	Zvornik	0.37940	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00587
25	Srbac	0.18720	$K_1 O_{ijk}$	$12 / O_{ijk}$	0.00290

Note: $K_1 = \dfrac{250}{224,796}$ and $K_2 = \dfrac{200}{155,090}$.

The LSMS design described here resulted in a total sample size of 5,400 households, but the design effect estimated by (7) is 2.13. This means that the "effective sample size" (Elliot, 1991) is only 2,531. In other words, estimates from the LSMS are only of equal precision to an equal-probability sample of 2,531 households. The design effects are estimated as 2.34 for RS and 1.98 for FBiH,

leading to effective sample sizes of 1,025 and 1,516 respectively, compared to nominal sample sizes of 2,400 and 3,000 households. (Note that these are estimates of the design effects solely due to differential selection probabilities. There will also be other components of the overall survey design effect, notably a component due to sample clustering. This will tend to be greater than one too, so the overall survey design effects are likely to be even greater than those estimated here.) The design effect estimates presented here suggest that the LSMS design is rather inefficient for the purpose of design-based inference.

4. Constraints on Sample Design

Following the LSMS, the next priority for the BiH survey programme was identified as being the implementation of a Household Budget Survey (HBS). However, the only potential sampling frame for the HBS was the list of households arising from the LSMS enumeration of 25 municipalities. There were a number of problems with the use of these lists as a sampling frame:

1. They had been compiled in 2000. Considering the significant levels of internal migration and construction in BiH - partly as a result of the policies mentioned in section 1 to encourage persons displaced during the war to return home - changes were likely to have been considerable by the time that the HBS would have been carried out (2003 at the earliest).

2. Use of a similar design to the LSMS would have resulted in statistically inefficient sample. Consequently, a very large sample size would have been necessary to achieve suitable levels of precision.

3. Though the design effect due to variable selection probabilities could have been reduced, relative to the LSMS, by better distributing the sample of households across the 25 municipalities, this would have resulted in high concentrations of sample in a few very small municipalities. This in turn would have been likely to

increase the design effect due to clustering. It was not at all clear that an overall reduction in design effect could have been achieved.

4. Restriction of the sample to just 25 municipalities introduces considerable variation to estimates, no matter what the selection probabilities of municipalities. It would be preferable to include a much larger number of municipalities in the sample.

5. For the HBS, it would be desirable to make some estimates for sub-entity units (e.g. Cantons in FBiH) and for Brčko. This would not be possible with such a small sample of municipalities.

All of these objections were important, but the out-of-date nature of the lists alone was probably sufficient to rule them out. An alternative was sought, but it had to be one that allowed for the following constraints:

- There are no population registers, housing registers, or postal address registers that can be used as sampling frames;

- There is considerable internal migration and rapid change amongst the housing stock;

- The only available auxiliary data with state-wide coverage for small areas are from the 1991 census. Even simple census measures such as population counts are unlikely to bear much relationship to the *de facto* situation in 2003, given the huge population movements described in section 1, so census data are unlikely to be of any use in sample design;

- There are no up-to-date state-wide small scale maps available that could be used to help define areas as PSUs for area sampling;

- The statistical institutions in BiH have limited resources and limited experience of general population sampling methods.

The alternative design that was developed is outlined in the next section.

5. An Area-Based Approach

The approach adopted for the HBS is one that can be easily adapted to other general population surveys. It is a 3-stage sample design, using a form of area sampling. The small areas used as primary sampling units (PSUs) are 1991 census enumeration areas (EAs). In essence, the approach is to select an equal-probability sample of EAs, to carry out a field listing of dwellings in each sampled EA, and then to select an equal-probability sample of dwellings from that list. In other words, the sampling fraction of dwellings is fixed across all EAs, so the sample size of dwellings in each EA will be proportional to the total listed number of dwellings.

The selection probability of EA i is simply:

$$P_i = \frac{m}{M}$$
- (8)

where m is the number of EAs to be sampled and M is the total number of EAs in BiH. The selection probability of dwelling j in EA i, conditional upon the selected sample of EAs is:

$$P_{j/i} = \frac{n}{N^*}$$
- (9)

where n is the total number of dwellings to be sampled and N^* is the total number of dwellings listed in the sample of m EAs. The product of (8) and (9) gives the overall selection probability of each dwelling and this can be easily seen to be a constant:

$$P_j = \frac{m \times n}{M \times N^*}$$
- (10)

Of course, N^* is a random variable rather than a pre-determined design parameter, but:

$$E(N^*) = \frac{m \times N}{M}$$
- (11)

Substituting (11) into (10), we have:

$$E\left(P_j\right)=\frac{m\times n\times M}{M\times m\times N}=\frac{n}{N}$$ - (12)

5.1 SAMPLING SMALL AREAS

The first step in the process was to develop a list (frame) of EAs. Though conceptually simple, there were several challenges to be overcome. There was imperfect agreement between the entities regarding the ownership of some EAs. There were several EAs the status of which had only recently been agreed. Recent changes in status had not necessarily been incorporated in the lists held by the statistical institutions. There were some EAs whose status had not yet been established, so these could not be allocated to an entity. There were some EAs that were "split" between the entities, with each entity responsible for part of the EA. With the co-operation of all the statistical institutions, over a period between February and June 2003 it was eventually possible to create a list of EAs that was believed to be comprehensive, i.e. cover the complete territory of BiH, and mutually exclusive. This was created by merging the lists supplied by each of the entities, and a separate list for Brčko.

After several iterations, many EAs continued to appear on the lists of both entities, so these were placed in a separate stratum for sampling purposes. It was felt that for most of these EAs it may have been possible to reach agreement on which entity they belonged to, but it would have been inefficient to seek to reach such agreement. Rather, they could be treated as a separate stratum and subsequent to sample selection agreement could be sought only for the (much smaller number of) EAs selected. The total numbers of EAs on the list, by stratum, are shown in Table 4. For each EA, the listing included a unique EA identification number, identifiers and names for higher-level geographical units (settlements and municipalities), and an urban/rural indicator from the 1991 census.

Table 4: Numbers of EAs in Sampling Strata

Stratum	Number of EAs
FBiH only	10,287
RS only	7,511
Both FBiH <u>and</u> RS	1,254
Brčko	371
Total	*19,423*

Having created a list of EAs, an equal-probability sample can be selected systematically, treating the four strata shown in Table 4 as implicit strata and with further implicit stratification by municipality and by urban/rural within municipality (and by settlement within urban/rural). The implicit stratification will improve the precision of estimation (i.e. produce a design effect due to stratification of less than one) to an extent dependent upon the correlation between the target variables and the strata. Selection is made with a suitable random start and fixed interval.

For each EA, a sketch map had been created at the time of the 1991 census as the basis for the field operations. Though a few of these maps had been lost during the war, the vast majority had survived intact. In the 25 municipalities sampled for the LSMS, updated maps had been produced in 2000. The maps could be located using the unique EA identification number present on the EA list. It was felt that these sketch maps should be sufficient for a field worker to be able to identify on the ground the boundary of the EA. If sufficient time and resources were available between the selection of EAs and the start of the listing field work, then updated maps could be produced for any selected areas where significant changes are expected. This would involve showing the <u>same</u> boundary, but with respect to current features.

5.2 SAMPLING DWELLINGS

For each selected EA, a field worker would identify on the ground the territory covered by the EA, using the sketch map, and would then proceed to list all

dwellings/households within that area. There are a number of possible approaches to listing, including:

1. Listing households;

2. Listing dwellings from observation.

The first approach would involve making contact at each dwelling to ascertain whether it is occupied and if so by how many households. Basic information would have to be collected about each household, sufficient to allow unambiguous identification at a subsequent visit in the event that the household should be sampled. This approach is intrusive, as it involves collecting personal information from each household. It is also relatively costly and time consuming, as interviewers must make multiple visits to each address until they have been successful in making contact and obtaining the required information.

The second approach is non-intrusive, as the interviewers need only observe dwellings externally. It may work well in areas where dwellings are numbered or otherwise identified externally and where most dwellings are occupied and contain a single household. Unfortunately, areas with these characteristics are relatively few in BiH. It is common for houses and flats not to have numbers and, especially in rural areas, for streets not to have names. For the reasons described in section 1, many dwellings are unoccupied and this is not always obvious from an external viewing. Similarly, many dwellings are occupied by multiple households and some households live in locations that might not appear to be residential dwellings.

For all these reasons, it was proposed that listing should be carried out by a semi-intrusive approach. Interviewers would be instructed to use non-intrusive methods wherever this seemed appropriate (e.g. a block of flats where each flat has a number and at least a large majority seem to be occupied), but to collect a family name for each household where this was necessary in order to identify the dwelling (e.g. rural areas without numbering systems). Where possible, it should be established whether or not a dwelling is occupied: unoccupied dwellings should not be included on the list. However, in cases where occupancy status is not

established for certain, the dwelling should be included on the list in order to avoid any possible under-coverage.

When the listing is completed in all sampled EAs, the lists would be concatenated (in order of stratification) and a systematic sample of dwellings selected, using an interval of $I = {N^* \mskip-5mu/\mskip-3mu n}$ and a random start between 1 and **I**.

5.3 SAMPLING HOUSEHOLDS

Because the listing is done by semi-intrusive methods, the unit listed will sometimes be a household and sometimes a dwelling. When a selected unit turns out to contain more than one household, a rule is required to determine which households to include in the survey. Including all households at multi-household dwellings would retain the equal-probability nature of the sample of households, but might occasionally cause unpredictable increases in interviewers' workloads. Randomly selecting a single household would introduce some (probably very modest) variation in selection probabilities, but would be procedurally simple. A compromise might be to include all households up to some predetermined limit, n (e.g. n=3) and randomly sub-select n households at each sampled dwelling containing more than n. One of these methods must be chosen prior to implementing the design in the field, but the choice does not affect the prior stages. A field test of the method, described in the next section, would provide data that should help in deciding which of these 3 methods for sampling households within dwellings is likely to be preferable.

6. Field Test

A field test of the procedures for listing households/ dwellings was carried out in May 2003. This was designed with two aims in mind:

a) to obtain estimates of parameters that were needed in order to determine sample size and sample distribution for future surveys (e.g. proportion of listed units that turn out to be unoccupied);

b) to test the procedures for all stages of the process, from selection of the EAs to return of the completed listing forms, for feasibility, accuracy and resource requirements.

6.1 DESIGN OF THE FIELD TEST

A systematic random sample of 50 EAs was selected from a list stratified by municipality within entity. This resulted in the selection of 27 EAs in FBiH, 22 in RS and 1 in Brčko. The frame of EAs, as described in section 5.1, had not been finalised at this time and as a consequence there were some problems with some of the selected EAs and 4 of them had to be replaced (purposive selections).

FBiHIS, RSIS and SODB employed field workers to carry out the listing in each sampled EA. These field workers were generally people who had worked previously as interviewers for the statistical institutions. In most cases there was a one-to-one correspondence between field workers and EAs, though a couple of field workers dealt with 2 EAs each, thus there were 48 field workers in total. The field workers were briefed in one of two briefing conferences, held in Sarajevo and Banja Luka on 9 May and 12 May respectively.

The field workers were instructed to complete the listing in a single day. The idea was to test whether it was feasible using the semi-intrusive method for the listing to be done on a single visit to the area. It was recognised that this would inevitably mean that it would often not be possible to establish for certain whether or not a dwelling was occupied and consequently that the list would include some vacant properties. This disadvantage was expected to be of negligible importance relative to the extra resources that would have been needed to obtain full occupancy information.

Subsequent to completion of the listing, a second phase of the field test took place in each EA. In this second phase, field workers attempted to make contact at every dwelling in the EA and collect basic information about the number of households at each listed dwelling and the age and gender structure of each household. This would provide some information about:

- possible under-coverage (households at dwellings that were not listed);

- the proportion of unoccupied dwellings;

- the possible effects of sub-sampling at multi-household dwellings;

- the average number of households per listed unit, which determines the number of units to sample in order achieve a desired sample size of households.

Copies of the interviewer instructions and the listing forms used at each of the two phases are included in the appendix to this working paper. Field listing (phase 1 of the test) in each of the 50 areas was completed by 20th May and data entry was completed on 22nd May. The second phase field work was completed by 3rd June and entry of the data from that phase was completed on 6th June.

6.2 RESULTS OF THE FIELD TEST

Overall, the field test was very successful and demonstrated that the proposed approach was feasible. The success of the second phase in obtaining information regarding virtually every household in every listed dwelling enables considerable confidence to be placed upon the conclusions regarding the coverage and characteristics of the first phase listing.

The data from the field test provided estimates of population parameters that would be needed in order to finalise the details of the sample design for the HBS, viz. the proportion of listed dwellings that turned out to be unoccupied and the distribution of number of households per dwelling.

In total, 3748 units were listed at phase 1, i.e. a mean of 75.0 per EA. Of these, 327 (8.7%) were found at phase 2 to be unoccupied. The proportion unoccupied varied over the 50 enumeration areas from 0 (in 19 of the EAs) to 75.8%, though there were only 6 EAs where it exceeded 20% (5 of these 6 were in FBiH). Only 75 phase 1 units (2.0%) were found at phase 2 to contain more than one household. There were 84 "extra" households at these 75 units. There were 27 EAs in which there were 0 multi-household units. The proportion of phase 1 units found to contain more than one household exceeded 6% in only 5 EAs and the maximum was 17.7%.

Consequently, the mean number of households per listed dwelling was 0.935, implying that for every 1,000 households required for a survey sample it would be necessary to select 1,070 dwellings. There were 11 EAs where the mean number of households per phase 1 unit fell below 0.9 and 2 where it was greater than 1.1. The mean number of households per EA was therefore 70.1.

If it were decided to sub-sample households at multi-household dwellings, then a larger number of dwellings would need to be selected initially. However, given that the largest number of households found at a listed dwelling was 4 and the mean number of households at multi-household dwellings was only 2.1, sub-sampling would seem unnecessary. It was concluded that all households would be included at multi-household dwellings.

The field test also identified some areas where improvements could be made:

- Identification of area boundaries by interviewers. There were some difficulties with this, particularly in areas where roads and/or buildings had either been demolished or built (or both) since 1991.

- Inclusion of dwellings outside the sampled area (over-coverage). At least one instance was found where the interviewer had listed dwellings on both sides of a road, where the area boundary ran along the middle of the road. Additionally, at least one case was found where businesses had been listed – another category of over-coverage.

- At phase 2, a total of 74 new households were found at units that appear not to have been listed at phase 1, representing 2.1% of all households identified at phase 2. Potentially, this suggests that 2.1% of households might be omitted by the listing procedure (under-coverage). However, it is also possible that many of these households were in fact at units listed at phase 1, but the interviewer failed to correctly record the link between the phase 1 and phase 2 units. Thus, this figure of 2.1% should perhaps be viewed as a *maximum* likely level of under-coverage. In other words, we should be confident that *at least* 97.9% of households will be included using this sampling method. Many of the excluded households were living in basements, attics, out-buildings and other places that were not obviously residential dwellings.

- Some addresses were not recorded/described in sufficient detail to allow unambiguous identification in the field.

The time needed to carry out the listing had been something of an unknown prior to the field test and a concern to the statistical institutions. Information on this was recorded by the field workers and provided some reassurance that the task was not excessive. In most areas the listing (phase 1) had been carried out on a single visit to the area. There were two areas where multiple visits over 3 or 4 days had been needed. These areas both contained a lot of newly-built properties. But across the other 48 areas included in the test sample, the mean time spent on the listing was just under six hours.

No problems were encountered in data entry. When checked against a sample of forms, the quality of data entry was seen to be high.

The field test also permitted comparisons of the number of households recorded in each EA in the 1991 census with the number listed in 2003. This was not of direct relevance to the proposed sampling method but provides some indication of the likely suitability of census counts as auxiliary data for estimation. There were 8 EAs where the number of households found by the field test was less than 67% of the number recorded in the 1991 Census and 10 where the number was more than 150% of the Census number. There was one extreme case where the number had

risen from 10 to 123. The overall correlation between the two measures was 0.38 (Pearson's correlation coefficient). Our prior expectations of a low correlation, for the reasons described in section 1, were therefore confirmed.

6.3 CONCLUSIONS OF THE FIELD TEST

The proposed approach to sampling for general population surveys appeared to be feasible. A number of minor improvements to the procedures were identified and would be implemented for the HBS:

- Identification of area boundaries by interviewers: It was decided to have cartographers draw up new maps where necessary for areas sampled for HBS, showing the existing boundaries in relation to current features.

- Over-coverage: The need to include only the dwellings on one side of the road, in cases where the EA boundary runs along the road, would be stressed in the training for the HBS fieldwork.

- Under-coverage: The need to be vigilant and comprehensive in terms of including potentially occupied basements, attics, out-buildings and other places that were not obviously residential dwellings would be stressed in the HBS instructions and training.

- Clear identification of dwellings: This would be addressed by improved training, though it was recognised that it was not always easy unambiguously to describe a dwelling, especially in rural areas where there is often no address system (numbers or street names).

Implementation of these amendments could be expected to result in a slight decrease in the average number of listed units per EA, but probably a slight increase in the overall proportion of unoccupied dwellings.

7. Household Budget Survey: Sample Design

The sampling frame of EAs was as shown in Table 4. It was agreed between all parties that 691 was an appropriate number of EAs to include in the HBS sample in order to provide accurate state-wide estimates. This number represents a compromise between various competing pressures, primarily those of budget *versus* statistical precision. Proportionate distribution of the 691 EAs over the 4 strata would have resulted in only 13 EAs being selected in Brčko. It was felt essential that at least basic HBS estimates should be available separately for Brčko, so ISTAT, the funders of the HBS data collection, decided to fund the inclusion of an extra 19 Brčko EAs in the sample.

Thus, the sample of EAs was selected independently from each of two explicit strata in July 2003. Stratum 1 consisted of the whole of BiH excluding Brčko and 678 EAs were selected from this stratum. Stratum 2 consisted of Brčko, 32 EAs being selected from this stratum.

In stratum 1, EAs were implicitly stratified (sorted) by settlement (numeric order of standard settlement identifier) within urban/rural within municipality within FBiH/RS/Both (see Table 4), prior to systematic selection of the sample EAs using a random start and fixed interval. Such implicit stratification should improve the precision of HBS estimates in so far as HBS measures are correlated with municipality, urban/rural and settlement. In stratum 2, EAs were implicitly stratified by settlement within urban/rural prior to systematic selection.

The resultant distribution of the sampled EAs over strata and entities is summarised in Table 5, where the effect of over-sampling Brčko can be seen.

It turned out that of the 45 EAs sampled from the "both entities" sub-stratum, 33 were in fact in FBiH and 12 were in RS. Additionally, one of the EAs selected from the FBiH sub-stratum was in RS. Thus, fieldwork would be carried out by FBiHIS in 398 EAs, RSIS in 280 and SODB in 32.

Almost all municipalities in BiH were represented in the sample. Only a couple of the smallest ones did not get selected at all. This contrasts with the LSMS sample, which was concentrated in just 25 municipalities.

Up-dated maps of a number of selected EAs were created during August 2003.

Table 5: HBS sample of EAs, by stratum and entity

Stratum	Entity	EAs in population		EAs in sample	
		Number	%	Number	%
1	FBiH only	10,287	53.0	366	51.5
	RS only	7,511	38.7	267	37.6
	Both FBiH <u>and</u> RS	1,254	6.5	45	6.3
2	Brčko	371	1.9	32	4.5
Total		*19,423*	*100.0*	*710*	

The field listing stage took place in August and September. A team of supervisors from each of the statistical institutions was briefed at two briefing conferences held in Sarajevo and Banja Luka in the last week of August. Each supervisor was responsible for a small number of interviewers and held a local briefing meeting for the interviewers in their area. Listing was completed by 12 September. All data on the listing forms – including names of householders and descriptions of dwellings – were keyed: data entry was completed by 3 October.

The total number of households listed was 39,081. The mean number of dwellings/ households listed per EA was rather lower for the HBS (55.0) than on the field test (75.0). This was felt to be partly due to a reduction in over-coverage (interviewers no longer including houses on the wrong side of roads that form EA boundaries) and partly due to a reduction in the numbers of empty dwellings listed, due to interviewers making greater efforts to establish occupancy for all dwellings. The mean number of units listed per EA varied between the 3 statistical institutions: 57.7 in FBiH, 50.8 in RS and 59.2 in Brčko.

ISTAT decided that an initial sample of 9,570 dwellings was required for the HBS. This includes allowances for ineligibles and for non-response, with a target of 6,000 fully co-operating households. The sample of households was selected systematically from throughout the entire list, resulting in a distribution that reflects the distribution on the complete list (Table 6). The final column of Table 6 shows the estimated distribution of all households in BiH, obtained by weighting the listed numbers by the inverse of the relative selection probabilities.

Table 6: Listed households and the HBS sample of households, by entity

	Sampled EAs	Listed households	Sampled households	%	Estimated population distribution %
FBiH	398	22,963	5,623	58.8	60.5
RS	280	14,223	3,483	36.4	37.5
Brčko	32	1,895	464	4.8	2.0
Total	*710*	*39,081*	*9,570*	*100.0*	*100.0*

The selected sample of 9,570 households will be systematically divided into 12 monthly samples. HBS field work is planned to begin on 1 December 2003 with the final monthly sample being issued to the field on 1 November 2004.

It is worth noting that for the HBS design, the design effect due to variable selection probabilities as estimated by (7) is 1.02 (considerably lower than the 2.13 reported in section 3.2 for LSMS). The design effect due to clustering should also be relatively modest. Of the sampled EAs, 25 turned out to contain 0 households (15 in FBiH and 10 in RS), so a responding sample of 6,000 households would consist of a mean of 8.76 households per EA. The design effect due to clustering is:

$$D\hat{e}ff_{CL} = 1 + (\bar{b} - 1)\rho \qquad \text{- (13)}$$

Thus, for a variable with $\rho = 0.05$, for example, $D\hat{e}ff_{CL} = 1.18$ (as $\bar{b} = 8.76$).

8. Conclusions

A practical and statistically efficient approach to sampling the general population of households in BiH has been developed. It has a number of noteworthy features that distinguish it from approaches in common use. These have arisen from necessary adaptation to the conditions in BiH. They may be applicable in other situations where:

- There are no sampling frames of households;

- There are no population counts for small areas;

- There is considerable population movement;

- There exists a comprehensive and mutually-exclusive set of small areas for which boundaries can be identified on the ground.

The noteworthy features are:

- Use of census EAs as PSUs, even though census data are out-of-date and cannot be used to inform the sample design;

- Use of equal probabilities both at the stage of selecting PSUs and at the stage of selecting households within PSUs. This contrasts with the more common approach of selecting PSUs with PPS and then selecting a fixed number of households per PSU (i.e. probability inversely-proportional to size at the second stage) (e.g. Kopecky *et al*, 2002). Unlike the standard 2-stage PPS design, which results in a fixed sample size per EA, the BiH design results in a sample size proportional to the population number of households in the EA;

- The frame of EAs does not require frequent up-dating;

- The listing procedure can be carried out immediately prior to the field work for any particular survey, ensuring that the frame of households is as up-to-date as possible.

Finally, it should be noted that the approach offers some flexibility in the distribution of the sample of households over geographic areas. For example, it has been suggested that a future Labour Force Survey (LFS) in BiH might require Cantons to be sampled at different rates in order to provide a minimum sample size per Canton (FBiH is divided for administrative purposes into 10 Cantons; there is no system of Cantons in RS, but geographical sub-divisions are likely to be of interest there too). This would simply require a minor adaptation to the design used for the HBS, whereby Cantons, identified as a set of municipalities, form the explicit sampling strata. This is exactly analogous to the over-sampling of Brčko for the HBS.

References

Cigar, N. (1995). *Genocide in Bosnia: The Policy of "Ethnic Cleansing".* Dallas: Texas A&M University Press.

Elliot, D. (1991) *Weighting for Non-response: A survey researcher's guide.* London: OPCS.

Glenny, M. (1996). *The Fall of Yugoslavia: The Third Balkan War.* New York: Penguin.

Holbrooke, R. (1998) *To End a War.* New York: Random House.

Kish, L. (1965) *Survey Sampling.* New York: John Wiley and sons.

Kopecky, M., Makalous, I. And Pavlickova, Z. (2002) The Labour Force Survey in the Czech Republic, *Statistics in Transition*, 5:4, 541-554.

Malcolm, N. (2002). *Bosnia: A Short History*. London: Pan Books.

Office of the High Representative of Bosnia and Herzegovina (2002) *Decision of the High Representative no. 57/02 of 21 October 2002: Law on Statistics of Bosnia and Herzegovina.* OHR: Sarajevo.

Papić, Ž (2001) The general situation in B-H and international support policies, chapter 1 in *International Support Policies to South-East European Countries: Lessons (not) Learned in B-H* (ed.s Ž. Papić *et al*), Sarajevo: Müller.

Rieff, D. (1995) *Slaughterhouse: Bosnia and the Failure of the West.* New York: Simon and Schuster.

UNHCR (1998) *Populations of Concern to UNHCR – 1997 Statistical Overview.* Geneva: UNHCR.

UNHCR (2003) *Returns Summary to Bosnia and Herzegovina from 01/01/1996 to 31/07/2003.* Sarajevo: UNHCR Sarajevo.

World Bank (2003) *Basic Information Document: Bosnia-Herzegovina Living Standards Measurement Study Survey 2001.* Available at http://www.worldbank.org/lsms/country/bih/docs/bihbinfoe.pdf.

Appendix: Field Test of Sampling Procedures
Instructions for Interviewers and Listing Forms

The following instructions are very important, please read them carefully before you begin your work. Your work consists of two distinct stages.

First Stage

The Map: Before you begin working in the field you must study the map of your Enumeration Area (EA) and become familiar with the boundaries of your EA. What the map shows in terms of buildings is not important, your job is to list whatever you actually find on the ground. The purpose of the map is ONLY to define the boundary of the area to be listed. It is not necessary to update the map with new buildings you encounter during this work.

Number of trips for Stage 1: The first stage must take place with ONE trip to the EA. In most EA's this is about half a day's work but in some larger EAs a whole day maybe necessary. The first stage must be completed within 7 days of you receiving your instructions and documents. During this one visit you should try and make contact with every dwelling/household[1] in the EA. In some cases it may not be possible to make contact with anyone. If this is the case just list the dwelling, if it looks like it is occupied. Do not list dwellings that are known to be unoccupied, but DO list dwellings for which you are not certain whether they are occupied.

If you do make contact with someone, ask if the property is occupied (if you have any doubt) and how many households live at that dwelling. If there is more than one household at the dwelling list each household separately (see Household Numbers 2 and 3 in Example 1).

It may sometimes be the case that you cannot make contact with anyone in a dwelling but a neighbour can provide you with this information, this is acceptable. If you have a particularly large EA and it does not look like you can visit all dwellings in one day it is more important to list every dwelling rather than making contact. Please record the number of hours spent during this one 1st stage visit on your "Codes to transfer to your 1st and 2nd Stage" Form.

If you do make contact with someone at Stage 1 <u>do not be tempted to gather the other information that needs to be collected during the second stage</u>. It may seem more effective to work like this, but for the purpose of the pilot it is vital that this does not happen. The pilot is a scientific experiment and the procedures must be undertaken precisely.

---

[1] **Dwelling** - a single house or flat, or a residential space that appears to have been intended for occupation by a single household. EXCLUDE Prisons, Military Barracks, Schools, Hospitals, Old Peoples Homes, Homes for the disabled, Shops, Hotels, Museums, Public Buildings, governmental, Factories, businesses

Household - the social unit comprised of one or a group of persons who join together to share shelter and food. In other words, the households is a group of persons who normally reside in the same dwelling unit or in part of it (live under the same roof), who are joined or not by blood ties, and who cook together (eat from the same pot).

When making contact in the household it is important that you:

- show your Identification Card and the letter
- do not take anyone else with you when you approach a household.

Codes: You can see from Example 1 that the First Stage form is a list where you record one dwelling/household on each line. The date, your interviewer code, entity code (1 For RS and 2 for FBiH, 3 for District of Brcko), municipality code and EA code must be entered on every completed page. The municipality and EA code for each enumeration area are found on the document "Codes to transfer to your 1st and 2nd Stage Forms".

Household Number: The first household you list is Household Number 1. Increment the Household Number by 1 until you have listed all the dwellings in your EA (see Example 1). You must list each dwelling that is, or may be, occupied. If you have any doubt whether a dwelling is occupied it is recommended that you knock on the door or ask neighbours to be certain.

Blocks of flats: List each flat separately (see Household Numbers 4,5, 6 and 7 in Example 1). If you find out that there are multiple households in a particular dwelling, each household must be listed separately (ideally, with family name or some other details in the 'other information' box to clarify the situation (see Household Numbers 2 and 3 in Example 1). If you do not know how many households live in a dwelling, just list the dwelling once.

For each dwelling you need to write the number of the dwelling, street address and postcode, plus any other information if needed, for example a description of where the dwelling is if it is difficult to locate (see Household Number 8 in Example 1).

Rural areas: If there is no street numbering system, house numbers or names you should describe the dwelling in as much detail as you can (ideally including a family name). To begin the listing chose a dwelling with strong distinguishing features. Describe other houses in relation to the previous one on the list (see Household Numbers 9 and 10 in Example 1).

When you have listed all the households/dwellings you must send the first stage list back to the Central Office. You cannot begin the second stage work until it has been acknowledged that your first stage form has been received.

Second Stage

Aim: The aim of the second stage is to obtain some basic information about EVERY household living at a dwelling that you listed at the first stage. The purpose of this is to enable us to assess how useful a listing exercise such as that undertaken at stage 1 (i.e. in a single visit) could be as a means of sampling households.

Number of Trips: During the second stage it will be necessary to make several trips to gather all the information necessary. During the first visit of this second stage you must attempt to contact EVERY household. At the second visit you will only be attempting contact with households where you did not get any information at all or where you only have proxy information (from a neighbour etc) and there is some degree of uncertainty. We would like you to record the number of trips made to the EA and the amount of time during each trip on your "Codes to transfer to your 1st and 2nd Stage" Form.

Length of trips: The number of households you attempt to contact will fall at each visit. Some households may need only one visit, other may need 5 or 6 attempts to successfully gather the information. Information from neighbours can be attempted after the third call at a dwelling. But if the information obtained from neighbours involves any uncertainty, you should continue to try to contact the actual dwelling/household**.   To minimise the number of calls it is strongly recommended that the first visits are made at weekends and in the evenings.**

Each household/dwelling has a separate Second Stage form. In others words, there is one second stage form for each 'unit' (dwelling or household) that was given a household number at stage 1.

During this second stage you should ideally gather information from the Head of Household[2]. If this is not possible, from someone else who can give you reliable details. The information gathered is:

- Is the dwelling/household occupied or unoccupied
- Number of households at the dwelling;
- Family name;
- Given name(s) of head or other reference person;
- Number of persons in household, split into under 18 males and females, over 18 males and females.
- Code for who provided the information (Head of Household, other household member, neighbour or other non-household member)

During the training a number of examples will be given to you and explained so that the completion of 1st and 2nd stage is completely clear to you.

---

[2] **Head of Household**: is the person the household designates as such, regardless of the reason (age, decision making, earnings, tradition, etc). The household members themselves based on their own criteria will identify the head of household. Most often, but not always, it will be the main provider for the household and someone who is familiar with all the activities and occupations of the household members. The head of household can be male or female.