LSMS
Living Standards
Measurement Study
Working Paper No. 129

Poverty Comparisons and Household Survey Design

Steven Howes and Jean Olson Lanjouw

Poverty Comparisons and Household Survey Design

The Living Standards Measurement Study

The Living Standards Measurement Study (LSMS) was established by the World Bank in 1980 to explore ways of improving the type and quality of household data collected by statistical offices in developing countries. Its goal is to foster increased use of household data as a basis for policy decisionmaking. Specifically, the LSMS is working to develop new methods to monitor progress in raising levels of living, to identify the consequences for households of past and proposed government policies, and to improve communications between survey statisticians, analysts, and policymakers.

The LSMS Working Paper series was started to disseminate intermediate products from the LSMS. Publications in the series include critical surveys covering different aspects of the LSMS data collection program and reports on improved methodologies for using Living Standards Survey (LSS) data. More recent publications recommend specific survey, questionnaire, and data processing designs and demonstrate the breadth of policy analysis that can be carried out using LSS data.

LSMS Working Paper
Number 129

Poverty Comparisons and Household Survey Design

Steven Howes and Jean Olson Lanjouw

The World Bank
Washington, D.C.

Manufactured in the United States of America
First printing April 1997

To present the results of the Living Standards Measurement Study with the least possible delay, the typescript of this paper has not been prepared in accordance with the procedures appropriate to formal printed texts, and the World Bank accepts no responsibility for errors. Some sources cited in this paper may be informal documents that are not readily available.

The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of the data included in this publication and accepts no responsibility whatsoever for any consequence of their use.

The boundaries, colors, denominations, and other information shown on any map in this volume do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

ISBN: 0-8213-3862-5
ISSN: 0253-4517

Steven Howes is an economist at the World Bank; this paper was written while he was in the Poverty and Human Resources Division of the Bank's Policy Research Department. Jean Olson Lanjouw is Assistant Professor in the Department of Economics at Yale University.

Library of Congress Cataloging-in-Publication Data

Howes, Stephen R., 1964–
Poverty comparisons and household survey design / Stephen R. Howes and Jean Olson Lanjouw.
p. cm. — (LSMS working paper, ISSN 0253-4517 ; no. 129)
Includes bibliographical references.
ISBN 0-8213-3862-5
1. Poverty—Statistical methods. 2. Household surveys—Statistical methods. I. Lanjouw, Jean Olson. II. Title. III. Series.
HC79.P6H68 1997
339.2'2'—dc21 96–53397
CIP

Table of Contents

Foreword

The ability to monitor poverty is crucial to assessing the success of policies designed to improve standards of living. With the LSMS household surveys and others now available, many developing countries now have the data-base required to undertake this policy monitoring. Earlier LSMS papers have shown how to approach the measurement of poverty statistically, so as to be able to distinguish real changes from sampling variation. This paper extends the earlier work to show how to take into account typical sample designs in calculating statistical measures of poverty change. It uses LSMS data sets both to show how household surveys differ greatly from the "simple random sample" paradigm and to illustrate the importance of basing statistical formulae on the actual sample design used.

Lyn Squire, Director
Policy Research Department

Abstract

Poverty comparisons - an increasingly important starting-point for welfare analysis - are almost always based on household surveys. They therefore require that one be able to distinguish underlying differences in the populations being compared from sampling variation: standard errors must be calculated. So far, this has largely been done on the assumption that the household surveys are simple random samples. But household surveys are more complex than this. We show that taking into account sampling design has a major effect on standard errors for well-known poverty measures: they can increase by around one-half. We also show that making only a partial correction for sample design (taking into account clustering, but not stratification, whether explicit or implicit) can be as misleading as not taking any account at all of sampling design.

Acknowledgments

We would like to thank, for their provision of data, information and/or comments: Benu Bidani, Gaurav Datt, Mark Foley, Paul Glewwe, Margaret Grosh, Dean Jolliffe, Peter Lanjouw, Martin Ravallion, Chris Scott, Kinnon Scott, Salman Zaidi, and Qing-hua Zhao. Mr Ranzam of Pakistan's Federal Bureau of Statistics also provided us with useful information.

I. Introduction

Has poverty increased or fallen? Is urban or rural poverty higher? Will some policy under consideration reduce or increase poverty? These are typical of the questions asked in poverty analyses. To provide answers, recourse is required to household surveys. But a survey is not a census. It is a sample, with a size typically numbering in the thousands of households, from which conclusions concerning populations typically numbering in the millions must be drawn. This leads to the fundamental problem that any comparative analysis must distinguish population differences from sampling variation. A series of recent papers have stressed the importance of this and have provided the tools by which standard errors can be calculated (Howes, 1993, Kakwani, 1993, Pudney and Sutherland, 1994, Ravallion, 1994). The problem with the current state of play is that, in presenting statistical methods and results for use in poverty comparisons, the assumption has been made that the household surveys being analyzed are simple random samples of the populations from which they are drawn. In fact, however, they are not. Household surveys are far more complex in their design.

This can best be seen by analogy. Consider each household in the population to be represented by a number written on a piece of paper. All pieces of paper are of equal size and are placed in a hat. Then a household survey would be a simple random sample if it were selected by blindly drawing numbers from the hat. Household surveys differ in a number of ways from this simple model:

- One may have many hats from which sub-samples are drawn: often populations are first divided into strata, each of which may be considered a separate hat or sub-population.
- There will probably be "hats within hats". A random selection of clusters, such as villages, is invariably first made from the population (or from each stratum). Households are then randomly drawn from these smaller clusters.
- Some numbers (households) have a higher probability of selection than others.
- The selection of numbers may not be "blind", that is, random. Instead, it may be systematic: the numbers may be lined up and every *n*th one chosen.

As we will see, this is only the start of a fairly long list of complexities which household surveys incorporate. What are the implications for statistical poverty (and welfare or inequality) analysis of these various features? This is the central question which this paper addresses. We present estimators of the variance of poverty measures appropriate for typical survey designs. And we assess the influence departures from a simple random sample are likely to have on the precision with which poverty estimates can be made. Which departures have a substantial impact and which can be safely ignored for the sake of convenience? We show that, under sample designs commonly in use, conventional formulae may lead to estimates of standard errors for poverty measures which are only two-thirds the size they should be. That is, ignoring sample design can make us think estimates are considerably more precise than they actually are.

It should be noted at the outset that the key results we make use of - relating to the variances of sample means - have been known since the fifties, and are presented in several

textbooks on the subject of sample design.[1] However, we have not found any work which links the general results available for complex survey designs to the typical features of household surveys, let alone to poverty analysis. Moreover, the fact that these general results have been almost completely overlooked in the empirical and theoretical literature on poverty measurement suggests that there is a need to set out clearly the formulae required and to provide a strong motivation for their use.[2]

In the next section, we provide a more formal and detailed treatment of the various ways in which household surveys can differ from the simple random sample model. In Section 3, we provide the basic formulae. Section 4 applies these formulae to poverty measures. Section 5 gives some examples of the importance of taking into account sample design. Section 6 concludes. The appendix provides proofs of the paper's key results.

---

1. Kish (1965) is the classic on this subject. It gives what is still probably the most comprehensive treatment, though not the simplest. Som (1973) provides a very clear presentation of results. Hansen, Hurwitz and Madow (1953) provide proofs. Levy and Lemeshow (1991) provide an introduction, as does Scheaffer, Mendenhall and Ott (1990).

2. Deaton (1994) provides an excellent introduction to and analysis of household surveys and to some extent fulfills these two aims. However, his treatment does not cover many of the common problems raised by the use of household surveys (such as when one has stratification and clustering, or raising factors and clustering, or all three, or when one is estimating per capita (rather than household) means). Scott and Amenuvegbe (1989, pp.55-57) cover - in relation to a survey of Mauritania - the joint use of clustering and (implicit) stratification (not raising factors). However, their intention is to provide only formulae, and they add neither motivation nor explanation. Rodgers and Rodgers (1992) use a alternative method to that provided in this paper (balanced repeated replications - see footnote 16), but provide only results. In general it would seem to be that poverty analysis of developing countries pays less attention to sample design and its implications than analysis of developed countries (see Rodgers and Rodgers, 1992, and Duncan and Rodgers, 1991, as examples for the United States). Incorporation of sample design also seems to be much more widespread among demographers than economists (see Cleland and Scott, 1987).

2. Household survey designs

In this section, we discuss eight sampling features one needs to be aware of when analyzing household surveys. We then (in 2.9) present examples showing how different surveys incorporate various of them. This section, like the rest of the paper, draws most of its examples from surveys conducted as part of the World Bank Living Standards Measurement Study (LSMS). This is simply because we are more knowledgeable about these surveys' designs than others. However, LSMS surveys do present a range of household survey designs. Since each survey in the series is carried out in conjunction with the statistical bureau of the country in which it is being conducted, different LSMS surveys incorporate different designs, depending on prevailing practice in the countries concerned.

2.1. Clustering

One feature which most household surveys share is that they are clustered. That is, the first selection (from the population or *sample frame*) is not of households, but of some higher level units such as villages or street blocks, known as clusters. This is the case for all nationwide household surveys, though some very small surveys (of one or several villages, say) are not clustered. As we will see, clustering leads to higher variances. Its justification is purely practical. By concentrating sampled households in a small number of geographical areas, clustering drastically reduces survey costs per household.

Under some sample selection procedures, a cluster can be selected more than once: that is, more than one group of households, say, can be selected from a single cluster. To avoid confusion, we refer to each selection of a group of households (or "ultimate sampling units" - see 2.3 below) from the cluster as a 'cluster take'.

2.2. Stratification

Many surveys are stratified, typically into geographical regions, such as urban/rural and provincial, but also by other characteristics. The difference between strata and clusters is simply explained. Both strata and clusters divide the sampling frame exhaustively and exclusively. If both are present, the clusters sub-divide the strata. All strata are included in the sample (each with its designated sample size), but only a selection of clusters are included in the sample. Stratification is very common, but not universal. Stratification with equal sample rates in the strata ensures a more representative sample overall, and so reduces variance. It also can also be used to ensure that one obtains sufficient observations from small sub-populations of interest.

Note that what we call 'stratification' here is sometimes referred to as 'explicit stratification' to distinguish it from implicit stratification, a form of systematic sampling discussed in 2.7. We analyze both forms of stratification, but when we use the term 'stratification' without qualification we are referring to the explicit variety.[3]

2.3. Ultimate sampling unit

The ultimate sampling unit is the smallest level of population unit sampled by the survey. For most household surveys, the ultimate sampling unit is the household. That is, after first selecting clusters, and then after possibly some intermediate selection stages (see 2.4), the final selection of elements is of households. But this is not universal. In the Nicaragua LSMS, the ultimate sampling unit was groups of five households: each cluster was divided into groups of five households (on the basis of geographical proximity) and a selection of these groups, rather than of individual households, was made. Of course, by selecting groups of households one is selecting individual households. However, the ultimate sampling unit is the lowest level at which sampling occurs (and hence is the unit in which sample size is measured). In the Nicaraguan case, there was no sampling below the group level: all households within any selected group were chosen.[4]

2.4. Number of random selection stages

The selection process of many samples is two-stage. That is, once clusters have been selected, the ultimate sampling units, typically households (see above), are selected directly from the clusters. However, more than two stages may also be used, especially in large countries. For example, in the Russian Longitudinal Monitoring Survey, *raions* (regions) provide the first-level clusters. Then voting districts are selected from the chosen *raions*. These serve as second-stage clusters. In a third stage, households are chosen from the selected voting districts. Note that the reference here is to the number of sampling stages, so stratification is never regarded as the first stage.

---

3. About half of the World Bank LSMS surveys are explicitly stratified. As far as we know, all the rest are implicitly stratified.

4. Compare the typical case in which the final sampling is of households. Then households are the ultimate sampling unit, and not individuals, even though by selecting households one is selecting individuals. By contrast, if one samples from a list of individuals or if one samples individuals within households (as in the case of some fertility surveys), then individuals do indeed become the ultimate sampling units.

2.5. Unequal probabilities of selection

Many household surveys are not self-weighting. This means that some households have a higher chance of being selected than others. Variable weights (known as *raising* or *expansion factors*) have to be used to prevent estimators being biased as a result. Formally, raising factors can be defined as a set of weights such that the weighted sum of the sample observations of a given variable is an unbiased estimator of the population total of the variable. There are three main reasons why a survey may not be self-weighting. First, when strata are used, the sample may not be distributed over the strata in accordance with the distribution of the population. Instead *disproportionate stratification* may be used, and some areas deliberately over-represented. These may be areas, typically urban, in which sampling is cheaper or they may be small sub-national political units, such as small provinces, for which one wants to ensure a minimum sample size. Second, even if the survey is intended to be self-weighting, it can end up not being so, owing, for example, to non-response. Finally, a common method of selecting clusters - discussed immediately below - often leads to the sample being non-self-weighting.

2.6. Selection of clusters according to size or estimated size

Some clusters are bigger than others in terms of the number of ultimate sampling units (typically households) they contain. It is a widespread practice to give larger clusters a higher probability of selection. If exact cluster sizes were known, then selection could be by *probability proportional to size (pps)*, which has the advantage of allowing an equal number of households to be drawn from each cluster at the second-stage with the sample remaining self-weighting (at least within each stratum). With household surveys, however, the exact size of a cluster at the time of sampling is generally not known prior to sampling. The most recent count may be the last census, which may be a decade or even more out of date.

If the discrepancies are expected to be small, they can be ignored and pps can be used (this was done in the Russian Longitudinal Monitoring Survey, for example). Alternatively, the discrepancies can be formally dealt with. In this case, one explicitly recognizes that one is using a proxy for size - typically the out-of-date census but possibly something as weakly correlated with cluster size as area - and is choosing clusters on this basis. This is known as selection by *probability proportional to estimated size (ppes)*. Under ppes, selected clusters are re-enumerated (re-listed). The new list (typically of households in the selected clusters) not only provides the basis for the selection of the ultimate sampling units. It also gives the desired exact measure of size of the selected clusters. The ratio of actual to proxied size can be used in two ways. If the size of the cluster take is held fixed, the ratio can be used to re-weight observations (to give, for example, greater weight to those observations from clusters whose actual exceeds their proxied size, and which are thus under-represented in the survey). Alternatively, the size of the cluster take can be adjusted up or downwards (from some average size) by this ratio. This has the disadvantage of resulting in differently-sized cluster takes, but the advantage of removing the need for cluster weights (see footnote 11 for a formal statement of this). Both methods allow

unbiased estimates to be derived from a survey even if only very rough or out-of-date information on cluster sizes is available. Both methods are commonly used and both require the same statistical treatment. Note that pps can be regarded as a special case of ppes in which estimated size equals, or is assumed to equal, actual size.

2.7. Systematic or random sampling

Systematic sampling is selection at a fixed interval, say s, from a list of units, starting from a random point not more than s units from the start of the list. Its usage in household surveys, at all stages, is widespread.

At the final sampling stage, its usage would seem to be universal. What happens here is that ultimate sampling units (say households) from the chosen clusters are listed. Then a sampling interval, s, equal to the total number of households in the cluster list divided by the desired cluster sample size, is defined. Then a random number is chosen less than or equal to s. That household is chosen, as is every sth household from then on until one reaches the end of the list.[5]

With regards to the choice of clusters (the first stage), the situation is more complex. In some textbooks (e.g., Scheaffer, Mendenhall and Ott, 1990, p.79), pps and ppes are defined to be random. However, in practice, in many surveys, systematic sampling is used at the first stage. These two methods work as follows. In both cases, all clusters in a stratum are placed in a list and their (estimated or actual) sizes, M_c, are cumulated, giving a series with typical element $R_i=\Sigma_{c=1}^{i}M_c$. Say one wants to choose n clusters from a total of N. Using random sampling, one would simply pick n random numbers between 1 and R_N. Using systematic sampling, one would define the sampling interval, $s=R_N/n$. A random number, r, less than s would be chosen and the list of numbers $\{r, r+s, r+2s, \ldots, r+(n-1)s\}$ compiled. Both methods - random and systematic - result in a list of n numbers, each of which picks out a cluster using the rule that a number in the range $(R_{i-1},R_i]$ picks out the ith cluster. Both methods make it possible to have clusters chosen more than once.

In interpreting systematic sampling, the ordering of the elements in the sample frame (list of clusters or households) is very important. There are two cases. In the first, the ordering of units is random with respect to the variables of interest: an example might be an alphabetical listing (though one can think of exceptions). In the second and more common case, the ordering is non-random: often neighboring clusters or households are placed next to one another. This is sometimes referred to as *implicit stratification* (see 1.2). The implications of these two methods of ordering are very different and are discussed in the next section (see 3.5).

---

5. For simplicity, we assume that s is an integer. See Kish (1965, pp.115-116) for details of the procedure to be followed when it is not.

Systematic sampling is popular because it is easier to implement than random sampling. Under systematic sampling, only the starting point of the sample selection need be generated randomly. After that, no more attention need be paid to random number generation. In addition, the implicit stratification method generally leads to a more representative sample, and so reduces variance (though it can also make variance estimation difficult - again see 3.5).

2.8. Sampling with or without replacement

Say one has random (as against systematic) sampling, and a household or cluster is removed from the sample frame once it has been selected in the sample, so that no unit can be selected twice. This is known as (random) sampling without replacement.[6] If the household or cluster is replaced, one has (random) sampling with replacement. (The random ppes sampling described in 2.7 is with replacement.) Since the with/without replacement issue only applies to random sampling, and since most household surveys use systematic sampling, the relevance of the distinction may seem unimportant. It is, in fact, very germane, as, in presenting formulae, we will be forced to approximate systematic by random sampling. The issue is discussed further in 3.5.

2.9. Examples

Table 1 gives some examples of some recent household surveys from various developing countries. All of them are clustered and use systematic sampling, but otherwise a diversity of features is present. Table 1 also presents the features which an imaginary household survey conducted as a simple random sample (srs) would possess to highlight the many differences between the srs and real-world designs. It should be noted that the complexity of surveys is not necessarily exhausted by the detail of Table 1. As notes 3 and 4 to the table illustrate, many surveys have methods of selection specific only to that survey.[7]

---

6. In practice, random sampling without replacement implies that once a random number has picked out a household or cluster any subsequent random numbers which identify the same element are ignored.

7. The Jamaican Survey of Living Conditions (JSLC, described in PRDPH, 1994) provides another illustration of the individual features which surveys can have, for which there is no solution without approximation. The survey is drawn from the Jamaican Labor Force Survey (LFS), a stratified, clustered, two-stage sample. The JSLC is based on a selection of the LFS's strata, turning these strata into new first-stage clusters. Yet the proportion of LFS strata chosen is high: 100% for one year for some parts of the country, and one- or two-thirds for other years and other parts. How is one to treat these LFS strata? Where less than 100% of them are chosen, they are, strictly speaking, clusters. Yet with the selection rate so high, the assumption of sampling with replacement would give a poor approximation. The without-replacement formulae, however, make informational demands which can not be met with the survey data publicly available (see 3.5). One alternative is to regard the survey as one not of all Jamaica, but of those LFS strata surveyed, and treat the LFS strata as JSLC strata. Applications of survey findings

Table 1 Features of sample design from some recent national household consumption surveys, and a comparison with simple random sampling

Sample	Clus-tered	Strat-ified	Ultimate sampling unit	Number of stages	Self-weighting[1]	Method for selection of clusters	Use of systematic sampling at first stage/second stage	If random sampling, with or without replacement
Simple random sampling	No	No	Household	One	Yes	N/A	No (only one stage)	Either[2]
Pakistan Integrated Household Survey 1991	Yes	Yes	Household	Two	No	ppes	Yes / Yes	N/A
Ghana Living Standards Survey 1987-88	Yes	No	Household	Two	Yes	ppes[3]	Yes / Yes	N/A
Nicaragua LSMS 1993	Yes	Yes	Groups of 5 households	Two	No	ppes	Yes / Yes	N/A
Russian Longitudinal Monitoring Survey 1993	Yes	Yes	Household	Three[4]	Yes	pps	Yes / Yes	N/A

Notes: 1. Most surveys are not self-weighting on account of differential non-response. In the table above, however, we are concerned with whether the surveys are intended to be self-weighting or not.

2. Some authors define srs to be sampling without replacement. Others allow it to be either with or without replacement. We follow the latter convention.

3. The Ghana survey used ppes, adjusting the size of the cluster take on the basis of the ratio of actual to estimated cluster size. However, an integer restriction was used so that cluster takes were either nought, once, twice (etc.) times the target cluster take size (in this case, 16 households). See Scott and Amenuvegbe (1989, 1991) and Section 5 of this paper for further discussion.

4. Most strata have three stages, a couple have only two.

---

to Jamaica as a whole would then depend on the (probably perfectly acceptable) judgement that the survey's sample frame differed little from the country as a whole.

3. Estimators of totals and means and their variances appropriate for complex survey designs

Before turning to poverty measures, we consider, in this section, the estimation of means of variables. We focus on means not because they are attractive welfare measures, but on account of their simplicity. As we show in Section 4, the estimation of poverty and welfare indices follows on as a direct analogue from the estimation of simple means.

Since a mean is the ratio of two totals, we begin with the estimation of (consumption or income) totals. The complexity of survey design is very easily incorporated into the calculation of unbiased estimators of population totals. Suppose that we have a sample of N households, subscripted by j, each with income of y_j. To each of the households is attached an expansion factor or weight, w_j. Given the correct choice of weights, an unbiased estimator of τ, the total income of the population, can be written

$$t = \sum_{j=1}^{N} w_j y_j \tag{1}$$

If the sample is self-weighting, then w_j is constant for all households. Using weights in this way is both simple and is commonly done, at least implicitly, to obtain means (dividing (1) by the sum of w_j).[8]

Matters are more complex when estimating variances. Now one must consider not just the use of expansion factors but rather all the departures from simple random sampling implied by the sample design. To make any progress, one must therefore be explicit about the type of sample design being dealt with. We assume (for the moment) that the survey being analyzed:

(i) is clustered,
(ii) uses ppes for the selection of clusters (with pps as a special case)
(iii) has any number of strata (with one stratum equivalent to no stratification),
(iv) has two stages,
(v) has households as the ultimate sampling unit,
(vi) was selected by random sampling with replacement at the first stage, and by systematic or random sampling with or without replacement but with equal probability of selection at the second stage, and,
(vii) is self-weighting or has cluster-level expansion factors due to the use of ppes and/or disproportionate stratification.

8. Typically, if the sample is self-weighting, one does not have the required expansion factors. This means totals cannot be estimated. Means, however, can be, and still using the method given in this paper. Since, if the weights are constant, they will cancel out in a ratio of totals (which a mean is), one can set the weights to an arbitrary constant, say one. See 3.4 for further detail on this.

The two key features of sample design - stratification and clustering - can be separated and dealt with independently. That is, whatever use is made of clustering, stratification is always dealt with in the same way, and vice versa. We first consider these separately (in 3.1 and 3.2), and then bring them together (in 3.3). In 3.4, we move from totals to means.

Some of the above assumptions may be questioned, in particular that of random sampling at the first stage, since, as claimed in 2.7, systematic sampling is far more common if not universal. However, the reasons for making these assumptions will soon become evident. In 3.5, we discuss extensions and, more importantly, the extent to which results based on these assumptions can serve as, and are needed as, approximations for results valid for alternative sampling designs.

3.1 Stratification

Suppose the population is divided into H strata. (If there is no stratification, H=1.) In the hth stratum, let t_h be an unbiased estimator of τ_h, the total income in that stratum. Then one can re-write equation (1) as

$$t=\sum_{h=1}^{H} t_h \tag{2}$$

The variance of t is given by

$$\begin{aligned} Var(\sum_{h=1}^{H} t_h) &= \sum_{h=1}^{H} Var(t_h) + 2\sum_{h=1}^{H} \sum_{h'\neq h} Cov(t_h, t_{h'}) \\ &= \sum_{h=1}^{H} Var(t_h) \end{aligned} \tag{3}$$

so that if $\hat{Var}(t_h)$ is an unbiased estimator of $Var(t_h)$, Var(t) can be estimated unbiasedly by summing over the $\hat{Var}(t_h)$.

Ignoring stratification will overestimate standard errors,[9] since the between-stratum variances, which should be ignored, will be included in the estimate. Intuitively, this is because, with stratification, one ensures that no part of the sampling frame goes unrepresented (Kish, 1965, pp.139-142). Note that stratification must be taken account of whether it is proportionate or disproportionate.

---

9. Of course, this true only in expectation. We take this qualifier to be implicit here and throughout the paper. By "ignoring stratification" we mean ignoring the fact that different observations come from different strata. We do not mean ignoring the fact that different observations are associated with different expansion factors, even if these expansion factors in fact come from disproportionate stratification (it can be, for example, wrongly assumed that they come from the use of a ppes selection of clusters).

3.2 Clustering

Given that equations (2) and (3) hold whatever use of clustering is made within each stratum, we can focus on the derivation of t_h and $\hat{Var}(t_h)$, that is, on estimators for a single stratum. Let there be N_h clusters in this stratum from which n_h cluster takes are chosen with replacement. Let M_{hc} be the true size of (number of households in) the cluster from which the cth cluster take is selected in the hth stratum and let Z_{hc} be the estimated size of the cth cluster.[10] The total estimated stratum population size is $Z_h = \Sigma_c Z_{hc}$. Randomly select one cluster with probability proportional to estimated size, i.e. with probability Z_{hc}/Z_h, and sample m_{hc} of the M_{hc} households. Define the cluster-level expansion factor:

$$w_{hc} = \frac{Z_h}{Z_{hc}} \frac{M_{hc}}{m_{hc}} \tag{4}$$

Note that w_{hc} takes into account any or both of disproportionate stratification and the use of ppes. If pps is used, $Z_{hc}=M_{hc}$ and, assuming m_{hc} to be constant within a stratum, the weights reflect only the former (and so will vary only between strata).[11]

Since the selection of clusters is random, any one cluster take can be used to form an unbiased estimator of total income, τ_h, namely t_{hc}:

$$t_{hc} = w_{hc} \sum_{i=1}^{m_{hc}} y_{hci} \tag{5}$$

Note that this is estimated mean income in the cth cluster times $Z_h*(M_{hc}/Z_{hc})$, the estimate of the size of the hth stratum based on cluster c. Hence t_{hc} is an unbiased estimator of τ_h. It is not, however, efficient. A more efficient estimator can be obtained by utilizing information from all cluster takes. From n_h cluster takes we can construct a combined unbiased estimator of τ_h as the unweighted mean of the t_{hc}:

$$t_h = \frac{1}{n_h} \sum_{c=1}^{n_h} t_{hc} \tag{6}$$

Note that if we define (for some h and all c) $w_j = w_{hc}/n_h$ then we can rewrite (6) in the form we

---

10. c is used to refer both to the cluster and to the cluster take. In reality, if more than one take can be selected from a cluster, the cth cluster take may come from cluster b. However, to avoid a proliferation of notation, this possibility is ignored in the exposition though it is covered by the results provided.

11. If ppes is used and m_{hc} is held constant then w_{hc} will vary across clusters. The alternative ppes method set out in 2.6 adjusts m_{hc} according to the ratio M_{hc}/Z_{hc} to hold w_{hc} constant within the hth stratum.

began with, namely, analogously to (1), as

$$t_h = \sum_{j \in h} w_j y_j \tag{7}$$

However, (6) is more useful for our purposes, since it leads directly to the variance of t_h. Let S_{hc} be the number of possible takes which could be drawn from the cth cluster if it were selected at the first stage. Define S_h to be the sum of S_{hc} over all N_h clusters. Then, two-stage sampling with replacement is no more than sampling with replacement n_h of these S_h elements. Hence

$$Var(t_h) = E(t_h - \tau_h)^2 = \frac{1}{S_h} \sum_{c=1}^{N_h} \sum_{s=1}^{S_{hc}} t_{hcs}^2 - \tau_h^2 \tag{8}$$

where each cluster is only counted once (from 1 to N_h). It follows then that an unbiased estimator of $Var(t_h)$ is given by

$$\hat{Var}(t_h) = \frac{1}{n_h(n_h-1)} \sum_{c=1}^{n_h} (t_{hc} - t_h)^2. \tag{9}$$

(See Appendix I.1 for a proof, which parallels the one used for simple random samples.) This is a remarkable and useful result, basic to the sampling literature, but rarely applied to poverty analysis. The result is remarkable for its simplicity in the face of all the complexities of sample design which it is capturing. It is useful for its very weak assumptions about sampling at the sub-cluster level. In particular, the formula takes no explicit account of sampling variability within each cluster. All that is assumed is that the t_{hc} are unbiased estimators of τ_h. t_{hc} may be obtained using random sampling with or without replacement or systematic sampling.[12]

Unlike the case of stratification, ignoring clustering will often lead to an *under*estimate of Var(t). If one (wrongly) assumed simple random sampling, one would assume

$$Var(t_h) = \frac{Var(t_{hci})}{m_h} \tag{10}$$

where $t_{hci} = y_{hci} w_{hc} m_{hc}$ is an estimator of τ_h based on the ith household in cluster c in stratum h and $m_h = \Sigma_{c=1}^{n_h} m_{hc}$. In fact the correct variance is given by (8). As shown in Appendix I.3, if there is two-stage sampling, and we make the simplifying assumption that sampling is with replacement at both stages, then (8) may also be written

12. (9) does assume that one has at least two cluster takes per stratum (so that n_h-1 > 0). If this condition is not satisfied, say because data is missing, one can construct "pseudo-strata" by joining neighboring strata until all are made up of at least two cluster takes.

$$Var(t_h) = \frac{Var(t_{hci})}{m_h}[1 + (m_h^* - 1)\rho_h], \tag{11}$$

where

$$\rho_h = \frac{Cov(t_{hci}, t_{hcj})}{Var(t_{hci})}, \qquad m_h^* = \frac{\sum_{c=1}^{n_h} m_{hc}^2}{m_h} \tag{12}$$

ρ_h is the intra-cluster correlation coefficient for the hth stratum. If $\rho_h > 0$ (that is, if there is homogeneity within the cluster with respect to the variable of interest), which will be the case in practice, then assuming simple random sampling will yield an estimator of $Var(t_h)$ which is biased downward. The intuition is the same as that for stratification, though here it works in the opposite direction: clustering leads to a less uniform coverage of the population, and so reduces precision.

Both (9) and (11) - the latter with caps added - provide estimators of $Var(t_h)$. The former is simpler, since it involves no computation of covariances. Moreover, if sampling at the second-stage is systematic, then any estimation of ρ_h based on a single sample will reflect the influence of the ordering on which the systematic selection was based as well as the degree of intra-cluster correlation, and so may be biased. Equation (9), by contrast, is valid for both systematic and random sampling at the second stage.

3.3 Stratification and clustering

For a sample which is both stratified and clustered, use the formulae for t_h and $\hat{Var}(t_h)$ in 3.2 in the formulae for t and $\hat{Var}(t)$ in 3.1. Bringing together our earlier remarks on the effects of clustering and stratification, for a given sample size and population variance, the variance of estimators is smaller the greater the homogeneity between clusters within strata and the greater the heterogeneity within clusters. The introduction of clustering increases within-stratum homogeneity since clusters differ by less than households do. This means that stratification becomes a more potent tool for increasing estimator precision once one has clustering (Kish, 1965, p.164). We will see this at work in Section 5.

3.4 Means

An estimated mean is a ratio of two totals, with the estimated total of the variable of interest as the numerator and estimated or actual population size as the denominator. The population size will either be the number of households (if one wants a household mean) or the number of individuals (if one is after a per capita mean). We can handle both cases using p_{hc},

an estimator of the population (however defined) in the hth stratum based on data from the cth cluster take in that stratum. Now, p_{hc} is exactly analogous to t_{hc} defined in (5): for a household mean, replace the y_{hci} in that equation by ones; for a per capita mean replace y_{hci} by household size (h_{hci}). Likewise, p_h will be defined analogously to t_h in (6) as the simple mean of the p_{hc} and an estimator of its variance will be derived identically (see (9)). Finally, p will be defined analogously to t in (2) as the sum of the p_h. Average income can then be estimated by $\bar{y}$, defined as

$$\bar{y}=\frac{t}{p} \tag{13}$$

Since this estimator is a ratio of two random variables, it will be biased but consistent. Using a Taylor's expansion, a consistent estimator of its variance is

$$\hat{Var}(\bar{y}) = \frac{1}{p^2} [\hat{Var}(t) + \bar{y}^2 \hat{Var}(p) - 2\bar{y} \hat{Cov}(t, p)]. \tag{14}$$

The covariance is the only new term. If there are strata, then the covariance for the entire sample will be the sum of the covariances within each stratum. The covariance within a single stratum is estimated by

$$\hat{Cov}(t_h, p_h) = \frac{\sum_{c=1}^{n_h} (t_{hc}-t_h)(p_{hc}-p_h)}{n_h(n_h-1)}. \tag{15}$$

These formulae may also be used to calculate means and their standard errors for sub-populations of interest. If whole clusters are contained within a sub-population, for example, urban or rural, then only those clusters are used. If clusters have elements from various sub-populations, for example, literate or illiterate households, then, for a given category, the relevant households are included and other households are given a weight of zero. In either case, but especially the second, the number of cluster takes per stratum may fall below two in which case neighboring strata may be merged (see footnote 12).

What if the sample is self-weighting, and one does not have the weights given by (4)? Then totals cannot be calculated, but means still can be. Since the (unknown) weights - the w_j of equation (1) - will not vary across clusters or strata, and since they will appear in both the denominator and numerator of (13), they will cancel out, and so can be ignored from the start. Hence, for a self-weighting sample, one can arbitrarily set w_{hc} to n_h for all c and h (and thus w_j to 1 for all observations - see p.11). More generally, as long as the weights give unbiased estimates of some multiple or fraction of the relevant population totals, one can use the formulae of this section to calculate means and their variances. This is important since often weights provided with survey data sets are normalized. As long as one is only interested in calculating means and not totals, the normalization is not important.

Note finally that having a self-weighting sample does **not** in general imply that $\hat{Var}(p)$ is zero. If one is analyzing a per capita mean, variation in household size will lead to a non-zero variance for p.[13] However, if one has a self-weighting sample selected using pps and one is estimating a household mean then there will be no variation in cluster sample size and the variance of p will be zero.[14] In this case, (14) simplifies to

$$\hat{Var}(\bar{y}) = \frac{1}{p^2}\hat{Var}(t). \tag{16}$$

3.5 Alternative assumptions

We now considering relaxing some of the initial assumptions. First, say there are more than two stages. This poses no problem as long as we assume the first-level cluster takes are selected with replacement. One takes the number of cluster takes as given by the number of these first-level cluster takes. Expansion factors may not then be constant within the clusters thus defined, since lower-level cluster takes may be chosen using ppes, but unbiased estimates of τ_h from each first-level cluster can still be obtained. Likewise if individuals or groups of households are the ultimate sampling units. We have already incorporated the case of no stratification (simply set $H=1$), but say there is no clustering. Then $Var(t_h)$ will depend on the precise sample design used within each stratum. Take the simplest case of stratified random sampling (with replacement). Then $Var(t_h)$ becomes the element variance of in the hth stratum divided by the number of households sampled from that stratum. Hence assumptions (i), (iv) and (v) are easily relaxed, while (ii) and (iii) are not restrictive.

What if the absence of self-weighting is due to non-response rather than the use of ppes (our assumption (vii))? If, as sometimes happens, non-response is corrected at the cluster level using weights, this is exactly analogous to having cluster weights due to ppes.[15] But in other cases, non-response is assumed to be a function of other-than-cluster-level variables, such as age groups. One is then left with weights which can vary within clusters. For discussion of this more

---

13. This must of course be taken account of even if one has a simple random sample. Hence the formula presented by Kakwani (1993) for individual-based poverty measures are invalid even if applied to simple random samples (they are valid only for household-based measures).

14. Another way of looking at this is that, if pps is used, the population size must be known (or at least assumed to be known) by the sample designer, though not necessarily the sample analyst.

15. This is assuming that the respondents have the same characteristics as those questioned. In many contexts this will not be the case and non-response will introduce sample selection biases in estimates. There is no general solution to this problem. Note that to correct for non-response in this way, one needs to know whether existing expansion factors assume target or realized cluster sample sizes. This may not always be clear.

difficult case see Hansen *et al.* (1953), Kish (1965), and Pudney and Sutherland (1994). The arguments of the latter authors would suggest there is little loss in accuracy from treating these as random weights and thus in the same way as expansion factors which vary within first-level clusters due to the use of multi-stage clustering (see the previous paragraph).

We now come to assumption (vi), the restrictive part of which is that the sampling of clusters is random and with replacement. In fact, we know from 2.7 that many if not all surveys use systematic sampling at the first stage. Why not assume this then? If the ordering on which the systematic sampling is based is approximately random, then systematic sampling will approximate random sampling. This means that the assumption of random sampling is valid for the first type of systematic sampling discussed in 2.7. But say that, as is much more common, sampling is systematic with non-random ordering - the case of "implicit stratification". Since the randomness enters only once (at the start) of a systematic draw, unless one has several systematic draws, one does not have enough variation in the sample to estimate variances. So, unless there is more than one sub-sample (known as "replicated sub-sampling"), variances simply cannot be estimated from systematic samples.[16] Since implicit stratification adds precision, we do know, however, that variances calculated under the assumption that the sample is random will tend to be upper bounds on the true variances and so will be conservative.[17] Moreover, we can do better than this. Implicit stratification leads to precision for the same reason as explicit stratification, and we can in fact treat samples derived using the former method as if they had been derived using the latter, using a method given by Kish (1965).

Say that all the clusters in the hth explicit stratum have been listed from 1 to N_h and n_h cluster takes have been selected systematically. Group pairs of cluster takes closest to each other into sub-strata (starting with the first two). Calculate $\hat{Var}(t_h)$ as the sum of estimated variances for each sub-stratum, where the latter are calculated using the pairs of clusters takes and (9).

---

16. There are some samples which are based on replicated sub-samples. For example, the Nicaraguan LSMS is based on four systematic samples of clusters within each stratum, each with a different random start). However, most surveys either are not designed in this way or, if they are, do not record which cluster is part of which replicate. Moreover, the small number of sub-samples involved will lead to unreliable variance estimates (Kish and Frankel, 1970, p.1088). Replicates can also be created by randomly sub-sampling from the sample once it has been collected. However, this must be done based on the design of the original sample. Taking random sub-samples which ignore any clustering or stratification present and basing variances on the variation in the mean estimates from these different sub-samples will lead to biased variance estimates. The method of "balanced repeated replications" (Kish and Frankel, 1970) is one type of sub-sampling which does take into account sample design. But in this case one might just as well, wherever possible, use the same assumptions used to replicate (post-survey) to calculate the variances analytically. It is only when this is not possible (when one has statistics more complex than unconditional means) that post-survey sub-sampling is appropriate.

17. The qualifier 'tend' is required since one can think of cases in which implicit stratification will reduce precision. This will be the case if the elements display periodicity: so that every sth element will be similar of if the ordering displays very strong trends (Kish, 1965, pp.120-121). Neither case will be relevant in the household survey context.

$\hat{Var}(t_h)$, thus calculated, would be an unbiased estimator of the hth stratum variance if that hth stratum had been divided into $n_h/2$ sub-strata - each containing, among others, the relevant pair of sampled clusters - and two cluster takes had been selected at random and with replacement from each of these sub-strata. Moreover, $\hat{Var}(t_h)$ will also be only slightly biased upwards as an estimator of $Var(t_h)$ given actual sample design. There would be no overestimation at all if the ordering of clusters within each sub-stratum was random. Since clusters which are close to each other in a list used for implicit stratification are unlikely to differ much, this assumption is unlikely to be seriously violated and so the overestimation will be small (see Section 5 for an example). Certainly, when there is implicit stratification, the assumption that the ordering is random only within each (artificial or created) sub-stratum is better than the assumption that it is random within each (explicit) stratum. In any case, working with $n_h/2$ sub-strata makes the number of clusters within sub-strata as small as possible subject to the requirement of at least two clusters be chosen per sub-stratum and so makes the overestimation as small as possible.[18] Note that this method requires that one knows the original ordering of clusters, from which the sample was systematically drawn. Unfortunately, this information is often not available to analysts. In this case, one has no choice but to ignore implicit stratification altogether.

So, the assumption of random sampling will lead to little loss of accuracy if applied in the manner suggested, and in any case we have no better alternative available. However, in assumption (vi), we assume not only random sampling at the first stage, but random sampling *with replacement*. Why not assume random sampling without replacement? Formulae are certainly available for this case. For simplicity, assume all clusters are of equal size and that equal-probability random sampling (without replacement) occurs at both stages. Then, as shown in Appendix I.2, an unbiased estimator of the variance of t_h is given by:

$$\hat{Var}(t_h) = \left(\frac{N_h - n_h}{N_h}\right)\frac{\sum_{c=1}^{n_h}(t_{hc}-t_h)^2}{n_h(n_h-1)} + \left(\frac{\bar{M}_h - \bar{m}_h}{\bar{M}_h}\right)\frac{\sum_{c=1}^{n_h}\sum_{i=1}^{\bar{m}_h}(t_{hci}-t_{hc})^2}{N_h n_h\ \bar{m}_h(\bar{m}_h-1)}, \tag{17}$$

where N_h (n_h) is the total (sample) number of clusters in stratum h and $\bar{M}_h$ ($\bar{m}_h$) the common population (sample) cluster size. (9) will generally exceed (17) (Kish, 1965, p.160) so why not use the latter? There are four reasons. First, (17) is more complex. Once cluster sizes are allowed to vary and pps or ppes used, the unequal probabilities of selection at the first stage must be taken account of (even if the sample if self-weighting) by a weighting-scheme involving single **and joint** inclusion probabilities for each cluster (see Shah *et al.*, 1993). Second, it won't make much difference. As N_h increases, the second term goes to zero. Third, (17) makes

18. Though Kish (1965, pp. 202 and pp. 285-286) provides this method, his exposition (in terms of collapsing pairs of strata) is slightly different. See Kish also for the case in which n_h is not even. Basically, the idea is to end up with as fine an explicit stratification as possible. Kish (p. 286) also provides an alternative method based on a double pairing (the first cluster with the second, the second with the third, and so on). This will give a more precise estimate, though whether it is worth the extra trouble is debatable.

greater informational demands. To make the finite population corrections one needs to know the total number of clusters, N_h, and the population of the clusters, $\bar{M}_h$, information that the statistical office which collected the data may know, but which the data analyst often will not. By contrast, the information required for the formulae we have presented will typically be available to the analyst. Finally, whatever happens at the first stage, sampling at the second will invariably be systematic, which means the second term of equation (17) will no longer be valid. The advantage of (9), by contrast, is that, to re-iterate, it is valid whatever the method of sampling used at the second stage (as long as it gives rise to unbiased estimators). For all these reasons, the use of formulae valid for sampling with replacement is widely recommended in the sampling literature (see Kish, 1965, p.160; Som, 1973) even if it does lead to a slight overestimation of variances.

4. Poverty and other welfare measures

We deal with the class of additive poverty measures for which the analogy with a simple mean is direct. Let $\pi(k_{hci})$ be a measure of household poverty for the ith household in the cth cluster in the hth stratum. There are two aspects to $\pi(k_{hci})$. First, we define $k_{hci}=f(y_{hci},X_{hci})$. y_{hci} has already been defined as total household income/consumption and X_{hci} is a vector of variables such as prices and household size which allow one to normalize y_{hci} such that households with equal k's are assumed to be equally well off. In the simplest case, k_{hci} will be per capita nominal income (y_{hci}/h_{hci}).

Second, a functional form needs to be chosen for π. The indicator function, $I(k_{hci} \leq k)$, where k is the poverty line, plays a crucial role here. If the expression inside the brackets is false then I(.) is zero. If it is true then I(.) equals one if one is interested in poverty among households; it equals h_{hci}, the number of people in the household, if one interested in the number of individuals poor.

We can now define some specific measures. For the headcount ratio, $\pi(k_{hci})=I(k_{hci} \leq k)$. For the poverty gap, $\pi(k_{hci})=I(k_{hci} \leq k)*(1-k_{hci}/k)$. More generally, the well-known and much-used FGT family is given by $\pi_\alpha(k_{hci})=I(k_{hci} \leq k)*(1-k_{hci}/k)^\alpha$ for $\alpha \geq 0$ (see Ravallion, 1994, for further details).

In order to calculate aggregate poverty measures and standard errors, simply replace y_{hci} in equation (5) with $\pi(k_{hci})$. $\bar{y}$, in equation (13), then becomes an estimator of aggregate poverty, with a variance estimator given by (14).

Use of the simple formulae given here also enables statistical poverty stochastic dominance analysis to be undertaken. First-class stochastic dominance is based on the headcount ratio and second-class is based on the poverty gap, both over some range of poverty lines. Combining the estimators with the testing method presented in Howes (1993) enables one to test for statistically significant stochastic dominance rather than just work with sample outcomes.

The attention paid by this paper to poverty measures is simply to provide a focus. One can, of course, think of a more general class of welfare measures, of which the negatives of poverty measures are one sub-class. The mean itself is one type of welfare measure, and there are many others within the class of separable functions with more desirable (egalitarian) properties. With some extensions to cater for mean-normalization, one can also derive variances for estimators of inequality indices.

5. Some examples

We illustrate the quantitative importance of these considerations with analyses of household surveys from two developing countries, Pakistan and Ghana. What we measure for both countries are *sample design effects*. These give the ratio of the standard error one obtains given a particular set of assumptions about sample design to the standard error one obtains assuming that the sample is a simple random one. This ratio has two interpretations. First, it gives the magnitude of error made when ignoring sample design. Second, it represents the efficiency gain or loss from moving away from simple random sampling. Since our interest is in analyzing rather than designing surveys, our interest is primarily in the former interpretation, though we also briefly address the second. For the second interpretation, one needs to be able to estimate what the standard error of the estimator of interest would be for the same population under some alternative sample design. The fact that one does not actually have that alternative sample design does in some cases introduce bias: for example, the standard estimator for the element variance is biased downward in a clustered sample. However, one can show that such biases are negligible for the cases considered here (Kish, 1965, Deaton, 1994).

We deliberately focus on two surveys which, although both LSMS surveys, are quite different in nature. Both are of course clustered and use systematic sampling and both are also two-stage with the household being the second-stage sampling unit. But in other important ways the two surveys differ.

The Pakistani sample is designed in a traditional, highly-stratified way (described in Howes and Zaidi, 1994). Its 4745 observations are collected from 300 clusters in 104 strata (see Table 2).[19] The Pakistani sample is also non-self-weighting. Its cluster-level weights reflect the usage of ppes as well as disproportionate stratification, with, in particular, urban households over-represented.

Table 2 Sample design for Pakistan and Ghana LSMS surveys

	Strata	Number of clusters	Cluster weights	Sample size	Average cluster sample size
Pakistan Integrated Household Survey, 1991	104	300	Yes	4745	15.82
Ghana Living Standards Survey, 1987-88	None	170	No	3181	18.71

The Ghanaian sample is made up of 3181 households from 170 clusters (see Scott and Amenuvegbe, 1989, for a detailed description). As mentioned in the notes to Table 1, ppes is

19. For simplicity, we delete one stratum since it only has one cluster. Alternatively, we could have followed the procedure recommended in 3.3.

used and cluster-take sizes are adjusted by the ratio of actual to estimated cluster size. However, the survey contains an innovation: this ratio was rounded off to the nearest integer to restrict all cluster-take sizes to a multiple of 16 (for reasons of administrative convenience). Scott and Amenuvegbe (1989 and 1991) show that this restriction introduces only a negligible bias so the standard ppes formula can be used. 141 cluster takes have a size of approximately 16 (a very small number of households is lost to data cleaning, 19 in all[20]), 28 of 32, and 1 of 48 (an additional 30 clusters originally selected were dropped since their ratio of actual to estimated size was less than half).

The Ghanian sample is not explicitly stratified at all. However, implicit stratification was used: clusters were selected on the basis of a geographical ordering. To deal with this, we follow the method recommended in 3.5, and approximate the effects of this implicit stratification by a fine explicit stratification. It is assumed, on the basis of informed advice, that the current ordering of clusters reflects their original ordering in the sampling frame. On this basis, the 170 clusters were divided into 85 strata, starting with the first two clusters in the first stratum.

As noted in 3.5, this procedure may leave a slight overestimation of the 'actual' standard errors, since the assumed stratification may not capture all of the precision induced by systematic sampling. Figure 1, however, shows that the procedure is likely to be very effective. It gives standard errors assuming that the implicit stratification can be captured by explicit stratification of differing degrees of fineness, beginning with no stratification (assuming a random ordering), then with 2, 3, 5, 10, 21, 42 and 85 strata, the latter being the finest possible degree of stratification consistent with the data.[21] It is clear from the figure that most of the gains from finer stratification are exhausted by the introduction of twenty strata. In particular, in this case from 42 to 85 strata actually slightly increases the estimated standard error (recall that the fall is in expectation). One can deduce from this that the approximation made of random ordering of clusters within each of the constructed 85 strata causes very little overestimation indeed.

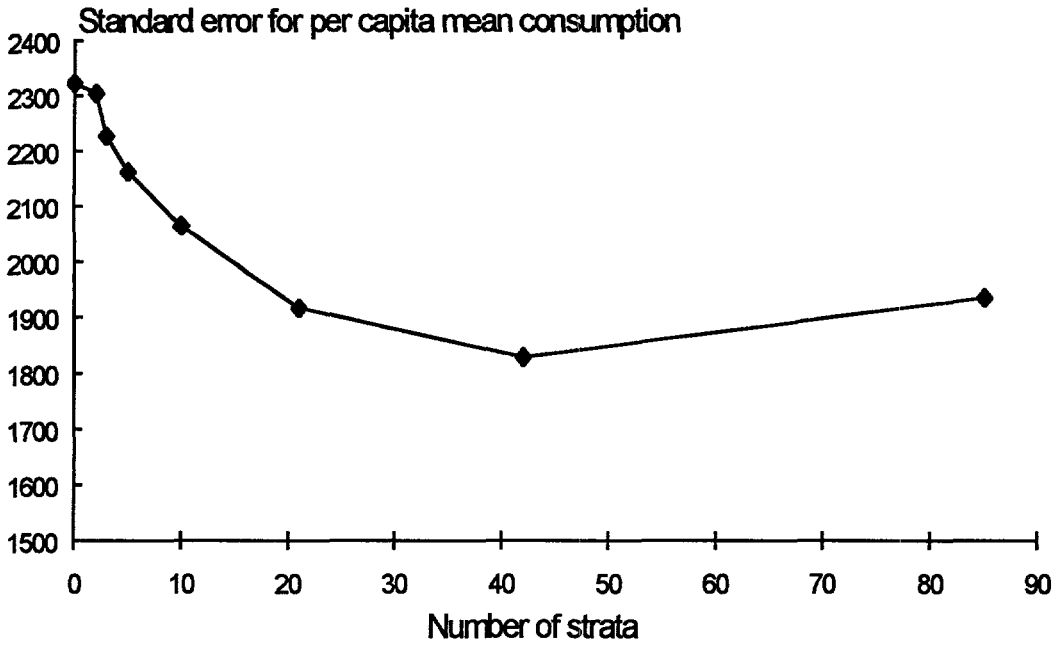
In Table 3 we present standard errors for both means (of household size and aggregate expenditure) and the three most used poverty measures: the headcount, poverty gap and FGT2 ($\alpha=2$) index. Both the poverty indices and the expenditure means are per capita. Since the exercise is purely illustrative we do not worry about equivalence scales or regional deflators, and arbitrary poverty lines are set at 3780 rupees per capita per month for Pakistan and of 37,900 cedis per capita per annum for Ghana. Both lines are chosen to put approximately 1/3 of the sample in poverty.

---

20. Since attrition due to cleaning is so low, no adjustment is made to correct for it. Likewise, no adjustment is made to the Pakistani data for non-response since (i) we are unsure whether the weights assume actual or target sample sizes (see footnote 15) and (ii) we tried both assumptions and it made no difference due to the very high response rates.

21. These degrees of stratification were obtained as described in the text. Where clusters could not be divided evenly among the strata, any residual clusters were added into the last stratum.

Figure 1 The impact on standard errors of approximating implicit by differing degrees of explicit stratification: an example from Ghana



For each measure, four standard errors are given. Each is an estimate of the standard error which would be correct for the measure of interest for some sample design. (The means are correct under all four sets of assumptions, since in all four cases, the same weights are assumed - though their rationale would have to change.) The first ("stratification and clustering") is correct for the set of assumptions outlined at the start of Section 3 and the degree of stratification and clustering found in each survey (with the necessary approximations made where necessary to fit the surveys to this model). The second ("clustering without stratification") would be correct for the same set of assumptions if the surveys were conducted without stratification. The third ("stratification without clustering") would be correct for the same set of assumptions except that instead of clustering one had random sampling of households (with replacement) within each stratum. The fourth ("simple random sample") would be correct if the surveys were simple random samples (with replacement) in all respects except that they were not self-weighting.

Table 3 Sample design effects for mean expenditure, household size and various poverty measures for two household surveys

		Mean household size		Mean expenditure		Headcount		Poverty-gap		FGT2 (α=2) index	
		Pak.	Ghana	Pak.	Ghana	Pak.	Ghana	Pak.	Ghana	Pak.	Ghana
Estimate		7.17	4.52	5934	71276	.336	.333	.092	.098	.0364	.0403
Standard errors	A. Clustering and stratification	.072	.081	125	1935	.012	.016	.0046	.0072	.0024	.0039
	B. Clustering w/o strat'n	.087	.098	158	2323	.017	.020	.0063	.0083	.0032	.0043
	C. Strat'n w/o clustering	.059	.053	86	918	.009	.010	.0033	.0042	.0018	.0024
	D. Simple random sample	.061	.055	91	1002	.009	.011	.0035	.0044	.0019	.0025
Ratio of standard errors	A/D	1.18	1.47	1.37	1.93	1.33	1.45	1.31	1.64	1.26	1.56
	B/D	1.43	1.78	1.74	2.32	1.89	1.82	1.80	1.89	1.68	1.72
	C/D	0.97	0.96	0.95	0.92	1.00	0.91	0.94	0.95	0.95	0.96
	A/B	0.83	0.83	0.79	0.83	0.71	0.80	0.73	0.87	0.75	0.91

Notes: See the text for explanation of the different standard errors.

We begin with Pakistan. For mean household size, the design effect is not that large. Standard errors increase by only 18% over a simple random sample (A/D). This is what one would expect as there will typically be a great deal of variation within any one cluster with respect to household size. But for mean expenditure and the poverty measures, the design effect is much larger. For the mean, it is 37%; for the headcount 33%; for the poverty gap 31%; and for the FGT2 index 26%. On average, standard errors increase by around a third. Ignoring the sample design of the Pakistan survey leads to calculated standard errors which are substantially smaller than the correctly estimated standard errors.

It is clear from the table that stratification is very important. With clustered but unstratified data, standard errors would be between 68% and 89% (depending on the welfare/poverty measure) greater than under srs (B/D). But note that stratification is only important when the sample is clustered. Take mean expenditure as an example. Without clustering, stratification reduces the standard error on mean expenditure by 5% from 91 to 86 rupees (C/D). With clustering, stratification reduces standard errors by four times as much, that is, by 21% from 158 to 125 rupees (A/B). Taking into account clustering but not stratification can be as misleading as taking neither into account. For the mean, the former strategy's estimate of 158 overestimates the correctly calculated standard error of 125 by as much as the latter

strategy's estimate of 91 underestimates it.[22]

The sample design effects from Ghana are qualitatively similar and quantitatively larger. Sample design increases the standard error of household size by 47% (A/D). For the other estimators, increases due to sample design range from a low of 45% for the headcount to 55-65% for the poverty gap and FGT2 index to 93% for mean expenditure. Again the importance of taking stratification (here implicit) into account is evident. With clustering, stratification reduces standard errors by 9-20% (A/B); without it, stratification reduces standard errors by only 4-9% (C/D).

Of course, whatever the standard error estimated, the t-ratio associated with its poverty measure will be very large. This should not surprise, since all it means is the rejection of the null hypothesis of zero poverty. As indicated by the paper's title, however, of greater interest are poverty comparisons. Here the inaccuracies in standard error computation are much more important. Take the case of Pakistan. Say we had another survey of identical size and design a few years later. Assuming that the new survey would give approximately the same standard errors, how much of a change in poverty would the new survey have to register for that change to be judged significant (at a level of .05), that is, not due to sampling variation? For the headcount, for a simple random sample, the, say, decrease would have to be from 33.6% to 31.1%, a decline of 6.6%. With clusters and strata, a 10.1% decline would be required from 33.6% to 30.2%. One can imagine many cases in which two survey points would register a change in poverty of 7-10%: at least in Pakistan, these are the cases in which taking account of sampling design is imperative.[23]

---

22. Of course, one might prefer the former on the grounds that it is better to err on the side of caution.

23. Rodgers and Rodgers (1992) calculate sample design effects for permanent (multi-year) poverty measures using U.S. data and the balanced repeated replication method mentioned in footnote 16 above. They find larger sample design effects of 50 to 200% (see their footnote 22). Comparing their results with those of this paper suggests that data source used by Rodgers and Rodgers is not highly stratified (so that the clustering effect is hardly offset at all), though the actual degree of stratification is not given in the paper.

6. Concluding comments

Household surveys are far from being simple random samples. Typically, they are clustered and highly stratified, to name only the two most important departures from the srs model. Although these departures have been ignored by recent papers on statistical poverty analysis, they are of considerable importance. The examples of the previous section show that not taking into account sample design can increase standard errors by around one-half.

The main recommendation of this paper is therefore that standard errors for poverty and welfare measures should be calculated taking proper account, where possible, of sample design. Both stratification (explicit or implicit) and clustering should be factored in. Taking account only of one, and not the other, can lead to inaccuracies as great as taking account of neither. Often the complexity of sample design will necessitate approximations, but these are much less egregious than just assuming simple random sampling.[24]

Examining sample design from a statistical perspective also leads to several ancillary conclusions about data availability, analysis and sample design itself. On data availability, the obvious point is that the information required to take account of sample design should be made available with the sample information itself. Thus each household should have a cluster identifier and, if appropriate, a stratum identifier. Not only should raising factors be made available, but their derivation should be documented, as should the sample design itself. All surveys should be describable in terms of the list of features presented in Section 2. In particular, if implicit stratification is used to pick clusters, the relationship of the cluster numbering to the original ordering of clusters should be given. For many surveys, precisely because attention has not been paid to sample design by end-users, these basic requirements are not fulfilled (see also Harpman and Scott, 1987, p.375).

On data analysis, we have been concerned in this paper with the calculation of standard errors for estimators of unconditional means (whether of original or transformed variables). What about standard errors for conditional means, such as are obtained by regression analysis? It has recently become much more widespread to take into account clustering when calculating standard errors for regression coefficients (using methods given in Deaton, 1994). Whether or not stratification should be similarly accounted for is more contentious. But assume that one either has a highly explicitly stratified or an implicitly stratified sample. In either case, one could regard this as a survey with n cluster takes and, approximately, n/2 strata. In such a case, it is unlikely that one would want to handle stratification by introducing dummy variables, as one might do with a smaller number of strata (a common strategy with urban/rural or provincial strata). However, if there is heterogeneity across strata, one would like to take account of the greater precision induced by this stratification. If one does not, and only factors in the clustering, one runs the risk of exaggerating the design effect, at least if the examples using

24. Standard data-processing software cannot always be relied upon to provide the correct estimators, but specialist packages are available (for example, Sudaan, PC Carp and Clusters).

unconditional means presented in this paper are any guide. There are methods available for incorporating both clustering and stratification into regression analysis (see Deaton, 1994; Shah *et al.,* 1993; and Kish and Frankel, 1970). The importance of implementing these methods is clearly a topic worthy of further research.

Finally, we turn to sample design, in particular to the question of stratification. Given the large precision gains induced by stratification when clustering is also being used, the current practice of using some degree of stratification at the first stage would certainly seem warranted, though one might argue over the optimal degree. The question remains, however, of whether explicit or implicit stratification should be availed of.[25] Many survey designers favor the latter on the grounds of simplicity (see Scott and Amenuvegbe, 1989, p.17, for example). But from an analytical point of view, one must favor the explicit variety, since, for the calculation of variances, one is forced to assume that this is what is used. Opposing this, implicit stratification will tend to produce slightly more precise estimates than even the finest possible degree of explicit stratification (two clusters per stratum), though here we have stressed that the qualifier 'slightly' must be emphasized. We are unclear which type of sampling is easier to implement, but suspect that there is little difference. We also suspect that the information required to take account of explicit stratification information is easier to record than that required for implicit stratification. The former requires a single, stratum-representing variable. The latter requires that significance be given to the ordering of the actual cluster numbers chosen, information which could easily be lost, especially in a complex list.[26] Taking all of these factors into account, we would argue that the precision gains of explicit stratification should first be exhausted before one turns to implicit stratification. That is, one should not use systematic sampling at the first stage until one has divided the sampling frame into N/2 strata, where N is the desired number of clusters. Only once one has done this should one consider use of systematic selection (to select the two clusters from each stratum) in order to achieve even greater precision (though by an unquantifiable amount) through implicit stratification.[27] If one uses fewer than N/2 strata, it is hard to see the rationale for any use of systematic sampling, since in this case one could increase precision in an analytically tractable way with more extensive use of explicit stratification.

---

25. This is of course in relation to the first stage. Implicit stratification at the second stage creates no problems at all.

26. Certainly we have never seen such information documented, though this may be because there has never been a demand for it.

27. Alternatively, one can, as Kish (1965, pp.284-285) recommends, divide each of the strata into half-strata and randomly pick one cluster from each. This avoids the use of systematic sampling altogether.

Appendix I: Proofs

We begin with a summary of the relevant notation required for these proofs. $t_{hci}=y_{hci}*Z_h*M_{hc}/Z_{hc}$ is an unbiased estimator of the population parameter τ_h based on observation i from cluster c in stratum h. t_{hc}, defined in (5), is the simple mean of t_{hci} (summed over i and divided by m_{hc}, the sample size drawn from cluster c). And t_h, defined in (6), is the simple mean of the t_{hc} (summed over c and divided by n_h, the number of clusters drawn from the hth stratum).

1. An unbiased variance estimator for two-stage cluster sampling with replacement

We prove that (9) is an unbiased estimator of $Var(t_h)$. In fact, the proof is trivial. We give it only to make clear the importance of the with-replacement assumption. Note that in the proof, as in the text, an important distinction is made between the summation $c=1$ to n_h and that $c=1$ to N_h. The latter is simply a list of all clusters in the stratum. In the former one is selecting n_h cluster samples, with the possibility that one is sampling from the same cluster more than once. Each of the N_h clusters contains S_{hc} possible sampling combinations (depending on the second-stage sampling size and rules). The sum of the S_{hc} over all c is S_h. Two-stage sampling with replacement is no more than the random selection with replacement of n_h of these S_h possible sampling combinations, each with population parameter t_{hc}.

Proof:

$$n_h E(\hat{Var}(t_h)) = \frac{1}{n_h-1}E(\sum_{c=1}^{n_h} t_{hc}^2)-\frac{n_h}{n_h-1}E(t_h^2) \tag{A.1}$$

$$E(\sum_{c=1}^{n_h} t_{hc}^2)=\sum_{c=1}^{n_h} E(t_{hc}^2)=n_h\frac{\sum_{c=1}^{n_h}\sum_{s=1}^{S_{hc}} t_{hcs}^2}{S_h} \tag{A.2}$$

$$E(t_h^2)=Var(t_h)+\tau_h^2 \tag{A.3}$$

$$Var(t_h) = Var(\frac{1}{n_h}\sum_{c=1}^{n_h} t_{hc}) = \frac{1}{n_h^2}\sum_{c=1}^{n_h}\sum_{c'=1}^{n_h} Cov(t_{hc},t_{hc'}) \tag{A.4}$$

Since sampling is with replacement $Cov(t_{hc},t_{hc'})=0$ for $c\neq c'$. So (A.4) simplifies to

$$Var(t_h) = \frac{1}{n_h^2}\sum_{i=1}^{n_h} Var(t_{hc}) = \frac{Var(t_{hc})}{n_h} \tag{A.5}$$

Now substitute (A.5) into (A.3). Substitute the result of this and (A.2) into (A.1) to give

$$\begin{aligned} n_h E(\hat{Var}(t_h)) &= \frac{n_h}{n_h-1}[\frac{1}{S_h}\sum_{c=1}^{N_h}\sum_{s=1}^{S_{hc}} t_{hcs}^2 - \frac{Var(t_{hc})}{n_h} - \tau_h^2] \\ &= \frac{n_h}{n_h-1}(\frac{1}{S_h}\sum_{c=1}^{N_h}\sum_{s=1}^{S_{hc}} (t_{hcs}-\tau_h)^2 - \frac{Var(t_{hc})}{n_h}) \\ &= \frac{n_h}{n_h-1}(Var(t_{hc}) - \frac{Var(t_{hc})}{n_h}) \\ &= Var(t_{hc}) \end{aligned} \tag{A.6}$$

Combining (A.6) with (A.5) completes the proof.

2. An unbiased variance estimator for two-stage cluster sampling without replacement

With single-stage sampling, without-replacement estimators are obtained simply by multiplying the with-replacement variance estimator by a degrees-of-freedom correction. But, with multi-stage sampling, matters are more complex. Once any one cluster is picked, one is removing from the draw not one, but S_{hc} elements, all from the one cluster, and so all correlated.

We assume that sampling is without replacement at both stages, and that clusters are all of equal and known size, $\bar{M}_h$, and that $\bar{m}_h$ is sampled from each cluster. In this case, for all c, $w_{hc}=N_h\bar{M}_h/\bar{m}_h$, so t_{hc} is simply the stratum population times mean income in cluster c. Relaxation of these assumptions is discussed at the end of the sub-section.

To prove that (17) is a unbiased estimator of $Var(t_h)$ in this case, it is necessary first to define the latter. So we prove first that

$$Var(t_h) = \frac{1}{N_h n_h}[\frac{N_h-n_h}{N_h-1}\sum_{c=1}^{N_h}(\tau_{hc}-\tau_h)^2 + \sum_{c=1}^{N_h} V(t_{hc})] \tag{A.7}$$

where $V(t_{hc})$ is the within-cluster variance of t_{hc} (see (A.16) for the case of sampling without replacement), and is to be distinguished from $Var(t_{hc})$.

Proof:

$$Var(t_h) = V_1\ [E_2(t_h)] + E_1\ [V_2(t_h)] \tag{A.8}$$

where the subscript 2 refers to the indicated operation with clusters fixed and the subscript 1 refers to the operation over all samples of clusters.

$$\begin{aligned} V_1\ [E_2(t_h)] &= V_1\ [\ E_2[\frac{\sum_{c=1}^{n_h} t_{hc}}{n_h}]\ ] \\ &= V_1\ [\frac{\sum_{c=1}^{n_h} E_2(t_{hc})}{n_h}] \\ &= V_1\ [\frac{\sum_{c=1}^{n_h} \tau_{hc}}{n_h}] \\ &= \frac{(N_h-n_h)}{N_h-1}\frac{1}{n_h}\ V_1\ (\tau_{hc}) \\ &= \frac{N_h-n_h}{N_h-1}\frac{1}{n_h}\ \frac{\sum_{c=1}^{N_h}(\tau_{hc}-\tau_h)^2}{N_h}. \end{aligned} \tag{A.9}$$

Note here the without-replacement correction $(N_h-n_h)/(N_h-1)$.

$$\begin{aligned} E_1\ [V_2(t_h)] &= E_1\ [V_2\ [\frac{\sum_{c=1}^{n_h} t_{hc}}{n_h}]] \\ &= E_1\ [\frac{\sum_{c=1}^{n_h} V_2(t_{hc})}{n_h^2}] \\ &= \frac{1}{n_h}\ \frac{\sum_{c=1}^{N_h} V(t_{hc})}{N_h}. \end{aligned} \tag{A.10}$$

Substituting (A.9) and (A.10) into (A.8) gives (A.7), which completes the proof.

We now prove that (17) is an unbiased estimator of $Var(t_h)$ as defined by (A.7).

Proof:
(17) is made up of two terms, the first of which is

$$\frac{N_h-n_h}{N_h}\frac{\sum_{c=1}^{n_h}(t_{hc}-t_h)^2}{n_h(n_h-1)} \tag{A.11}$$

The expected value of (A.11) can be derived as follows

$$\begin{aligned}
&E_1\,\{E_2\,[\frac{1}{n_h-1}\sum_{c=1}^{n}(t_{hc}-t_h)^2\,]\}\\
&= E_1\,[\frac{1}{n_h-1}\,E_2\,[\,\sum_{c=1}^{n_h}t_{hc}^2 - \frac{1}{n_h}(\sum_{c=1}^{n_h}t_{hc})^2\,]\\
&= E_1\,[\frac{1}{n_h-1}\,E_2\,[\,\sum_{c=1}^{n_h}t_{hc}^2 - \frac{1}{n_h}(\sum_{c=1}^{n_h}t_{hc}^2 + 2\sum_{c=1}^{n_h}\sum_{c<c'}^{n_h}t_{hc}t_{hc'})]]\\
&= E_1\,[\frac{1}{n_h-1}\,E_2\,[\,(1-\frac{1}{n_h})\sum_{c=1}^{n_h}t_{hc}^2 - \frac{2}{n_h}\sum_{c=1}^{n_h}\sum_{c<c'}^{n_h}t_{hc}t_{hc'}]]\\
&= E_1\,[\frac{1}{n_h-1}\,[\,(1-\frac{1}{n_h})\sum_{c=1}^{n_h}(V_2(t_{hc})+\tau_{hc}^2) - \frac{2}{n_h}\sum_{c=1}^{n_h}\sum_{c<c'}^{n_h}\tau_{hc}\tau_{hc'}]]\\
&= E_1\,[\frac{1}{n_h-1}\,[\,(1-\frac{1}{n_h})\,(\sum_{c=1}^{n_h}V_2(t_{hc})) + \sum_{c=1}^{n_h}\tau_{hc}^2 - \frac{1}{n_h}\,(\sum_{c=1}^{n_h}\tau_{hc})^2\,]]\\
&= E_1\,[\frac{1}{n_h}\sum_{c=1}^{n_h}V_2(t_{hc}) + \frac{1}{n_h-1}\,[\sum_{c=1}^{n_h}\tau_{hc}^2 - \frac{1}{n_h}\,(\sum_{c=1}^{n_h}\tau_{hc})^2]]\\
&= \frac{1}{N_h}\sum_{c=1}^{N_h}V(t_{hc}) + \frac{1}{N_h-1}\sum_{c=1}^{N_h}(\tau_{hc}-\tau_h)^2
\end{aligned} \tag{A.12}$$

From (A.12) and (A.9), the expectation of (A.11) is

$$A \equiv \frac{1}{N_h n_h}\frac{N_h-n_h}{N_h-1}\sum_{c=1}^{N_h}(t_{hc}-\tau_h)^2 + \frac{N_h-n_h}{N_h}\frac{1}{N_h n_h}\sum_{c=1}^{N_h}V(t_{hc}) \tag{A.13}$$

From (A.7) and (A.13)

$$Var(t_h)=A + \frac{1}{N_h^2}\sum_{c=1}^{N_h} V(t_{hc}) \tag{A.14}$$

Since we can estimate A by (A.11), what is needed is an estimator for the second term of the RHS of (A.14). Consider the estimator

$$\frac{1}{N_h n_h}\sum_{c=1}^{n_h} \frac{\bar{M}_h-\bar{m}_h}{\bar{M}_h} \frac{\sum_{i=1}^{\bar{m}_h}(t_{hci}-t_{hc})^2}{\bar{m}_h(\bar{m}_h-1)} \tag{A.15}$$

That (A.15) estimates the second term of (A.14) can be seen as follows

$$\begin{aligned}
&E[\frac{1}{N_h n_h}\sum_{c=1}^{n_h} \frac{\bar{M}_h-\bar{m}_h}{\bar{M}_h}\sum_{i=1}^{\bar{m}_h} \frac{(t_{hci}-t_{hc})^2}{\bar{m}_h(\bar{m}_h-1)}] \\
&= \frac{1}{N_h}E_1[\frac{1}{n_h}\sum_{c=1}^{n_h} \frac{\bar{M}_h-\bar{m}_h}{\bar{M}_h}\sum_{i=1}^{\bar{m}_h} \frac{E_2[(t_{hci}-t_{hc})^2]}{\bar{m}_h(\bar{m}_h-1)}] \\
&= \frac{1}{N_h}E_1[\frac{1}{n_h}\sum_{c=1}^{n_h} V(t_{hc})] \\
&= \frac{1}{N_h^2}\sum_{c=1}^{N_h} V(t_{hc})
\end{aligned} \tag{A.16}$$

Adding the two terms given by (A.11) and (A.15) gives the estimator (17) which completes the proof.

Several points are in order. First, (A.8) holds whether sampling is with or without replacement. Say sampling were with replacement at the first stage. Then (A.7) would hold except without the $(N_h-n_h)/(N_h-1)$ correction factor (compare the formula in Som, 1973, p.243). The next sub-section gives yet another expression for $Var(t_h)$ under sampling with replacement.

Second, if sampling was with replacement at the first stage, then (A.12) would still hold except that, in the second term of the last line, the $1/(N_h-1)$ would be replaced by $1/N_h$. Combining this with the remarks of the previous paragraph shows that (9) is an unbiased estimator of (A.8) where sampling is with replacement at the first stage. So this sub-section provides an alternative proof to that given in the sub-section above for the main result of the paper.

Note what happens as N_h increases. (A.15) tends to zero (since it is a sum over n_h, divided by N_h). Likewise the term $(N_h-n_h)/(N_h)$ in (A.11) would tend to unity. Hence the without replacement estimator of $Var(t_h)$ - (17) - tends to the with replacement estimator - (9) - as N_h increases. This is exactly what we should expect. Under sampling with replacement, as N_h increases the chances of picking the same cluster more than once tends to zero, which is its value under sampling without replacement.

One could easily drop the assumption of a constant number of elements, m_h, being sampled from all clusters. Unequal sample sizes across clusters are only sensible if the populations of the clusters themselves differ. This is problematic because, if cluster sizes differ, invariably pps or ppes is used, which introduces unequal probabilities of selection at the first stage. In this case, as mentioned in the text, the formulae get much more complex.

3. Expressing the variance of t_h in terms of the intra-cluster correlation coefficient.

We assume that sampling is random and with replacement at both stages. We show that, under these assumptions, $Var(t_{hc})$ can be written in the form given in (11), that is, in terms of the intra-cluster correlation coefficient.

The proof is based on the t_{hci}. Since sampling is with replacement the cth cluster can be picked more than once. If it is, we assume (just for the purposes of this sub-section) that all samples from the one cluster are referred to by a single c.[28] For single-stage sampling with replacement the covariance between t_{hci} and t_{hcj} would be zero for $i \neq j$. However, with cluster sampling, this covariance is only zero for t_{hci} and $t_{hc'j}$ where $c \neq c'$. Two observations from the same cluster will have non-zero covariance (typically positive), defined to equal to $\rho_h Var(t_{hci})$, where ρ_h is the intra-cluster correlation coefficient.

To summarize then, we have

$$\begin{aligned} &Var(t_{hci}) = E[(t_{hci}-\tau_h)^2] \\ &Cov(t_{hci},t_{hcj}) = E[(t_{hci}-\tau_h)(t_{hcj}-\tau_h)] = \rho_h Var(t_{hci}) \quad for\ i \neq j \\ &Cov(t_{hci},t_{hc'j})=0 \quad for\ c \neq c' \end{aligned} \tag{A.17}$$

t_h can be written as follows:

28. This makes both n_h and m_{hc} random variables. But since the proof is conducted solely in relation to the second sampling-stage units, this does not cause any extra difficulty.

$$t_h = \frac{\sum_{c=1}^{n_h}\sum_{i=1}^{m_{hc}} t_{hci}}{m_h} \tag{A.18}$$

where m_h is the sample size (in terms of ultimate sampling units) in the hth stratum (see (10) for a definition). The variance of t_h can also therefore be re-rewritten:

$$\begin{aligned} Var(t_h) &= (m_h)^{-2}E\left\{\sum_{c=1}^{n_h}\sum_{i=1}^{m_{hc}} t_{hci} - E[\sum_{c=1}^{n_h}\sum_{i=1}^{m_{hc}} t_{hci}]\right\}^2 \\ &=(m_h)^{-2}E\{\sum_{c=1}^{n_h}\sum_{i=1}^{m_{hc}} t_{hci}-m_h\ \tau_h\}^2 \end{aligned} \tag{A.19}$$

Using the assumptions given in (A.17), this simplifies to

$$\begin{aligned} Var(t_h) &= m_h^{-2}\{ E[\sum_{c=1}^{n_h}\sum_{i=1}^{m_{hc}} (t_{hci} - \tau_h)^2]+E[\sum_{c=1}^{n_h}\sum_{i=1}^{m_{hc}}\sum_{j\neq i}^{m_{hc}} (t_{hci} - \tau_h)(t_{hcj} - \tau_h)]\ \} \\ &= m_h^{-1}Var(t_{hci}) + m_h^{-2}[\sum_{c=1}^{n_h} m_{hc}(m_{hc}-1)]\rho_h Var(t_{hci}) \\ &= m_h^{-1}Var(t_{hci})[1 + (m_h^* - 1)\rho_h]. \end{aligned} \tag{A.20}$$

where m_h^* has already been defined in (12).

References: The main reference for section 1 of this Appendix is Hansen, Hurwitz and Madow (1953); for section 2, Scheaffer, Mendenhall and Ott (1990); for section 3, Deaton (1994).

References

Ainsworth, M. and Munoz, J. (1986) "The Cote d'Ivoire Living Standards Survey: design and implementation", LSMS Working Paper No.26, The World Bank.

Cleland, J. and Scott, C., eds. (1987) *The World Fertility Survey: an assessment*, Oxford University Press.

Deaton, A. (1994) *The Analysis of Household Surveys: Microeconometric Analysis for Development Policy*, World Bank.

Duncan, G. and Rodgers, W. (1991) "Has Children's Poverty Become More Persistent", *American Sociological Review*, 56, 538-550.

Hansen, M.H., Hurwitz, W.N. and Madow, W.G. (1953) *Sample Survey Methods and Theory, Volume II - Theory*, John Wiley and Sons, New York.

Harpman, T. and Scott, C. (1987) "Sample Implementation" in Cleland, J. and Scott, C. (eds.)

Howes, S. (1993) "Income Distribution: measurement, transition and analysis of urban China, 1981-1990", PhD thesis, London School of Economics.

Howes, S. and Zaidi, S. (1994) "Notes on Some Household Surveys from Pakistan in the eighties and nineties", mimeo, The World Bank.

Kakwani, N. (1993) "Statistical Inference in the Measurement of Poverty", *Review of Economics and Statistics*, 632-639.

Levy, P. and Lemeshew, S. (1991) *Sampling of Populations: methods and applications*, John Wiley and Sons, New York.

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

Kish, L. and Frankel, M. (1970) "Balanced Repeated Replications for Standard Errors", *Journal of the American Statistical Association*, 65 (331), 1071-1093.

PRDPH, World Bank (1994) "Jamaica Survey of Living Conditions (JSLC), 1988-1992: Basic Information", mimeo.

Pudney, S. and Sutherland, H. (1994) "How Reliable are Microsimulation Estimates? An investigation of the role of sampling error in a UK tax-benefit model", *Journal of Public Economics*.

Ravallion, M. (1994) *Poverty Comparisons*, Harwood Academic Publishers.

Rodgers, J. and Rodgers, J. (1992) "Chronic Poverty in the United States" *The Journal of Human Resources*, 28(1), 26-54.

Scheaffer, R., Mendenhall, W. and Ott, L. (1990) *Elementary Survey Sampling*, Fourth Edition, Duxbury Press, California.

Scott, C. and Amenuvegbe, B. (1989) "Sample Designs for Living Standards Surveys in Ghana and Mauritania", LSMS Working Paper No.49, World Bank.

Scott, C. and Amenuvegbe, B. (1991) "Reconciling Fixed Interviewer Workloads With Self-Weighting Sampling When Size Measures are Defective", *Journal of Official Statistics*, 7(3), 367-373.

Som, R.K. (1973) *A Manual of Sampling Techniques*, Heinemann, London.

LSMS Working Papers

No. 1 Chander, Grootaert, and Pyatt, *Living Standards Surveys in Developing Countries*

No. 2 Visaria, *Poverty and Living Standards in Asia: An Overview of the Main Results and Lessons of Selected Household Surveys*

No. 3 United Nations Statistical Office, *Measuring Levels of Living in Latin America: An Overview of Main Problems*

No. 4 Scott, de Andre, and Chander, *Towards More Effective Measurement of Levels of Living, and Review of Work of the United Nations Statistical Office (UNSO) Related to Statistics of Levels of Living*

No. 5 Scott, de Andre, and Chander, *Conducting Surveys in Developing Countries: Practical Problems and Experience in Brazil, Malaysia, and the Philippines*

No. 6 Booker, de Andre, and Chander, *Household Survey Experience in Africa*

No. 7 Deaton, *Measurement of Welfare: Theory and Practical Guidelines*

No. 8 Mehran, *Employment Data for the Measurement of Living Standards*

No. 9 Wahab, *Income and Expenditure Surveys in Developing Countries: Sample Design and Execution*

No. 10 Saunders and Grootaert, *Reflections on the LSMS Group Meeting*

No. 11 Deaton, *Three Essays on a Sri Lanka Household Survey*

No. 12 Musgrove, *The ECIEL Study of Household Income and Consumption in Urban Latin America: An Analytical History*

No. 13 Martorell, *Nutrition and Health Status Indicators: Suggestions for Surveys of the Standard of Living in Developing Countries*

No. 14 Birdsall, *Child Schooling and the Measurement of Living Standards*

No. 15 Ho, *Measuring Health as a Component of Living Standards*

No. 16 Sullivan, Cochrane, and Kalsbeek, *Procedures for Collecting and Analyzing Mortality Data in LSMS*

No. 17 Grootaert, *The Labor Market and Social Accounting: A Framework of Data Presentation*

No. 18 Acharya, *Time Use Data and the Living Standards Measurement Study*

No. 19 Grootaert, *The Conceptual Basis of Measures of Household Welfare and Their Implied Survey Data Requirements*

No. 20 Grootaert, Cheung, Fung, and Tam, *Statistical Experimentation for Household Surveys: Two Cases Studies of Hong Kong*

No. 21 Wood and Knight, *The Collection of Price Data for the Measurement of Living Standards*

No. 22 Grootaert and Cheung, *Household Expenditure Surveys: Some Methodological Issues*

No. 23 Ashenfelter, Deaton, and Solon, *Collecting Panel Data in Developing Countries: Does it Make Sense?*

No. 24 Grootaert, *Measuring and Analyzing Levels of Living in Developing Countries: An Annotated Questionnaire*

No. 25 Grootaert and Dubois, *The Demand for Urban Housing in the Ivory Coast*

No. 26 Ainsworth and Muñoz, *The Côte d'Ivoire Living Standards Survey: Design and Implementation*

No. 27 Grootaert, *The Role of Employment and Earnings in Analyzing Levels of Living: A General Methodology with Applications to Malaysia and Thailand*

No. 28 Deaton and Case, *Analysis of Household Expenditures*

No. 29 Glewwe, *The Distribution of Welfare in Côte d'Ivoire in 1985*

No. 30 Deaton, *Quality, Quantity, and Spatial Variation of Price: Estimating Price Elasticities from Cross-sectional Data*

No. 31 Suarez-Berenguela, *Financing the Health Sector in Peru*

No. 32 Suarez-Berenguela, *Informal Sector, Labor Markets, and Returns to Education in Peru*

No. 33 van der Gaag and Vijverberg, *Wage Determinants in Côte d'Ivoire*

No. 34 Ainsworth and van der Gaag, *Guidelines for Adapting the LSMS Living Standards Questionnaires to Local Conditions*

No. 35 Dor and van der Gaag, *The Demand for Medical Care in Developing Countries: Quantity Rationing in Rural Côte d'Ivoire*

No. 36 Newman, *Labor Market Activity in Côte d'Ivoire and Peru*

No. 37 Gertler, Locay, Sanderson, Dor, and van der Gaag, *Health Care Financing and the Demand for Medical Care*

No. 38 Stelcner, Arriagada, and Moock, *Wage Determinants and School Attainment among Men in Peru*

No. 39 Deaton, *The Allocation of Goods within the Household: Adults, Children, and Gender*

No. 40 Strauss, *The Effects of Household and Community Characteristics on the Nutrition of Preschool Children: Evidence from rural Côte d'Ivoire*

No. 41 Stelcner, van der Gaag, and Vijverberg, *Public-Private Sector Wage Differentials in Peru, 1985-86*

No. 42 Glewwe, *The Distribution of Welfare in Peru 1985-86*

No. 43 Vijverberg, *Profits from Self-Employment: A Case Study of Côte d'Ivoire*

No. 44 Deaton and Benjamin, *The Living Standards Survey and Price Policy Reform: A Study of Cocoa and Coffee Production in Côte d'Ivoire*

No. 45 Gertler and van der Gaag, *Measuring the Willingness to Pay for Social Services in Developing Counties*

No. 46 Vijverberg, *Nonagricultural Family Enterprises in Côte d'Ivoire: A Descriptive Analysis*

No. 47 Glewwe and de Tray, *The Poor during Adjustment: A Case Study of Côte d'Ivoire*

No. 48 Glewwe and van der Gaag, *Confronting Poverty in Developing Countries: Definitions, Information, and Policies*

No. 49 Scott and Amenuvegbe, *Sample Design for the Living Standards Surveys in Ghana and Mauritania/Plans de sondage pour les enquêtes sur le niveau de vie au Ghana et en Mauritanie*

No. 50 Laraki, *Food Subsidies: a Case Study of Price Reform in Morocco* (also in French, 50F)

No. 51 Strauss and Mehra, *Child Anthropometry in Côte d'Ivoire: Estimates from Two Surveys, 1985 and 1986*

No. 52 van der Gaag, Stelcner, and Vijverberg, *Public-Private Sector Wage Comparisons and Moonlighting in Developing Countries: Evidence from Côte d'Ivoire and Peru*

No. 53 Ainsworth, *Socioeconomic Determinants of Fertility in Côte d'Ivoire*

No. 54 Gertler and Glewwe, *The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru*

No. 55 Levy and Newman, *Rigidité des salaires: Données microéconomiques et macroéconomiques sur l'ajustement du marché du travail dans le secteur moderne* (in French only)

No. 56 Glewwe and de Tray, *The Poor in Latin America during Adjustment: A Case Study of Peru*

No. 57 Alderman and Gertler, *The Substitutability of Public and Private Health Care of the Treatment of Children in Pakistan*

No. 58 Rosenhouse, *Identifying the Poor: Is "Headship" a Useful Concept?*

No. 59 Vijverberg, *Labor Market Performance as a Determinant of Migration*

No. 60 Jimenez and Cox, *The Relative Effectiveness of Private and Public Schools: Evidence from Two Developing Countries*

No. 61 Kakwani, *Large Sample Distribution of Several Inequality Measures: With Application to Côte d'Ivoire*

No. 62 Kakwani, *Testing for Significance of Poverty Differences: With Application to Côte d'Ivoire*

No. 63 Kakwani, *Poverty and Economic Growth: With Application to Côte d'Ivoire*

No. 64 Moock, Musgrove, and Stelcner, *Education and Earnings in Peru's Informal Nonfarm Family Enterprises*

No. 65 Alderman and Kozel, *Formal and Informal Sector Wage Determination in Urban Low-Income Neighborhoods in Pakistan*

No. 66 Vijverberg and van der Gaag, *Testing for Labor Market Duality: The Private Wage Sector in Côte d'Ivoire*

No. 67 King, *Does Education Pay in the Labor Market? The Labor Force Participation, Occupation, Occupation, and Earnings of Peruvian Women*

No. 68 Kozel, *The composition and Distribution of Income in Côte d'Ivoire*

No. 69 Deaton, *Price Elasticities from Survey Data: Extensions and Indonesian Results*

No. 70 Glewwe, *Efficient Allocation of Transfers to the Poor: The Problem of Unobserved Household Income*

No. 71 Glewwe, *Investigating the Determinants of Household Welfare in Côte d'Ivoire*

No. 72 Pitt and Rosenzweig, *The Selectivity of Fertility and the Determinants of Human Capital Investments: Parametric and Semiparametric Estimates*

No. 73 Jacoby, *Shadow Wages and Peasant Family Labor Supply: An Econometric Application to the Peruvian Sierra*

No. 74 Behrman, *The Action of Human Resources and Poverty on One Another: What We Have Yet to Learn*

No. 75 Glewwe and Twum-Baah, *The Distribution of Welfare in Ghana, 1987-88*

No. 76 Glewwe, *Schooling, Skills, and the Returns to Government Investment in Education: An Exploration Using Data from Ghana*

No. 77 Newman, Jorgensen, and Pradhan, *Workers' Benefits from Bolivia's Emergency Social Fund*

No. 78 Vijverberg, *Dual Selection Criteria with Multiple Alternatives: Migration, Work Status, and Wages*

No. 79 Thomas, *Gender Differences in Household Resource Allocations*

No. 80 Grosh, *The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions*

No. 81 Deaton and Paxson, *Patterns of Aging in Thailand and Côte d'Ivoire*

No. 82 Ravallion, *Does Undernutrition Respond to Incomes and Prices? Dominance Tests for Indonesia*

No. 83 Ravallion and Datt, *Growth and Redistribution Components of Changes in Poverty Measure: A Decomposition with Applications to Brazil and India in the 1980s*

No. 84 Vijverberg, *Measuring Income from Family Enterprises with Household Surveys*

No. 85 Deaton and Grimard, *Demand Analysis and Tax Reform in Pakistan*

No. 86 Glewwe and Hall, *Poverty and Inequality during Unorthodox Adjustment: The Case of Peru, 1985-90*

No. 87 Newman and Gertler, *Family Productivity, Labor Supply, and Welfare in a Low-Income Country*

No. 88 Ravallion, *Poverty Comparisons: A Guide to Concepts and Methods*

No. 89 Thomas, Lavy, and Strauss, *Public Policy and Anthropometric Outcomes in Côte d'Ivoire*

No. 90 Ainsworth and others, *Measuring the Impact of Fatal Adult Illness in Sub-Saharan Africa: An Annotated Household Questionnaire*

No. 91 Glewwe and Jacoby, Estimating the Determinants of Cognitive Achievement in Low-Income Countries: The Case of Ghana

No. 92 Ainsworth, *Economic Aspects of Child Fostering in Côte d'Ivoire*

No. 93 Lavy, *Investment in Human Capital: Schooling Supply Constraints in Rural Ghana*

No. 94 Lavy and Quigley, *Willingness to Pay for the Quality and Intensity of Medical Care: Low-Income Household in Ghana*

No. 95 Schultz and Tansel, *Measurement of Returns to Adult Health: Morbidity Effects on Wage Rates in Côte d'Ivoire and Ghana*

No. 96 Louat, Grosh, and van der Gaag, *Welfare Implications of Female Headship in Jamaican Household*

No. 97 Colombe and Demery, *Household Size in Côte d'Ivoire: Sampling Bias in the CILSS*

No. 98 Glewwe and Jacoby, *Delayed Primary School Enrollment and Childhood Malnutrition in Ghana: An Economic Analysis*

No. 99 Baker and Grosh, *Poverty Reduction through Geographic Targeting: How Well Does It Work?*

No. 100 Datt and Ravallion, *Income Gains for the Poor from Public Works Employment: Evidence from Two Indian Villages*

No. 101 Kostermans, *Assessing the Quality of Anthropometric Data: Background and Illustrated Guidelines for Survey Manager*

No. 102 van de Walle, Ravallion, and Gautam, *How Does the Social Safety Net Work?: The Incidence of Cash Benefits in Hungary, 1987-89*

No. 103 Benefo and Schultz, *Determinants of Fertility and Child Mortality in Côte d'Ivoire and Ghana*

No. 104 Behrman and Lavy, *Children's Health and Achievement in School*

No. 105 Lavy and Germain, *Quality and Cost in Health Care Choice in Developing Countries*

No. 106 Lavy, Strauss, Thomas, and De Vreyer, *The Impact of the Quality of Health Care on Children's Nutrition and survival in Ghana*

No. 107 Hanushek and Lavy, *School Quality, Achievement Bias, and Dropout Behavior in Egypt*

No. 108 Feyistan and Ainsworth, *Contraceptive Use and the Quality, Price, and Availability of Family Planning*

No. 109 Thomas and Maluccio, *Contraceptive Choice, Fertility, and Public Policy in Zimbabwe*

No. 110 Ainsworth, Beegle, and Nyamete, *The Impact of Female Schooling on Fertility and Contraceptive Use: A Study of Fourteen Sub-Saharan Countries*

No. 111 Oliver, *Contraceptive Use in Ghana: The Role of Service Availability, Quality, and Price*

No. 112 Montgomery, Kouamé, and Oliver, *The Tradeoff between Number of Children and Child Schooling: Evidence from Côte d'Ivoire and Ghana*

No. 113 Pradhan, *Sector Participation Decisions in Labor Supply Models*

No. 114 Beegle, *The Quality and Availability of Family Planning Services and Contraceptive Use in Tanzania*

No. 115 Lavy, Spratt, and Leboucher, *Changing Patterns of Illiteracy in Morocco: Assessment Methods Compared*

No. 116 Lavy, Palumbo, and Stern, *Health Car in Jamaica: Quality, Outcomes, and Labor Supply*

No. 117 Glewwe and Hall, *Who Is Most Vulnerable to Macroeconomic Shocks? Hypotheses Tests Using Panel Data from Peru*

No. 118 Grosh and Baker, *Proxy Means Tests for Targeting Social Programs: Simulations and Speculation*

No. 119 Pitt, *Women's Schooling, the Selectivity of Fertility, and Child Mortality in Sub-Saharan Africa*

No. 120 Grosh and Glewwe, *A Guide to Living Standards Measurement Study Surveys and their Data Sets*

No. 121 van de Walle, *Infrastructure and Poverty in Viet Nam*

No. 122 Ravallion, *Poverty Comparisons: A Guide to Concepts and Methods*

No. 123 Ii, *The Demand for Medical Care: Evidence from Urban Areas in Bolivia*

No. 124 Hentschel and Lanjouw, *Constructing and Indicator of Consumption for the Analysis of Poverty: Principles and Illustrations with Reference to Ecuador*

No. 125 Leibrrandt, Woolard, Woolard, *The Contribution of Income Components to Income Inequality in South Africa: A Decomposable Gini Analysis*

No. 126 Grosh, Muñoz, *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*

No. 127 Dow, *Unconditional Demand for Curative Health Inputs: Does Selection on Health Status Matter in the Long Run?*

Distributors of World Bank Publications

Prices and credit terms vary from country to country. Consult your local distributor before placing an order.

ARGENTINA
Oficina del Libro Internacional
Av Cordoba 1877
1120 Buenos Aires
Tel (54 1) 815-8354
Fax (54 1) 815-8156

AUSTRALIA, FIJI, PAPUA NEW GUINEA, SOLOMON ISLANDS, VANUATU, AND WESTERN SAMOA
D A Information Services
648 Whitehorse Road
Mitcham 3132
Victoria
Tel (61) 3 9210 7777
Fax (61) 3 9210 7788
E-mail service@dadirect com au
URL http.//www dadirect com au

AUSTRIA
Gerold and Co
Weihburggasse 26
A-1011 Wien
Tel (43 1) 512-47-31-0
Fax (43 1) 512-47-31-29
URL http.//www gerold co/at online

BANGLADESH
Micro Industries Development Assistance Society (MIDAS)
House 5, Road 16
Dhanmondi R/Area
Dhaka 1209
Tel (880 2) 326427
Fax (880 2) 811188

BELGIUM
Jean De Lannoy
Av du Roi 202
1060 Brussels
Tel (32 2) 538-5169
Fax (32 2) 538-0841

BRAZIL
Publicações Tecnicas Internacionais Ltda
Rua Peixoto Gomide, 209
01409 Sao Paulo, SP
Tel (55 11) 259-6644
Fax (55 11) 258-6990
E-mail postmaster@pti uol br
URL http.//www uol br

CANADA
Renouf Publishing Co Ltd
5369 Canotek Road
Ottawa, Ontario K1J 9J3
Tel (613) 745-2665
Fax (613) 745-7660
E-mail renouf@fox nstn ca
URL http.// www fox nstn ca/~renouf

CHINA
China Financial & Economic Publishing House
8, Da Fo Si Dong Jie
Beijing
Tel (86 10) 6333-8257
Fax (86 10) 6401-7365

COLOMBIA
Infoenlace Ltda.
Carrera 6 No. 51-21
Apartado Aereo 34270
Santafé de Bogotá, D C.
Tel (57 1) 285-2798
Fax. (57 1) 285-2798

COTE D'IVOIRE
Center d'Edition et de Diffusion Africaines (CEDA)
04 B P 541
Abidjan 04
Tel (225) 24 6510,24 6511
Fax (225) 25 0567

CYPRUS
Center for Applied Research
Cyprus College
6, Diogenes Street, Engomi
PO Box 2006
Nicosia
Tel (357 2) 44-1730
Fax (357 2) 46-2051

CZECH REPUBLIC
National Information Center
prodejna, Konviktska 5
CS – 113 57 Prague 1
Tel (42 2) 2422-9433
Fax (42 2) 2422-1484
URL http.//www nis cz/

DENMARK
SamfundsLitteratur
Rosenoerns Allé 11
DK-1970 Fredeniksberg C
Tel (45 31) 351942
Fax (45 31) 357822

EGYPT, ARAB REPUBLIC OF
Al Ahram Distribution Agency
Al Galaa Street
Cairo
Tel (20 2) 578-6083
Fax (20 2) 578-6833

The Middle East Observer
41, Sherif Street
Cairo
Tel (20 2) 393-9732
Fax (20 2) 393-9732

FINLAND
Akateeminen Kirjakauppa
PO Box 128
FIN-00101 Helsinki
Tel (358 0) 12141
Fax (358 0) 121-4441
URL. http.//www akateeminen com/

FRANCE
World Bank Publications
66, avenue d'Iena
75116 Paris
Tel (33 1) 40-69-30-56/57
Fax (33 1) 40-69-30-68

GERMANY
UNO-Verlag
Poppelsdorfer Allee 55
53115 Bonn
Tel (49 228) 212940
Fax (49 228) 217492

GREECE
Papasotiriou S A
35, Stoumara Str
106 82 Athens
Tel (30 1) 364-1826
Fax (30 1) 364-8254

HAITI
Culture Diffusion
5, Rue Capois
C P 257
Port-au-Prince
Tel (509 1) 3 9260

HONG KONG, MACAO
Asia 2000 Ltd
Sales & Circulation Department
Seabird House, unit 1101-02
22-28 Wyndham Street, Central
Hong Kong
Tel (852) 2530-1409
Fax (852) 2526-1107
E-mail sales@asia2000 com hk
URL http.//www asia2000 com hk

INDIA
Allied Publishers Ltd
751 Mount Road
Madras - 600 002
Tel (91 44) 852-3938
Fax (91 44) 852-0649

INDONESIA
Pt Indira Limited
Jalan Borobudur 20
PO Box 181
Jakarta 10320
Tel (62 21) 390-4290
Fax (62 21) 421-4289

IRAN
Ketab Sara Co Publishers
Khaled Eslamboli Ave ,
6th Street
Kusheh Delafrooz No 8
PO Box 15745-733
Tehran
Tel (98 21) 8717819, 8716104
Fax (98 21) 8712479
E-mail ketab-sara@neda net ir

Kowkab Publishers
PO Box 19575-511
Tehran
Tel (98 21) 258-3723
Fax (98 21) 258-3723

IRELAND
Government Supplies Agency
Oifig an tSoláthair
4-5 Harcourt Road
Dublin 2
Tel (353 1) 661-3111
Fax (353 1) 475-2670

ISRAEL
Yozmot Literature Ltd
PO Box 56055
3 Yohanan Hasandlar Street
Tel Aviv 61560
Tel (972 3) 5285-397
Fax (972 3) 5285-397

R O Y International
PO Box 13056
Tel Aviv 61130
Tel (972 3) 5461423
Fax (972 3) 5461442
E-mail royil@netvision net il

Palestinian Authority/Middle East Index Information Services
PO B 19502 Jerusalem
Tel (972 2) 6271219
Fax (972 2) 6271634

ITALY
Licosa Commissionaria Sansoni SPA
Via Duca Di Calabria, 1/1
Casella Postale 552
50125 Firenze
Tel (55) 645-415
Fax (55) 641-257
E-mail licosa@ftbcc it
Url http.//www ftbcc it/licosa

JAMAICA
Ian Randle Publishers Ltd
206 Old Hope Road
Kingston 6
Tel 809-927-2085
Fax 809-977-0243
E-mail irpl@colis com

JAPAN
Eastern Book Service
3-13 Hongo 3-chome, Bunkyo-ku
Tokyo 113
Tel (81 3) 3818-0861
Fax (81 3) 3818-0864
E-mail svt-ebs@ppp bekkoame or jp
URL http.//www bekkoame or jp/~svt-ebs

KENYA
Africa Book Service (E A) Ltd
Quaran House, Mfangano Street
PO Box 45245
Nairobi
Tel (254 2) 223 641
Fax (254 2) 330 272

KOREA, REPUBLIC OF
Daejon Trading Co Ltd
PO Box 34, Youida
706 Seoun Bldg
44-6 Youido-Dong, Yeongchengpo-Ku
Seoul
Tel (82 2) 785-1631/4
Fax (82 2) 784-0315

MALAYSIA
University of Malaya Cooperative Bookshop, Limited
PO Box 1127
Jalan Pantai Baru
59700 Kuala Lumpur
Tel (60 3) 756-5000
Fax (60 3) 755-4424

MEXICO
INFOTEC
Av San Fernando No 37
Col Tlalpan Guerra
14050 Mexico, D F
Tel (52 5) 624-2800
Fax (52 5) 624-2822
E-mail infotec@rtn net mx
URL http.//rtn net mx

NEPAL
Everest Media International Services (P)Ltd
GPO Box 5443
Kathmandu
Tel (977 1) 472 152
Fax (977 1) 224 431

NETHERLANDS
De Lindeboom/InOr-Publikaties
PO Box 202
7480 AE Haaksbergen
Tel (31 53) 574-0004
Fax (31 53) 572-9296
E-mail lindeboo@worldonline nl
URL http.//www worldonline nl/~lindeboo

NEW ZEALAND
EBSCO NZ Ltd
Private Mail Bag 99914
New Market
Auckland
Tel (64 9) 524-8119
Fax (64 9) 524-8067

NIGERIA
University Press Limited
Three Crowns Building Jericho
Private Mail Bag 5095
Ibadan
Tel (234 22) 41-1356
Fax (234 22) 41-2056

NORWAY
NIC Info A/S
Book Department
PO Box 6125 Etterstad
N-0602 Oslo 6
Tel (47 22) 57-3300
Fax (47 22) 68-1901

PAKISTAN
Mirza Book Agency
65, Shahrah-e-Quaid-e-Azam
Lahore 54000
Tel (92 42) 735 3601
Fax (92 42) 758 5283

Oxford University Press
5 Bangalore Town
Sharae Faisal
PO Box 13033
Karachi-75350
Tel (92 21) 446307
Fax (92 21) 4547640
E-mail oup@oup khi erum com pk

Pak Book Corporation
Aziz Chambers 21
Queen's Road
Lahore
Tel (92 42) 636 3222, 636 0885
Fax (92 42) 636 2328
E-mail pbc@brain net pk

PERU
Editorial Desarrollo SA
Apartado 3824
Lima 1
Tel (51 14) 285380
Fax (51 14) 286628

PHILIPPINES
International Booksource Center Inc
1127-A Antipolo St
Barangay, Venezuela
Makati City
Tel (63 2) 896 6501, 6505, 6507
Fax (63 2) 896 1741

POLAND
International Publishing Service
Ul Piekna 31/37
00-677 Warzawa
Tel (48 2) 628-6089
Fax (48 2) 621-7255
E-mail books%ips@ikp atm com pl
URL http.//www ipscg waw pl/ips/export/

PORTUGAL
Livraria Portugal
Apartado 2681
Rua Do Carmo 70-74
1200 Lisbon
Tel (1) 347-4982
Fax (1) 347-0264

ROMANIA
Compani De Librarii Bucuresti S A
Str Lipscani no. 26, sector 3
Bucharest
Tel (40 1) 613 9645
Fax. (40 1) 312 4000

RUSSIAN FEDERATION
Isdatelstvo <Ves Mir>
9a, Lotpachniy Pereulok
Moscow 101831
Tel (7 095) 917 87 49
Fax (7 095) 917 92 59

SINGAPORE, TAIWAN, MYANMAR, BRUNEI
Asahgate Publishing Asia Pacific Pte Ltd
41 Kallang Pudding Road #04-03
Golden Wheel Building
Singapore 349316
Tel (65) 741-5166
Fax (65) 742-9356
E-mail ashgate@asianconnect com

SLOVENIA
Gospodarski Vestnik Publishing Group
Dunajska cesta 5
1000 Ljubljana
Tel (386 61) 133 83 47, 132 12 30
Fax (386 61) 133 80 30
E-mail belicd@gvestnik si

SOUTH AFRICA, BOTSWANA
For single titles
Oxford University Press Southern Africa
PO Box 1141
Cape Town 8000
Tel (27 21) 45-7266
Fax (27 21) 45-7265

For subscription orders
International Subscription Service
PO Box 41095
Craighall
Johannesburg 2024
Tel (27 11) 880-1448
Fax (27 11) 880-6248
E-mail iss@is co za

SPAIN
Mundi-Prensa Libros, S A
Castello 37
28001 Madrid
Tel (34 1) 431-3399
Fax (34 1) 575-3998
E-mail libreria@mundiprensa es
URL http.//www mundiprensa es/

Mundi-Prensa Barcelona
Consell de Cent, 391
08009 Barcelona
Tel (34 3) 488-3492
Fax (34 3) 487-7659
E-mail barcelona@mundiprensa es

SRI LANKA, THE MALDIVES
Lake House Bookshop
100, Sir Chittampalam Gardiner Mawatha
Colombo 2
Tel (94 1) 32105
Fax (94 1) 432104
E-mail LHL@sri lanka net

SWEDEN
Wennergren-Williams AB
P O Box 1305
S-171 25 Solna
Tel (46 8) 705-97-50
Fax (46 8) 27-00-71
E-mail mail@wwi se

SWITZERLAND
Librairie Payot Service Institutionnel
Côtes-de-Montbenon 30
1002 Lausanne
Tel (41 21) 341-3229
Fax (41 21) 341-3235

ADECO Van Diermen EditionsTechniques
Ch de Lacuez 41
CH1807 Blonay
Tel (41 21) 943 2673
Fax (41 21) 943 3605

TANZANIA
Oxford University Press
Maktaba Street
PO Box 5299
Dar es Salaam
Tel (255 51) 29209
Fax (255 51) 46822

THAILAND
Central Books Distribution
306 Silom Road
Bangkok 10500
Tel (66 2) 235-5400
Fax (66 2) 237-8321

TRINIDAD & TOBAGO, AND THE CARRIBBEAN
Systematics Studies Unit
9 Watts Street
Curepe
Trinidad, West Indies
Tel (809) 662-5654
Fax (809) 662-5654
E-mail tobe@trinidad net

UGANDA
Gustro Ltd
PO Box 9997, Madhvani Building
Plot 16/4 Jinja Rd
Kampala
Tel (256 41) 254 763
Fax (256 41) 251 468

UNITED KINGDOM
Microinfo Ltd
PO Box 3
Alton, Hampshire GU34 2PG
England
Tel (44 1420) 86848
Fax (44 1420) 89889
E-mail wbank@ukminfo demon co uk
URL http.//www microinfo co uk

VENEZUELA
Tecni-Ciencia Libros, S A
Centro Ciudad Comercial Tamanco
Nivel C2
Caracas
Tel (58 2) 959 5547, 5035, 0016
Fax (58 2) 959 5636

ZAMBIA
University Bookshop, University of Zambia
Great East Road Campus
PO Box 32379
Lusaka
Tel (260 1) 252 576
Fax (260 1) 253 952

ZIMBABWE
Longman Zimbabwe (Pte)Ltd
Tourle Road, Ardbennie
PO Box ST125
Southerton
Harare
Tel (263 4) 6216617
Fax (263 4) 621670

03/20/97

THE WORLD BANK
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Telephone: (202) 477-1234
Facsimile: (202) 477-6391
Telex: MCI 64145 WORLDBANK
MCI 248423 WORLDBANK
Cable Address: INTBAFRAD
WASHINGTONDC
World Wide Web: http://www.worldbank.org/
E-mail: books@worldbank.org

13862
9 780821 338629
ISBN 0-8213-3862-5