Scientific Reports

NUMBER 55 APRIL 1984

ALBERT M. MARCKWARDT

Response Rates, Callbacks and Coverage: The WFS Experience

INTERNATIONAL STATISTICAL INSTITUTE
Permanent Office. Director: E. Lunenberg
428 Prinses Beatrixlaan, PO Box 950
2270 AZ Voorburg
Netherlands

WORLD FERTILITY SURVEY
Project Director: Halvor Gille
35–37 Grosvenor Gardens
London SW1W 0BS
United Kingdom

The World Fertility Survey is an international research programme whose purpose is to assess the current state of human fertility throughout the world. This is being done principally through promoting and supporting nationally representative, internationally comparable, and scientifically designed and conducted sample surveys of fertility behaviour in as many countries as possible.
The WFS is being undertaken, with the collaboration of the United Nations, by the International Statistical Institute in cooperation with the International Union for the Scientific Study of Population. Financial support is provided principally by the United Nations Fund for Population Activities and the United States Agency for International Development.
This publication is part of the WFS Publications Programme which includes the WFS Basic Documentation, Occasional Papers and auxiliary publications. For further information on the WFS, write to the Information Office, International Statistical Institute, 428 Prinses Beatrixlaan, Voorburg, The Hague, Netherlands.

L'Enquête Mondiale sur la Fécondité (EMF) est un programme international de recherche dont le but est d'évaluer l'état actuel de la fécondité humaine dans le monde. Afin d'atteindre cet objectif, des enquêtes par sondage sur la fécondité sont mises en oeuvre et financées dans le plus grand nombre de pays possible, Ces études, élaborées et réalisées de façon scientifique, fournissent des données représentatives au niveau national et comparables au niveau international.
L'Institut International de Statistique avec l'appui des Nations Unies, a été chargé de la réalisation de ce projet en collaboration avec l'Union Internationale pour l'Etude Scientifique de la Population. Le financement est principalement assuré par le Fonds des Nations Unies pour les Activités en matière de Population et l'Agence pour le Développement International des Etats-Unis.
Cette publication fait partie du programme de publications de l'EMF, qui comprend la Documentation de base, les Documents Non-Périodiques et des publications auxiliaires. Pour tout renseignement complémentaire, s'adresser au Bureau d'Information, Institut International de Statistique, 428 Prinses Beatrixlaan, Voorburg, La Haye, Pays-Bas.

La Encuesta Mundial de Fecundidad (EMF) es un programa internacional de investigación cuyo propósito es determinar el estado actual de la fecundidad humana en el mundo. Para lograr este objetivo, se están promoviendo y financiando encuestas de fecundidad por muestreo en el mayor número posible de países. Estas encuestas son diseñadas y realizadas científicamente, nacionalmente representativas y comparables a nivel internacional.
El proyecto está a cargo del Instituto Internacional de Estadística en cooperación con la Unión Internacional para el Estudio Científico de la Población y con la colaboración de las Naciones Unidas. Es financiado principalmente por el Fondo de las Naciones Unidas para Actividades de Población y por la Agencia para el Desarrollo Internacional de los Estados Unidos.
Esta publicación ha sido editada por el Programa de Publicaciones de la EMF, el que incluye Documentación Básica, Publicaciones Ocasionales y publicaciones auxiliares. Puede obtenerse mayor información sobre la EMF escribiendo a la Oficina de Información, Instituto Internacional de Estadística, 428 Prinses Beatrixlaan, Voorburg-La Haya, Países Bajos.

Scientific Reports

Response Rates, Callbacks and Coverage: The WFS Experience

ALBERT M. MARCKWARDT
WFS Central Staff

The recommended citation for this publication is:

Marckwardt, Albert M. (1984). Response Rates, Callbacks and Coverage: the WFS Experience. *WFS Scientific Reports* no 55. Voorburg, Netherlands: International Statistical Institute.

Printed in Great Britain at the Alden Press
Oxford London and Northampton

Contents

Preface

A critical assessment of survey experience and data quality is an integral part of the WFS programme. This assessment aims at ensuring that analyses are carried out with as full an understanding as possible of the quality and reliability of the data and at drawing lessons for the better conduct of future surveys.

This report contains an analysis and appraisal of WFS experience of sample implementation, with particular reference to response rates and coverage errors. Important implications for future survey work emerge. For instance, Dr Marckwardt adduces evidence that sample units which required the extra effort of re-visits differ appreciably in their characteristics from units which were successfully interviewed at the first visit. This type of evidence is rarely available for developing country surveys and thus I believe that this report has made an important contribution to survey methodology.

HALVOR GILLE
Project Director

Acknowledgements

I am indebted to several colleagues at the WFS London headquarters for having read and commented on various drafts of this study. In particular, I would like to thank John Cleland, Chris Scott and Martin Vaessen for their valuable suggestions. Also, I thank Vijay Verma, of the United Nations Statistical Office, for his useful commentary on an early draft. Finally, I should note that much of the work on callbacks was inspired by suggestions made originally by Alan Sunter.

1 Introduction

The achievement of high response rates is of great importance in survey research. It is axiomatic that confident generalizations can be made about the universe of study only if probability sampling has been adhered to and a high level of response obtained. The reason for this is that non-response to household surveys may be, and in practice is generally found to be, selective of certain segments of the population. If non-response is significant, the resulting mix of respondents to the survey will then not reflect the composition of the universe, and findings will be biased. But the achievement of high rates of response can be costly, often requiring various callbacks to households where for one reason or another an interview was not obtained on the first visit.

Non-response represents the failure to obtain information from a sample unit. Non-response may be either total or partial. By total non-response we mean the failure to obtain an interview. Partial non-response, often called item non-response, occurs when the respondent fails to answer a particular question, or the interviewer fails to record the answer. Our attention here will be restricted to total non-response.

In this document we will be concerned with: (a) documenting the response rates achieved in the countries participating in the World Fertility Survey; (b) analysing, for a selected group of countries, differential non-response to the individual interview; (c) studying, for another group of countries, the benefits of callbacks to households where an interview was not obtained at the first visit; (d) presenting some evidence on field work efficiency and the cost, in time, of callbacks; and, finally (e) examining the reasons for serious shortfalls in the number of completed interviews relative to the number planned in many countries, even when respectable response rates are achieved.

The limitation to 'selected' countries for most topics is not for reasons of economy, but simply because the data are not readily available for many countries. Some WFS country reports have presented only very sketchy documentation of response rates, and these were often based on manual tabulations of field records. In the preparation of recoded data files for tabulating results, it has been normal WFS practice to exclude cases of non-response. This is true of both individual and household data files. Hence, at this late date it is very difficult to document reliable response rates and their differentials for many countries. Nevertheless, from the information available, useful findings do emerge and it is possible to make a number of generalizations.

2 Response Rates

2.1 RESPONSE RATES ACROSS COUNTRIES

In general, non-response to household surveys is not as serious a problem in developing countries as in developed western countries, where it can run as high as 20–30 per cent (United Nations 1982). It will be seen that the WFS experience supports this generalization. Operationally, WFS surveys involved first interviewing a sample of dwellings or households to obtain a list of women eligible for an intensive interview, and then either taking a subset or a complete set of these for administration of the detailed interview. Response rate can hence be calculated for both the household and individual interviews; these are presented in tables A1 and A2, respectively, for the 42 countries which have participated in the WFS developing-country programme.

As noted above, non-response represents the failure to obtain an interview from a sample unit. Problems may arise in defining what the sample units really are; this topic will be discussed fully in the next section. In column 1 of table A1, response rates have been calculated on the basis of 'located' units, ie the denominator excludes vacant, destroyed and unlocated dwellings or households. The following five columns, which horizontally sum to 100 per cent, represent the ultimate disposition of all selected sample units. Depending on the nature of the sampling units and on the sample design, the 'correct' response rate for a given survey may be represented by the figure in column 1, by that in column 6, or by a figure somewhere between the two. For purposes of the present discussion we will restrict our attention to the figures in column 1. For the sake of precision, the term 'located' is used to refer to 'currently occupied dwellings' and '*de facto* households'.

If we exclude Portugal, the only European country, response rates for the household interview were over 85 per cent in all 39 countries for which we have data, above 90 per cent in 37, and above 95 per cent in 29. Refusal rates were generally well below one per cent, the only known exception being Jamaica. Hence the largest component of non-response was non-contact, either because of the temporary absence of household members or for some other reason (eg inaccessibility).

Response rates for the individual interview were also generally very high: above 90 per cent in all but two of the 40 countries for which data are available, and above 95 per cent in 25 countries. Refusal rates were once again uniformly low, and in no country did they constitute the bulk of non-response. However, it is interesting to note that refusals were much more significant in western countries than in African or eastern countries. Refusal rates of above one per cent were recorded in Portugal, Jamaica, Trinidad and Tobago, Colombia and Costa Rica. But, as in the case of the household interview, the primary reason for non-response in all countries was non-contact, defined broadly as the failure to secure an interview for reasons other than the attitude of the selected woman. These include physical absence, mental handicaps, and language problems. In most countries the first of these factors predominated in the incidence of non-contact.

The 'true' response rates for the surveys of individual women can be estimated as the product of the response rates for the household and individual interviews. However, this assumes that non-interviewed households contain, on average, the same number of eligible respondents as do interviewed households. Indirect evidence, based on the characteristics of households that required callbacks, suggests that non-contacted households have fewer members. This is corroborated by data from the WFS in Peru and Ecuador and from other surveys in various countries in which interviewers have been instructed to estimate household size of non-interviewed households by obtaining information from neighbours. Non-interviewed households have typically one to three fewer members than interviewed households. Hence, the product of the response rates for the two interviews will generally underestimate the true response rate. We have seen that in a large majority of WFS countries the response rates for both interviews were over 95 per cent. This implies true response rates of over 90 per cent for these countries. These are respectable figures by any standard.

For comparative purposes, it is interesting to cite some response rates of surveys conducted in association with the WFS in the developed countries of Europe and America. The comparative analysis of these surveys is being conducted by the UN Economic Commission for Europe, which has kindly provided some as yet unpublished data. Response rates for these countries vary from a high of 99 per cent in Bulgaria to a low of 59 per cent in Belgium. Few generalizations can be made about these response rates since the sampling methodology varied so greatly from country to country. In many countries no household interview was conducted; women were selected directly from population registers or census records, the completeness and timeliness of which varied greatly. In half a dozen countries the methodologically suspect practice of substitution was permitted, and for several of these it is not clear how the original non-contacts were handled in calculating the response rate. But one thing is clear: refusals, particularly in Western European countries, are of much greater significance in the total non-response than is true for WFS developing countries. Among women successfully contacted, refusal rates were around 20 per cent in Belgium and the Netherlands, and above 10 per cent in France, Great Britain, Norway and Spain. Even in the US, where a relatively good response rate of 88 per cent was obtained, one-half of the non-response was attributable to refusals.

2.2 ON THE COMPUTATION AND COMPARABILITY OF RESPONSE RATES

Allusion has been made to the difficulty of making meaningful comparisons of response rates if the sampling methodology differs. In the case of WFS developing-country surveys the problem has to do with the nature of the listing units on the one hand, and the timing of the listing operation on the other.

The listing units can vary along a continuum from purely social units to purely structural units. The dwelling unit was the favoured listing unit in the Latin American surveys. In several Asian countries, among them Bangladesh, Korea and the Philippines, the listing units were households. In other Asian and most African countries, the listing units were a mixture of dwellings and households: householder names were utilized to identify dwellings, and later all resident households in selected dwellings were interviewed, even if the original household had moved away since listing. This hybrid procedure may or may not be equivalent to a dwelling unit listing, depending on whether vacant dwellings are included in the listing or excluded from it. (Vacant dwellings can be included in a predominantly household listing either by using the name of the absent owner/usual occupier or by instructing listers to enter a description of the dwelling on the listing form.) If vacant dwellings are included, the procedure is no different from a dwelling unit listing. But regardless of the formal definition of the listing units utilized in the surveys' sampling manuals, in practice there was a tendency everywhere to favour structural units in urban areas and social units in rural areas. This arises from the anonymity of urban life, on the one hand, and the difficulties of preparing reliable sketch maps in rural areas, on the other.

The timing of the listing operation also has a bearing on response rates. In the case of pure household listings, it is essential that there be little time lag between the listing and interviewing operations. This will eliminate two potential sources of non-response bias: households that move away and households that are dissolved. It will also eliminate another source of potential bias: the exclusion of newly formed households in the intervening period. (This really is a coverage problem, but is dealt with here out of convenience.) Hence, the longer the gap between listing and interviewing, the greater the bias due to both non-response and selectivity. In samples of dwelling units, the timing of the operations is not so crucial, unless vacant dwellings have been given no chance of selection. If they have been excluded and if there is a significant delay, there will be a bias arising from the exclusion of newly formed households in dwellings that were vacant at the time of listing but occupied at the time of interview, and from the construction of new dwellings.

Correct calculation of response rate can be achieved only if one knows precisely the listing units utilized and has some idea of the time lag between listing and interviewing operations. In two of three WFS household samples, the interval between listing and interviewing was either extremely short (two to three days for Korea) or moderate (four months on average in Bangladesh) and thus the problems of bias are minor. The Philippines represents a particularly unusual and interesting case. Referring to table A1, it might be assumed that a response rate of 98.5 per cent was achieved in connection with the Philippines household survey. This is based on the assumption that the proper denominator was 'located households', based on the further assumption that the listing units were of the hybrid household-dwelling type. In fact, the listing units were private households. In the case of true household samples, such as this, the listing and interviewing units coincide, so the proper denominator for calculating response rates is all listed households. As already observed, in true household samples it is essential that there be virtually no gap between the listing and interviewing operations. A quick glance at column 6 of table A1 for the Philippines shows that only 86.4 per cent of the selected households were interviewed. What went wrong? In point of fact, a reference date of 1 August 1977 was used in listing households (during November); any households that had not been resident on that date were excluded from the listing. Interviewing took place over the period February to June 1978. Therefore, the average time lag between eligibility for the listing and the interviewing operation was nine months. It is little wonder that so many of the listed households could no longer be found.

This high non-response in the Philippines, resulting from the delay of nine months, has apparently introduced serious distortions not only in response and coverage rates but also in certain types of data, particularly nuptiality data. These arise because newly formed households and households that moved are excluded from the survey universe. It is revealing that while for the period August 1976 to January 1977, 185 first marriages were recorded as having occurred, for the similar period 12 months later (after the closing date of 1 August but before the start of interviewing) only 93 first marriages were recorded. This apparent shortfall of recent marriages undoubtedly reflects the exclusion of newly formed households from the sample universe but may in part also be due to yet a further biasing procedure utilized in the survey: women resident in selected households at time of interview who had not been resident on 1 August were enumerated on the household schedule but were exempted from the individual interview. A comparison of the age distributions of the excluded women (ever-married, ages 15–49) with that of the interviewed women is revealing (see table 1). There is evidently a dearth of recently married women in the sample selected for the individual interview. Might this not also affect the data on recent fertility? The example of the Philippines underlines the importance of avoiding any delay between the listing and interviewing operations in the case of true household samples.

Table 1 Comparison of age distribution of excluded and interviewed women

Age	Excluded (%)	Interviewed (%)
15–19	17.0	3.0
20–24	34.2	13.2
25–29	22.7	19.1
30–34	10.3	18.5
35–49	15.8	46.2
Total	100.0	100.0
N of cases	348	9268

In the case of hybrid household–dwelling samples, of crucial importance in the calculation (and hence, comparability) of response rates is whether or not vacant dwellings have been given a chance to be selected. If there is a significant time interval between the listing and interviewing operations, and if vacant dwellings have not been sampled, a bias will arise from the exclusion of newly formed households in dwellings that were vacant at the time of listing but occupied at the time of interviewing. Many WFS hybrid household–dwelling listings did not include vacant dwellings. In such cases, dwellings that were occupied at time of listing but vacant at the time of interview represent non-response. In those cases where vacant dwellings were sampled, any dwellings found to be vacant at time of interview are classed as blanks in the frame, and may be excluded from the denominator in calculating response rates.

There are some further practical difficulties in defining non-response in the hybrid household–dwelling or simple dwelling samples. Consider the following classification of outcomes of the Kenya household survey:

Selected household units	10 763
Dwelling vacant	590
Family away for duration of fieldwork	292
Non-African household	305
Household not located	277
Flooding, language problems	228
Household not at home	144
Refusal	33
Other	3
Successful household interviews	8 891

In Kenya, vacant dwellings were listed and given a chance of selection. Hence, the response rate as reported in the First Country Report and as presented in table A1 (92.8 per cent) is calculated over a base which excludes 'dwelling vacant', 'family away' and 'non-African household'. Non-African households were defined to be outside the sample domain so they raise no problem. Nor do vacant dwellings. The difficulty lies with the three categories 'family away', 'not located' and 'not at home'. As the Kenya survey used a *de facto* sample, the exclusion from the base of 'family away' is theoretically correct. Such families have a chance of selection as visitors in other sample clusters. It is also theoretically correct to include the 'not at home' category in the base for calculating response rates. These denote households that are in current residence but have not been successfully contacted because they are repeatedly not at home when the interviewer calls. In practice, however, the distinction between the two categories is subtle and difficult to enforce with consistency. Most WFS surveys did not even attempt the distinction but instead counted both as non-response. The third category 'not located' is also problematic because of its ambiguity. It may indicate poor fieldwork by listers or interviewers; the household may exist but has been inadequately described by listers or inadequately traced by interviewers. If this is so, then these cases count as non-response. Alternatively, 'not located' households may represent demolished dwellings or families that have migrated out without being replaced by a new household. In these circumstances they should not count as non-response. In the Kenya survey, the cautious decision was taken to classify 'not located' households as non-response. In other WFS surveys, they have been combined with the 'vacant, destroyed' category and thus typically excluded from the response-rate denominator.

There is one further item of interest in the Kenyan classification of outcomes, namely 'flooding, language problems'. In several WFS surveys, including that in Kenya, whole clusters or large portions of them have had to be abandoned for a miscellany of special reasons such as inaccessibility, flooding, political instability, inability to communicate, etc. In the Kenyan case, one entire cluster was omitted due to linguistic problems. One other cluster was abandoned when only half-complete, due to flooding. If such exclusions occur after the listing stage, it is simple (and correct) to include them under non-response. But what happens if they occur before listing? The Philippines is a case in point. Initially, 742 sample areas were selected. Of these, 13 were excluded before the listing operation for reasons of accessibility or security. Listing was successfully carried out in the remaining 729 areas. During the main fieldwork, 11 areas could not be covered due to security problems, and the interviews for another two areas were lost in transit; thus, in the end, interviews from 716 sample areas were processed. How were the excluded 26 areas treated? The initially excluded 13 (ie unlisted) were defined as being outside the survey universe. The remaining 13 (listed) areas were treated as non-response. This seems to have been the general practice in those few countries where entire areas could not be interviewed: if they had already been listed, they were counted as non-response; if they had not yet been listed, they were defined out of the survey universe. In Pakistan, seven areas could not be worked due to inaccessibility; the universe was redefined to include 91.8 per cent of the national population rather than the originally defined 93.2 per cent.

In the Latin American countries the sampling units were generally dwelling (housing) units rather than households. In those countries where an *ad hoc* sample was prepared for the WFS, the listing operation was usually carried out two to three months prior to interviewing. The listing unit in these surveys was the dwelling unit, whether occupied or unoccupied. But in a number of countries the WFS benefited from the existence of a master sampling frame. In the case of these master samples, intended for a relatively long lifespan, the listed unit is normally any structure, or part thereof, which could potentially someday be destined for residential use. In table A3, Costa Rica, Pakistan and Ecuador are examples of *ad hoc* WFS dwelling unit samples, while Peru exemplifies a master dwelling sample. There is a striking difference in the proportion of units classified as 'address not a dwelling' in Peru as compared to the other three, though this, of course, should not be counted as non-response.

There is no easy solution to the problem of achieving comparability among sets of WFS household response rates. This arises chiefly because of a lack of documentation concerning such matters as the nature of the listing units, whether or not vacant dwellings were sampled, and the criteria used by interviewers to distinguish between vacancies and 'not-at-homes'. Perhaps the best that can be done is to present upper and lower limits, as is done in tables A1 and A3. The 'correct' response rates in the case of dwelling-unit samples are probably somewhat lower than those based

on the number of ODUs, but significantly higher than those based on all dwellings. In the case of true household samples, they are necessarily closer to those based on selected households than to those based on located households. And in the case of hybrid household/dwelling samples, they are probably closer to those based on located households than to those based on selected households, if vacant households have been sampled, as illustrated by Kenya. If vacant households have not been sampled, they are probably closer to those based on selected households.

Response rates to the individual interview tend to be more comparable across countries, but here too there is also one complicating factor, namely, whether a *de facto* or a *de jure* residence criterion has been used to define eligibility. A *de jure* criterion should generally produce higher non-response due to the absence of usual residents. However, the use of a *de facto* criterion can also create some non-response in that visitors who slept at the residence the night before the household interview become eligible for the individual interview. By the time of the interviewer's visit for the individual interview they might have moved on. On balance, it has been found across many countries that the number of non-residents detected and successfully interviewed under a *de facto* criterion is approximately one-half that of the number of usual residents who are reported away at any point in time (Verma 1981). Hence, a *de jure* criterion tends to produce more non-response due to not-at-homes.

Peru provides a laboratory experiment for measuring non-response under the two criteria. The official criterion used there was *de facto*. However, early on in the field work it was noted that some normally resident women were being excluded simply because they had not slept at their usual residence the night before the interviewer's visit. To protect sample size projections, a decision was made to interview such women. Hence, in practice, Peru used a theoretically incorrect double criterion: a woman became eligible if either she had slept there or usually lived there. The resulting response rates were as follows.

Live	Slept	Response rate (%)	Base
1 Yes	Yes	97.4	5531
2 Yes	No	19.9	336
3 No	Yes	95.4	195
De jure (1 + 2)		93.0	5867
De facto (1 + 3)		97.3	5726
Combined (1 + 2 + 3)		93.0	6062

What happened was that the inclusion of 336 non-sleeping usual residents, ie absentees, produced but 67 interviews, a very low response rate of only 20 per cent. In any event, this illustrates the probable relative magnitude of non-response under the two residence criteria.

In table A2 the residence criterion used in each survey is identified. There are many gaps in the figures, particularly for Africa and the Middle East. Perhaps because of these gaps it is difficult to observe any tendency for *de jure* countries to experience higher rates of absence than *de facto* countries, as might be expected.

2.3 DIFFERENTIAL NON-RESPONSE TO THE INDIVIDUAL INTERVIEW

As the bulk of non-response is encountered at the household interview rather than the individual interview, it would be both important and desirable to compare the characteristics of non-respondent households with those of respondent households. Unfortunately this is not possible because of lack of data on non-respondent households. However, it is possible to examine the characteristics of non-respondents to the individual survey in successfully interviewed households to see if non-response at this stage is selective of certain types of women. If it is, and if it has been of significant incidence, then the survey results will certainly contain biases. In WFS surveys, varying amounts of social and demographic information have been collected concerning household members on the household schedule. This information enables one to compare the characteristics of women who later responded to the individual interview with those of women who were non-respondents; or, what amounts to the same thing, to calculate rates of non-response by characteristic. Data concerning differential non-response in selected countries are presented in table A4.

Non-response appears to generally be selective of older women, especially those over the age of 45. Detailed cross-tabulations (not shown) reveal that it is also selective of very young women, particularly if they are single or childless. Non-response rates are lowest among women of ages 25–34, presumably because they are tied to the home by the presence of young children. Non-response is selective of women who are not in a marital union, be they single, separated, widowed or divorced. It is also selective of women other than the household head or wife of the household head. Such women are mostly daughters of the head.

Surprisingly, non-response appears not to be related to the educational achievement of women in any consistent manner. Women at both the very lowest and very highest levels may show slightly higher rates of non-response than others, but differences are small. Even more surprisingly, non-response shows no consistent pattern by area of residence. This is in contrast to developed countries, where refusal rates soar in major urban areas. Perhaps if refusals become a significant component of non-response in the future, this pattern will be followed in developing countries. It is significant that even now such a pattern does appear in Costa Rica and Fiji.

While fertility data were collected on the household schedule in some countries, particularly in the more recent African surveys utilizing an expanded household sample, proper matching of the household and individual data has yet to be undertaken. Thus we are unfortunately unable to compare the fertility of respondents and non-respondents to the individual interview.

To sum up, non-response is selective of the oldest women; of the youngest if they are single or childless; and of women not currently married. This having been said, its impact on survey results is certainly trivial. With response rates to the individual interview averaging around 95 per cent, a difference of five per cent in non-response rates between subgroups in the population will affect, or bias, the resulting distributions by only one-quarter of one per cent. This bias is much less than the concomitant sampling

errors for almost all estimates from samples of the size used in the WFS, typically 3000–10 000 women.

2.4 THE EFFECT OF CALLBACKS ON RESPONSE RATES

A callback is defined as a later visit to a household or a person to obtain an interview where the interviewer was unsuccessful on her first or prior visit. According to standard household survey procedures, a callback should not be recorded if the interviewer simply knocks on the door again an hour later after an intervening interview. A callback, to be counted as such, should be made at a distinctly different time of day, or on a different day. In all probability, widely different standards were applied in the field in the WFS participating countries in the definition and recording of callbacks. It is a fact that in some countries interviewers were instructed not to record an unsuccessful callback if it occurred on the same day as the prior visit. This can be attributed to a perhaps unfortunate recommendation made in the WFS model Supervisors' Instructions (*WFS Basic Documentation* no 5, p 33): 'It is recommended that when a team is staying within a sample area, all attempts made by an interviewer to contact a respondent in the course of a single day should be counted as a single visit or callback.' The purpose of the recommendation was to prevent interviewers from using up their quota of calls to a household (normally four) in a single day. Fortunately for the present analysis, this practice was not followed in all WFS surveys. Where it was followed, often as many as 98 per cent of all household interviews are recorded as having been obtained on a first visit. For such countries no meaningful analysis of callback response can be made.

Another problem is the distinction of calls made for the purpose of obtaining a household interview and calls made to achieve an individual interview. Theoretically, calls should not be recorded on the individual interview until after successful completion of the household interview. But, in fact, for several countries one can detect a positive correlation between the number of calls at the household and individual levels, or more specifically, a concentration of cases on the diagonal. This would suggest that some calls recorded on the individual questionnaire were actually calls directed to obtaining a household interview. But despite the inadequacies of the callback data for a number of countries, it is still useful to analyse those countries with apparently good data to evaluate the benefits and costs of callbacks.

The first question to be asked is what effect callbacks have in boosting ultimate response rates. Data presented in table A5 suggest that the effect can be substantial. In Ghana, for example, only 76 per cent of selected households were successfully interviewed on the first visit. Subsequent visits resulted in raising the final response rate to close to 89 per cent. Similar benefits were recorded in Ecuador and Peru, with somewhat lesser benefits in the case of Costa Rica. Interviewers are, of course, selective in their callbacks. After a first visit they discard all addresses that are not a dwelling, all obviously vacant dwellings, and all non-existent households. Hence, another measure of the utility of callbacks is their 'payoff', measured by the percentage of units eventually successfully interviewed among those to which a second visit was made. These figures, presented in the final column of table A5, are dramatic evidence of the usefulness of callbacks. Roughly three-quarters of initially non-interviewed units were eventually interviewed among those sent back to the field in these countries. Callbacks obtain results!

3 Callbacks

3.1 CALLBACKS AND HOUSEHOLD SAMPLE COMPOSITION

It has been seen how callbacks help to improve ultimate response rates. A related question of some importance is what effect they have on the composition of the sample. If respondents picked up on second or subsequent visits are markedly different from those interviewed on the first visit, callbacks can be said to have an added payoff. Once again, of course, if almost all interviews are obtained on a first call, any differences between first-call respondents and callback respondents would have a negligible impact on sample composition. In developed countries the question is not trivial. In the Norwegian Fertility Survey of 1977, for instance, only 40 per cent of the ultimate interviews were obtained on a first visit; an additional 35 per cent were obtained on a second visit. Up to eight visits were made by interviewers. The mean number of live births per woman declined monotonically with call number: 1.84 on the first call, 1.56 on the second call, 1.08 on the sixth call, and 0.75 on the eighth call (Thomsen and Siring 1980).

Differences in household composition according to whether the interview was achieved on a first call or not, further broken down by urban/rural residence, are presented in table A6. It would have been interesting to break down further the call numbers, but the number of cases simply will not permit such analysis. The percentage of interviews obtained on a callback in urban and rural areas, respectively, of the four countries are as follows: Peru, 20 and 5 per cent; Costa Rica, 9 and 3 per cent; Pakistan, 6 and 3 per cent; and the Philippines, 15 and 9 per cent.

The composition of households by age, marital status, educational attainment and number of members is shown in the left-hand portion of table A6. The specific characteristics presented (eg percentage under age 10) have been selected for being the only ones that show some differentiation. First call response is selective of households containing children under the age of 10 years in urban areas, but not in rural areas. Callback response appears selective of households containing widowed, divorced or separated persons, except in the Philippines. In all four countries callback response is selective of households containing better-educated members; the differential is particularly strong in urban areas. In Costa Rican urban areas, for example, only 14 per cent of members of first-call households had some post-secondary education, as compared to 24 per cent of members of households interviewed on a later call. In general, and as might be expected, households interviewed on a first call are larger in size (number of members) than those to which a callback was required. The exception of the Philippines is indeed strange.

Looking at only that part of the household that is of interest in a fertility survey, ie women in the childbearing ages of 15 to 49 years, callbacks for the household interview do not appear to be selective consistently by either age or marital status. Interesting differences which appear in one country are reversed in another. But once again, callbacks select out households containing better educated women, particularly in urban areas. Education is hence the only variable on which callback response is consistently selective across the four countries.

3.2 CALLBACKS AND INDIVIDUAL SAMPLE COMPOSITION

As pointed out earlier, callbacks may also be required to obtain an individual interview after successful completion of the household interview. So once again it is of interest to examine the selectivity of callback reponse, this time in connection with obtaining the individual interview. The relevant figures for this discussion are presented in table A7. A much wider selection of countries is available for examination since for this analysis alone it is possible to use the Standard Recode Files developed by the WFS.

In several countries, callback response is selective of older women (ie over the age of 35), notably in Peru, Pakistan, Thailand and Malaysia. In the other countries it is less selective by age. There appears to be no generalization that can be made about marital status. In Peru, callback response is selective of women not currently in union, while in Haiti the reverse is true. In the other countries callbacks appear to be unrelated to marital status. All in all, callbacks do not appear to be related to these particular demographic characteristics of the women concerned.

The same is not true of social characteristics. In all eight countries callback response is selective of the best educated women, particularly in urban areas. The selectivity is especially pronounced in Peru, Costa Rica, Pakistan, the Philippines and Thailand. It is somewhat attenuated in the cases of Haiti, Ghana and Malaysia. Callbacks are also highly related to the labour force status of the women. They are selective of women who currently work away from the home, in both urban and rural settings in all eight countries. Only in Ghana is this selectivity not quite so pronounced.

The number of births that women have had in the past five years is a variable serving as a proxy for both current fertility and for child-care burden. Given the usual relationship between this variable and women's labour force status, one would expect callback response to be selective of women who have had fewer births in the past five years. This is indeed the case. In Peru, for example, urban women interviewed on a first call have had 26 per cent more births than those interviewed on a callback; the comparable figure for Thailand is 38 per cent. The differences are less pronounced in other countries, and sometimes disappear altogether for rural areas. Nevertheless, it is of great interest that callbacks are related to this most crucial of variables in a fertility survey. The similarity of findings from these

Table 2 Births to women in urban areas in five years before interview

Urban areas	Percentage callbacks	Births in last five years		
		First call result	Overall result	Difference
Pakistan	3.8	1.17	1.16	0.01
Costa Rica	5.7	0.69	0.69	0.00
Ghana	7.7	1.03	1.03	0.00
Peru	13.1	1.11	1.08	0.03
Haiti	18.6	0.81	0.80	0.01
Philippines	19.4	1.05	1.03	0.02
Thailand	24.4	0.83	0.77	0.06
Malaysia	35.0	0.99	0.93	0.06

developing countries to those cited above from Norway suggests that this is a universal phenomenon.

What effect do callbacks have on the composition of the individual sample and thus on the substantive results? In the case of differential non-response to the individual interview, it was found that the effects on sample distributions were negligible, in large part because of the generally very high response rates. In the case of callbacks we can compare the results of first call interviews with overall results to measure the impact of callbacks. Taking the variable most highly related to callbacks, namely, the percentage of women in urban areas who work away from home, the first call estimate for Pakistan is 4.5 per cent while the overall result is 5.0 per cent. For Peru the figures are 19.5 vs 22.3 per cent, and for Thailand, 29.0 vs 37.6 per cent. The pattern is evident: callbacks make a difference in precisely those countries where they have most often been required to obtain successful interviews. This is seen clearly in table 2 which shows the mean number of births in the five years before interview to women in the urban areas. The countries are ordered by the percentage of callbacks required in these same urban areas. While the correlation is far from perfect, the pattern is once again clear: where callbacks are frequent, they make a significant impact on findings because women picked up on callbacks are somewhat different from those successfully interviewed at a first visit.

It has been seen that women interviewed on a callback have some characteristics in common with non-respondents to the individual interview. In general, both non-response and callback response are selective of women not currently in a marital union and of women who are highly educated. In the case of callback response we concluded that much of the selectivity was attributable to the phenomenon of women working away from home. Can a similar assumption be made about non-respondents? We have data from only one country in which information on labour force status was collected on the household schedule, namely Ecuador. In table 3 women are classified according to whether or not an individual interview was achieved, and if so, whether it was on a first visit or a second or later visit. The figures are revealing. The lowest non-response and yet the highest callback rate occurs among women who work in the modern sector. In contrast, housewives are much more likely to be interviewed on a first call, but in the end are marginally more likely to be non-respondents compared with women working in the modern sector. Clearly, in the case of

Table 3 Women interviewed or not on first or later visit

	N of cases	All (%)	Interviewed		Not interviewed
			First call	Later call	
All	7241	100.0	85.8	9.5	4.7
Works	2397	100.0	82.4	13.4	4.2
Does not work	4844	100.0	87.5	7.5	5.0
Housewife	3613	100.0	89.3	6.4	4.3
Student	982	100.0	84.0	12.2	3.8
Works, modern sector	1414	100.0	81.0	15.6	3.4
Works, traditional sector	983	100.0	84.3	10.3	5.4
Other	249	100.0	73.9	6.4	19.7

Modern sector: Professional, clerical, sales, skilled and unskilled.
Traditional sector: Farmers, farm labourers, domestic and other service workers.
Other: Looking for work, retired, incapacitated, etc.

Ecuador it would have been wrong to assume that non-respondents are similar to callback respondents as regards their labour force status.

3.3 SOME EVIDENCE ON THE COST OF CALLBACKS

Callbacks have their costs as well as their benefits. They obviously slow down the working rate of an interviewing team, and may even require that the team stay on an extra day in a given locality. However, a team is not necessarily defenceless: it can reorganize its working schedule to respond to the need for callbacks. If callbacks are distributed uniformly over the period of stay in a cluster, and are not allowed to pile up to the end, an extra day may not be required. Unfortunately, we have no solid data on the real monetary cost of callbacks. To obtain such data would require elaborate pre-planned accounting procedures, something not attempted in WFS surveys. Some indirect evidence on the topic is presented in table A8. The data are for clusters outside of the capital city of each country, on the presumption that an 'extra day' has little meaning in the home city of interviewers. (In the case of Peru, the data refer to rural sierra districts.)

On the left side of the table is shown the percentage distribution of interviews obtained by serial day in the cluster, for clusters grouped according to the number of productive days spent in them by the interviewing team, separately for Peru, Costa Rica, Ghana, the Philippines and Thailand. The meaning of 'cluster' varies somewhat among these countries. In the Philippines and Thailand they are very small, averaging around 15 interviews. The clusters in the other three countries are of intermediate size, 25–50 interviews. (In the case of Costa Rica they are actually groups of clusters interviewed in single outings from San José.) Despite this variability, a common pattern emerges:

productivity falls on the final working day in a cluster, and on the final two days in five- and six-day clusters. This effect is particularly pronounced for Thailand, where regardless of how short or how long the stay in the cluster, productivity on the final day never reached 50 per cent of the average productivity per cluster-day.

Perhaps part of the explanation of this phenomenon lies in callbacks. (A host of other factors may also be involved: packing up, planning for the next area, awaiting transport, etc.) In the right-hand portion of table A8 data are presented on the proportion of interviews achieved after a callback by serial day in the cluster. As might be expected, callback interviews as a proportion of all interviews rise dramatically on the final day in a cluster, and on the final two days in five- and six-day clusters. Once again, the effect is particularly marked for Thailand, where well over half the interviews on a final cluster-day in three- and four-day clusters are achieved after a callback. Looking at it another way, proportionately three times more interviews are callback interviews on the final day than on other days in Thailand. This proportion is not so extreme in other countries, but still is of some magnitude. Only in the Philippines is there evidence that interviewing teams tried to spread out the volume of callbacks somewhat uniformly over the period of stay in a cluster.

The circumstantial evidence suggests that in these countries callbacks have played a significant role in extending the stay of interviewers in some clusters. This is very clear in the case of three- and four-day clusters in Thailand, where under 15 per cent of completed interviews were taken on a final cluster-day, and over half of these were callback interviews. The impact of callbacks on interviewer productivity appears to be somewhat less in the other countries examined. However, it is significant that in the case of the Philippines, where the volume of callbacks was the greatest, the percentage of interviews achieved after a callback rises monotonically with the number of days spent in clusters. The conclusion is that callbacks undoubtedly have their costs as well as benefits.

4 Coverage

4.1 ACHIEVED VS PLANNED SAMPLE SIZES AND COVERAGE ERRORS

In developed countries' surveys, probably the most important component of non-sampling error is non-response bias. In the developing world, on the other hand, non-response is still not much of a problem, as we have seen. Here, rather, the problems are of a different nature: response errors and coverage errors. The problem of response errors, or response reliability, is the focus of a special investigation being undertaken by the WFS. Results will be coming out soon. (For some preliminary results, see O'Muircheartaigh and Marckwardt 1981.) The topic to be dealt with here is that of coverage errors.

An obvious, though superficial, indication of possible coverage problems in WFS surveys comes from a simple comparison of achieved sample sizes of the surveys with the number expected, as specified in the original agreement between the country concerned and the funding agency, the so-called 'project document'. These figures are presented in table A9. Of the 38 countries for which an explicit expected number of individual interviews was specified in the project document, shortfalls were experienced in 30. They are particularly large in the case of some African countries, and cannot be explained away by non-response. The magnitude of the shortfalls indicated in table A9 should, however, not be taken too literally. Often a very approximate figure was chosen for the project document, and this was later altered in the light of certain logistic or financial considerations.

An error in coverage occurs when some unit in the population which should have been included is omitted from the sample, or, more uncommonly, when some unit which has not been selected in the sample is erroneously included. Under-coverage is not to be confused with non-response, which has to do with omission from the survey of a unit selected in the sample, rather than omission from the sample. (For an excellent discussion, see Chapter III, United Nations 1982.) There is scant evidence from the WFS surveys of over-coverage, but abundant evidence of under-coverage. Yet though the evidence is there, coverage errors are very difficult to measure precisely. Unlike response rates, which can be measured from the results of the survey itself, the measurement of coverage errors involves external evidence. The evidence can come from additional, independent fieldwork. In practice, this has not occurred more than a couple of times in the WFS programme. Most commonly the evidence has arisen from application of analytical techniques: extrapolation and comparison of survey results with census data; or an examination of the face validity of an age distribution.

Coverage errors can occur at the two listing stages: households or dwellings, and household members. At the household or dwelling stage the most common problems encountered concern the boundaries of the sample area, on the one hand, and the completeness of the listing within the area, on the other. Careless mapping or faulty instructions are the cause of most boundary errors; ill-trained or lazy listers are the principal cause of under-coverage within accurately delimited sample areas.

As for the listing of household members, some errors may arise from genuine problems in defining the household in *de jure* samples. But this has not been a major problem in WFS surveys. The principal 'error' has been the biased reporting of age and/or marital status information on women in the household, making them appear ineligible for the individual interview when in fact they are eligible. In some surveys there is evidence of omission of women from household member listings. The effect is the same: a shortfall in eligible respondents. Evidence of this phenomenon will be presented below. Whether such systematic age misplacement is the fault of the respondent or that of the interviewer has not been investigated; an educated guess is that it is generally due to a deliberate attempt by the interviewer to lighten her workload.

In table A10 data have been assembled on the components of the shortfall in individual interviews in nine countries. The figures may differ somewhat from those in table A9, because here we are concerned with either the explicit or implicit assumptions governing the sample design. To better understand the nature of the breakdown presented in table A10, we illustrate the calculations involved with reference to one of the countries, namely Costa Rica. The exposition of sampling assumptions, expectations and results in that country's report is unusually explicit.

Costa Rica sample (aim: to achieve 4000 individual interviews)

ASSUMPTIONS AND EXPECTATIONS	
Dwelling units in sample universe, 1973 census	342 155
Annual rate of growth of population (%)	2.2
Dwelling units at time of survey (1976) (assuming population growth rate)	365 238
Desired sample dwelling units	4 870
Sampling fraction (4870/365 238 = 1/75)	0.0133
Occupancy rate of dwelling units, 1973 census	0.937
Occupied dwelling units (4870*0.937)	4 563
Response rate to household interview (%)	96.0
Number of household interviews	4 380
Eligible women per household, 1973 census	0.9563
Number of eligible women	4 189
Response rate to individual interview (%)	96.0
Number of individual interviews	4 021
ACHIEVEMENTS	
Number of addresses	4 724
Address not a dwelling (%)	2.6
Number of dwellings	4 601

Occupancy rate of dwelling units	0.938
Number of occupied dwelling units	4 317
Response rate to household interview (%)	98.3
Number of household interviews	4 244
Eligible women per household	0.9590
Number of eligible women	4 070
Response rate to individual interview (%)	96.7
Number of individual interviews	3 935

REASONS FOR SHORTFALL

		as % of expected
Actual take less expected take (3935–4021)	− 86	− 2.1
Under-coverage/over-projection of DUs [{(4870–4601)}*0.937*0.96*0.9563*0.96]*–1	− 222	− 5.5
Under-projection of occupancy rate [{(4601*0.937)–4317}*0.96*0.9563*0.96]*–1	+ 5	+ 0.1
Under-projection of household response rate [{(4317*0.96)–4244}*0.9563*0.96]*–1	+ 92	+ 2.3
Under-projection of women per household [{(4244*0.9563)–4070}*0.96]*–1	+ 11	+ 0.3
Under-projection of women's response rate [4070*0.96–3935]*–1	+ 28	+ 0.7

The mechanics of predicting a sample take involve the following elements: the number of dwellings (households) and the number of eligible women per dwelling (or, alternatively, simply the universe of eligible women) from a prior census or survey; an assumed rate of growth since that time; if the estimates are based on dwellings, an assumed occupancy rate; and finally, assumed response rates to both the household and individual interviews. The sampler then works backward from the desired sample take to establish the overall sampling fraction. In the end, if the assumptions utilized in establishing the sampling fraction have been realistic, and if there have been no serious deficiencies in the execution of fieldwork (mapping, listing and interviewing), the number of women actually interviewed should nearly coincide with the desired sample take.

In the specific case of Costa Rica, it will be noted that satisfactory projections and allowances were made for the occupancy rate, women per household, and non-response rates. The only source of the shortfall in individual interviews came from either an under-coverage of dwelling units or an over-projection of the same. It was perhaps an error to have applied the last inter-censal annual population growth rate (1963–73, 2.2 per cent) in projecting the growth of dwelling units for the subsequent period (1973–76), in the face of a dramatic and sustained fall in the birth-rate that commenced in the late sixties. However, the implied annual growth rate of 0.3 per cent in dwelling units, calculated from the sample outcome, is unacceptably low, indicating that there was some under-coverage of dwelling units.

Now, returning to table A10, the first component of shortfall, titled 'Number of households' (the fifth row of figures), stems from either the under-coverage of households (dwellings) or an over-projection of households based on a prior census. Although in some countries there might have been an over-projection made at the time of designing the sample, in most cases the major part of the deficit is clearly attributable to under-coverage.

The second component of the shortfall in individual interviews is titled 'Household non-response'. As can be seen in table A10, in eight of the nine countries this component is a positive quantity. This arises because in designing a survey we ordinarily make an explicit allowance for non-response. Clearly, response rates were generally higher than had been anticipated. (In some countries an allowance for under-coverage was built into the study design; in these cases, such allowance has been included in the second component on the grounds that it makes no sense to measure coverage failures from an imperfect standard.)

The third component is titled 'Number of women'. This measures the shortfall in the number of eligible women appearing in the sample. In most surveys some assumption was made about the number of eligible women per household (dwelling) at the time of designing the sample, and this entered into the fixing of the household sample size. In this type of design, any real change in household composition from the assumed conditions will, of course, affect sample size. But in some other surveys an estimate was made of the total universe of eligible women, and the sampling fraction was fixed accordingly. Such designs are less vulnerable to changes in household composition. But any design is vulnerable to age mis-declaration. In seven of the nine countries there was a substantial shortfall relative to expectations in the number of eligible women.

The fourth component of the shortfall in individual interviews is simply the non-response of eligible women. This, in practice, is the least significant of the four components. It is a positive component in those cases where an overly generous allowance for non-response was built into the sample design. It is a negative quantity where no allowance or an explicit or implicit under-allowance for non-response was made.

In the present context our interest is focused on the first and third components of the shortfall: coverage of households and coverage of women. A few case studies of each should prove illuminating.

In Kenya the relative shortfall in the number of households was not a particularly grave problem, especially when compared to the shortfall in women. Nevertheless, it worried the survey organizers because the WFS exercise was just one of a series of surveys in the National Integrated Sample Survey Programme, and they aspired to ever-better results. They organized one of the very few post-survey field-checking operations mounted in connection with WFS surveys. The results of this operation confirmed that '... faulty sample implementation in the field leading to wrong identification of some cluster boundaries, as well as incomplete household listing and structure numbering are the major causes of under-coverage' (Kenya Fertility Survey, First Report: 31).

In Pakistan, the largest component of the shortfall in individual interviews is attributable to the deficit in the number of households. This is a typical case in which it proves to be impossible to disentangle under-coverage from over-projection. Growth from the time of the 1972 census

to mid-1975 was assumed to be 15 per cent in urban areas and 10 per cent in rural areas. Such large allowances for 'growth' were made because many demographers suspected a heavy under-enumeration in the 1972 census. The actual shortfall of household units in the sample amounts to 1223, or nearly 20 per cent of the expected 6152 units. Even had the 1972 census been complete and had there been no growth since that time, the sample of households would still be short by some seven per cent. Hence, there were clearly some serious errors in coverage. The true magnitude of the errors escapes measurement, but it is not unreasonable to place the lower limits of under-coverage at 10–15 per cent.

There has undoubtedly been some under-coverage of households in nearly every WFS survey, even those where, in the end, the targeted number of individual interviews was achieved. Such a case is Paraguay. The target set in the sample design was 4514 individual interviews; 4622 were actually taken. In the 1972 census, 428,111 private occupied dwelling units were counted. After taking into account a six per cent exclusion from the survey universe (El Chaco) and a presumed 2.55 per cent annual growth rate in the intervening six years, a sampling fraction of 0.0094 was set with the expectation of obtaining 4400 private occupied dwelling units in the sample. In fact, only 4218 were encountered, a 4.1 per cent shortfall. But since allowances for non-response of 10 per cent at the household level and of five per cent at the individual level were built into the design, and since actual non-response amounted to only 4.5 per cent and 4.4 per cent respectively, the targeted number of individual interviews was achieved. It should be noted that in both the census and the fertility sample, exactly 1.20 eligible women were encountered per household, a laudable feat in quality control by the Paraguayan field supervisors.

Is it fair to characterize the 4.1 per cent shortfall in households in Paraguay as under-coverage? Might not the 2.55 per cent annual projected growth rate have been an over-estimate? This will not be known until the next census is published. The results of the survey field work would imply an annual growth rate of only 1.83 per cent, which seems on the low side for Paraguay. So there probably was some under-coverage. And this would be above and beyond any under-coverage in the 1972 census.

Returning to the topic of under-coverage of eligible women, and to table A10, perhaps one of the most startling figures is the deficit of eligible women in Kenya. The sample was designed with the expectation of encountering 1.24 eligible women per household, a figure obtained from the 1969 census. It was unfortunate that at the time of designing the Fertility Survey (KFS), results from the 1977 National Demographic Survey (NDS) were not yet available. This survey revealed that the number of eligible women per household had dropped to 1.09. Using this figure as a reference, it can be calculated that of the total interview deficit of 2466 attributable to a lack of eligible women, 1278 or 52 per cent would be due to an apparent real decline in household size. However, the Fertility Survey recorded only 0.95 women per household. It is believed that the difference between the NDS and the KFS results is due to both omission of women and to age misplacement of women in the latter survey. From a careful analysis of the age distribution it is estimated that of the remaining shortfall of 1188 interviews, roughly 40 per cent is due to the omission of women, and the other 60 per cent due to age misplacement. There is evidence of a deficit of women of ages 15–24 and 45–49, and a corresponding surplus of girls of ages 10–14 and of women 50–59. This evidence is based upon a comparison of the NDS and KFS age distributions, contained in the principal KFS report.

Table 4 Age distribution at upper age limit in six countries: figures from the household survey

Country	% of females aged		Sex ratio[a] at ages	
	45–49	50–54	45–49	50–54
Fiji	2.9	3.7	139	86
Indonesia	3.8	4.6	116	78
Nepal	3.6	4.1	100	95
Turkey	4.1	4.3	115	86
Jamaica	1.8	2.5	103	67
Panama	3.0	4.8	113	68

[a]Number of males per 100 females.

There is evidence of age shifting across boundaries that define eligibility in quite a few of the WFS surveys. This occurs at both the lower and upper age limits, but is more pronounced for the latter. This phenomenon is not restricted to countries where women ordinarily have difficulty in remembering dates and ages. Some of the more extreme examples, where there are actually more women classed 50–54 than 45–49, are presented in table 4. The fact that such heaping at ages 50–54 occurred for women and not for men suggests that the phenomenon was not unrelated to the fact that women were thereby classified as ineligible for the individual interview.

Surveys involving a large two-phase operation provide an ideal laboratory for studying the phenomenon of under-coverage of women. In Lesotho the first phase consisted of a large (almost 20 000) household survey. A self-weighted sample of these households was later drawn for the Fertility Survey, the members were re-listed, and eligible women were interviewed. Comparison of the Phase 1 and Phase 2 household member listings is illuminating (see Timaeus and Balasubramanian, forthcoming). Phase 1 enumerated 0.83 eligible women per household, a fall from the figure of 0.90 of the 1966 census which had been used in designing the sample. Phase 2 enumerated only 0.75. Phase 1 registered a figure of 7.6 per cent visitors; Phase 2 only 2.7 per cent. Of all *de facto* females, Phase 1 showed 44.6 per cent to be in the eligible age range; the Phase 2 figure was 41.6 per cent. It is possible to allocate the deficit of individual interviews attributable to a shortfall in eligible women as follows:

Total deficit	714
Changed household composition since census	342
Under-enumeration of visitors	199
Probable biased reporting of age or marital status	173

A comparison of the Phase 1 and Phase 2 age distributions indicates that in the latter, women actually of ages 40–49 were shifted into ages 50–59 by interviewers.

To suggest that such age misplacement was a uniform characteristic of WFS surveys would be misleading. The case of Paraguay has already been mentioned. In Costa Rica, the 1973 census counted 0.9563 eligible women per

dwelling. The 1976 Fertility Survey counted 0.9590. But not only good examples emerge from Latin America. In Panama, the 1970 census enumerated 0.855 eligible women per dwelling; the 1975 Fertility Survey counted only 0.795. While this is not the place for invidious comparisons, it does seem that adequate coverage of women was achieved in those surveys characterized by recruitment of high-level supervisory personnel, good training of field workers and, particularly, close supervision of interviewers. But it should also be kept in mind that countries where age often has to be estimated by interviewers are especially prone to the phenomenon of age shifting across the boundaries that define eligibility.

5 Conclusions

5.1 SUMMARY OF FINDINGS

Response rates in WFS developing country surveys have been good. If these rates are defined in terms of located units, response rates for both the household surveys and the individual surveys were generally over 95 per cent. This implies true response rates (the product of the two) of over 90 per cent in the majority of countries. There are, nevertheless, some conceptual complications in comparing response rates across countries using the simple criterion of located units. These arise because the nature of the listing unit (household vs dwelling) varies across countries, as does the relative timing of the listing and interviewing operations. The implications are discussed in the text. The conclusion is that it is best to publish response rates based on both selected and located units. Normally, the correct response rate for a true household sample is closer to the figure based on selected units, while the correct response rate for a dwelling sample is closer to the figure based on located units.

Characteristics of non-respondents to the individual interview may be examined from data collected on household members in the household interview. In general, non-response to the interview is highest among women not currently in a marital union. Non-response is selective of the oldest women, and of the youngest if they are single or childless. There is a suggestion of a U-shaped relationship with educational attainment, with non-response being highest at the lower and upper extremes. Non-response differentials, while of some magnitude, are not large enough to bias survey results seriously when the response rates achieved have generally been well over 90 per cent.

The practice of 'calling back' to obtain an interview where the interviewer was unsuccessful in her first (or prior) visit has a definite impact on response rates. Though the data are very scanty, figures for four countries reveal that roughly three-quarters of initially non-interviewed units (both households and women) were eventually successfully interviewed among those to whom a later visit was made. Callback response is selective of smaller households (ie those having fewer members), and of households whose members are well educated. For the individual interview, callback response is selective of better-educated women and of those who work away from home. As a result, it is also selective of women of low recent fertility. We conclude that where callbacks have most often been required to obtain successful interviews, they make a significant impact on survey findings because women picked up on callbacks are different from those successfully interviewed on a first visit.

In a number of WFS countries, though by no means in all, the surveys have fallen short of their targeted number of interviews. In some cases this was due to deficient information upon which to base a sampling fraction; in others it was perhaps due to overly generous estimates of population growth at the time of designing the sample. But it is clear that in most cases the shortfalls are due to errors in coverage. In many countries the shortfalls due to inadequate coverage completely overshadow the shortfalls due to non-response in their magnitude and as sources of bias. These shortfalls have occurred at the two listing stages: households and household members. The problems at the household listing stage most frequently concern the boundaries of the sample area, on the one hand, and the completeness of the listing within the area, on the other. The quality of basic cartographic material available in a country inevitably plays a large role in this. The coverage of women for the individual survey has often been incomplete because of biased recording of age or marital status in the household member listing. The best documented bias here is the overstatement of age for or by women in their forties: many are recorded as being in their fifties and thereby become ineligible for the individual interview. On the bright side, in a number of WFS country surveys there is no evidence of such biased reporting; this has occurred in those countries where women are conscious of their age and where supervisors have exercised close control over interviewers.

5.2 RECOMMENDATIONS FOR FUTURE SURVEYS

Production of this analysis, while in the end a rewarding experience, was at many stages a most frustrating one. This arose because of lack of data, on the one hand, and faulty data documentation, on the other. Concerning the lack of data, what apparently has happened in many countries is that in the process of cleaning the data, cases of non-interview have been deleted from the raw data files. Yet the dictionaries for these files contain the full range of codes for outcome of the attempted interview (completed, not at home, refused, etc). This experience with faulty documentation, added to the results of the foregoing analyses, leads the author to make a set of suggestions for future data collection projects of the WFS type.

(1) More emphasis should be placed on the documentation of response rates right from the beginning. The model tabulation plan should make provision for, and the First Country Reports should contain a few tables on, differential non-response and on callbacks. This material should form an integral part of the description of field activities in the chapter on methodology.

(2) In the same chapter, the sample description should include an explicit reference to (a) the nature of the listing units, ie named households vs dwellings vs hybrid household–dwellings; (b) whether empty dwellings were listed and sampled; and (c) the time interval between listing and the use of the list for interviewing.

(3) Since, typically, raw data files are not rectangular, there is no good reason to delete cases of non-interview. At least one version of the cleaned full file should be retained, as well as a copy of the full original (dirty) file.
(4) In the result code for the household interview, separate categories should be adopted for 'dwelling demolished', 'dwelling not found', 'dwelling occupied but family away for duration of field work', and 'dwelling occupied, family in residence, but failed to make contact'. Though in practice these distinctions are difficult to draw, in theory they are important in the proper computation of response rates.
(5) Greater attention should be paid to monitoring the household (dwelling) listing operation. This is a phase of the survey process typically left in the hands of local field staff. Central control has been minimal, no doubt in the belief that it was a simple and straightforward task. But one of the clearest outcomes of this study has been the sizable under-coverage of households in many countries.
(6) In the training of supervisors, greater emphasis should be placed on the biases that arise when interviewers, whether deliberately or not, push potential respondents out of the eligible age range. There should be a greater amount of monitoring by supervisors, through callbacks, of women in borderline ages.
(7) There should be no relaxation in the importance attached to achieving high response rates and the related willingness to undertake callbacks to achieve this.

References

Central Bureau of Statistics, Ministry of Economic Planning and Development (1980). *Kenya Fertility Survey 1977–78. First Report*, Nairobi.

O'Muircheartaigh, Colm and Albert M. Marckwardt (1981). An Assessment of the Reliability of WFS Data. *World Fertility Survey Conference 1980: Record of Proceedings 3*. Voorburg, Netherlands: International Statistical Institute.

National Census and Statistics Office, University of the Philippines Population Institute, Commission on Population and National Economic and Development Authority (1979). *Republic of the Philippines Fertility Survey 1978. First Report.* Manila.

Thomsen, Ib and Erling Siring (1980). On the Causes and Effects of Non-Response: Norwegian Experiences. *Artikler Fra Statistisk Sentralbyra* no 121.

Timaeus, I. and K. Balasubramanian (forthcoming). Evaluation of the Lesotho Fertility Survey of 1977. *WFS Scientific Reports*.

United Nations (1982). National Household Survey Capability Programme: Non-Sampling Errors in Household Surveys: Sources, Assessment and Control; Preliminary Version. UN/TCD/SO, New York.

Verma, V.K. (1981). Sampling for National Fertility Surveys. *World Fertility Survey Conference 1980: Record of Proceedings 3*. Voorburg, Netherlands: International Statistical Institute.

World Fertility Survey (1975). Supervisors' Instructions. *Basic Documentation* no 5.

Appendix A - Detailed Tables

Table A1 Response rates and sample disposition, household interview

Country	Response rate[a]	Sample disposition					Number of successful interviews
		Vacant destroyed, etc.	Not at home	Refused	Other	Interviewed	
Africa							
Benin	-	-	-	-	-	-	20 030
Cameroon	93.8	-	-	-	-	-	37 870
Ghana	98.2	9.8	0.9	0.7	0.0	88.6	6 016
Ivory Coast	97.5	12.5	1.3	0.3	0.6	85.3	3 754
Kenya	92.8	11.0	1.3	0.3	4.8	82.6	8 891
Lesotho	99.7	10.0	0.1	0.1	0.1	89.7	18 244
Nigeria	93.4	1.4	5.3	1.1	0.1	92.1	8 624
Senegal	-	-	-	-	-	-	-
Egypt	97.4	6.3	1.2	--- 1.2 ---		91.3	10 079
Mauritania	98.6	-	-	-	-	-	14 827
Morocco	96.7	8.5	2.9	0.1	0.0	88.5	17 126
Sudan (N)	95.2	9.2	4.1	0.3	0.0	86.4	12 028
Tunisia	96.2	9.3	2.5	0.9	0.0	87.3	5 735
Asia and Pacific							
Jordan	96.2	-	-	-	-	-	14 490
Syria	96.2	0.2	3.4	0.0	0.4	96.0	14 670
Turkey	85.2	5.6	------- 14.0 ---------			80.4	5 142
Yemen A.R.	98.2	6.0	0.6	0.4	0.7	92.3	13 255
Bangladesh	98.2	4.7	0.1	0.0	0.0	95.2	5 853
Iran	92.6	-	-	-	-	-	5 677
Nepal	94.8	5.2	0.0	0.0	0.0	94.8	5 655
Pakistan	99.4	6.0	0.5	0.1	0.0	93.4	4 901
Sri Lanka	99.7	10.3	0.0	0.2	0.1	89.4	8 149
Fiji	94.2	3.5	0.7	0.2	4.6	91.0	4 901
Indonesia	96.7	0.0	------- 3.3 ---------			96.7	10 156
Korea, Rep. of	98.5	0.0	------- 1.5 ---------			98.5	20 932
Malaysia	99.8	4.1	0.1	0.1	0.0	95.7	7 755
Philippines	98.5	12.3	0.9	0.4	0.0	86.4	12 742
Thailand	98.7	3.5	------- 1.3 ---------			95.2	4 301
Americas							
Colombia	95.8	0.0	1.3	0.7	2.2	95.8	9 796
Ecuador	96.2	13.3	2.8	0.4	0.0	83.5	5 825
Paraguay	95.5	8.8	3.8	0.3	0.0	87.1	4 030
Peru	96.2	14.1	2.9	0.4	0.0	82.6	7 395
Venezuela	96.9	-	-	-	-	-	8 560
Costa Rica	98.3	8.6	1.0	0.6	0.0	89.8	4 244
Dominican Rep.	94.7	4.5	0.3	0.2	4.5	90.5	10 921
Mexico	96.0	12.4	2.7	0.6	0.2	84.1	13 080
Panama	99.3	7.8	0.6	0.1	0.0	91.5	4 771
Guyana	97.6	2.7	0.9	0.7	0.7	95.0	4 433
Haiti	89.6	9.8	2.1	0.1	7.1	80.9	3 008
Jamaica	92.9	11.0	2.9	3.2	0.2	82.7	4 613
Trinidad & Tob.	96.1	4.1	3.3	0.4	0.0	92.2	4 583
Europe							
Portugal	80.1	10.1	7.4	2.8	7.7	72.0	10 888

[a] Calculated on a base which excludes vacant, destroyed and unlocated dwellings or households.
- Figures not available

Table A2 Response rates for the individual interview

Country	De facto de jure [a]	Response rate	Reason for failure			Number of successful interviews
			Absence	Refusal	Other	
Africa						
Benin	F	-	-	-	-	4018
Cameroon	F	90.0	-	-	-	8219
Ghana	F	96.3	2.3	0.1	1.3	6125
Ivory Coast	F	85.0	1.3	0.3	13.4	5179
Kenya	F	95.8	2.2	0.5	1.5	8100
Lesotho	F	97.8	1.7	0.1	0.4	3603
Nigeria	F	96.0	1.1	0.7	2.2	9727
Senegal	F	89.7	8.4	0.1	1.8	3985
Egypt	J	97.9	0.9	0.2	1.0	8788
Mauritania	F	91.0	-	-	-	3504
Morocco	F	98.7	0.4	0.1	0.8	5801
Sudan (N)	F	97.2	1.2	0.1	1.5	3115
Tunisia	F	93.0	5.5	1.1	0.4	4123
Asia and Pacific						
Jordan	F	96.3	-	-	-	3612
Syria	F	96.2	3.5	0.2	0.1	4487
Turkey	J	92.9	-	-	-	4431
Yemen A.R.	F	92.8	2.8	0.9	3.5	2605
Bangladesh	F	98.0	1.0	0.5	0.5	6513
Iran	F	-	-	-	-	4878
Nepal	F	97.9	-	-	-	5940
Pakistan	F	99.0	0.5	0.0	0.5	4996
Sri Lanka	F	99.4	-	-	-	6812
Fiji	F	97.5	1.4	0.2	0.9	4928
Indonesia	J	97.7	1.7	0.0	0.6	9155
Korea, Rep. of	F	94.9	2.3	0.3	2.5	5430
Malaysia	F	99.2	0.1	0.1	0.6	6316
Philippines	J	96.5	2.4	0.3	0.8	9268
Thailand	F	94.4	-	-	-	3778
Americas						
Colombia	F	94.7	2.3	1.0	2.0	5378
Ecuador	J	95.3	3.2	0.3	1.2	6797
Paraguay	F	95.6	3.6	0.8	0.0	4622
Peru	F,J	93.0	6.2	0.4	0.4	5640
Venezuela	J	93.1	5.6	0.5	0.8	4361
Costa Rica	J	96.7	2.0	1.1	0.2	3935
Dominican Rep	F	96.8	1.1	0.3	1.8	3115
Mexico	J	95.3	3.5	0.5	0.7	7310
Panama	F	97.5	2.4	0.1	0.0	3701
Guyana	J	95.6	1.4	0.4	2.6	4642
Haiti	F	94.6	4.8	0.1	0.5	3365
Jamaica	J	93.6	2.2	2.8	1.4	3096
Trinidad & Tobago	J	97.2	1.2	1.2	0.4	4359
Europe						
Portugal	F	92.6	4.2	2.0	1.2	5148

[a] Residence criterion for the individual interview.

- Figure not available.

Note: Some figures are provisional, particularly those for the African countries, pending final editing of data tapes.

Table A3 Detailed outcome of fieldwork at the dwelling level

	Peru	Costa Rica	Pakistan	Ecuador
Original sample addresses	8331	-	-	6744
Extra dwellings[a]	618	-	-	235
Total addresses	8949	4724	5246	6979
Sample yield	100.0 %	100.0 %	100.0 %	100.0 %
Address a dwelling	93.7	97.4	98.7	98.0
Address not a dwelling	6.3	2.6	1.3	2.0
Occupancy	100.0 %	100.0 %	100.0 %	100.0 %
Dwelling occupied (ODU)	91.7	93.8	95.2	88.5
Dwelling unoccupied	8.3	6.2	4.8	11.5
ODUs ÷ total addresses	85.9 %	91.4 %	94.0 %	86.7 %
Successful household interviews	7395	4244	4901	5825
As % of ODUs	96.2 %	98.3 %	99.4 %	96.2 %
As % dwellings	88.2 %	92.2 %	94.6 %	85.1 %
As % of addresses	82.6 %	89.8 %	93.4 %	83.5 %

[a] Extra dwellings encountered by interviewers through application of the half-open interval.

Table A4 Differentials of individual interview non-response rates

Characteristic from HH schedule	Non-response to individual interview				
	Peru	Costa Rica	Philippines	Syria	Fiji
Total	7.0 %	3.3 %	3.5 %	3.8 %	2.5 %
Age:					
<25	8.3	3.7	4.6	3.6	2.6
25-34	4.6	2.4	2.9	3.2	2.5
35-44	7.3	3.7	3.6	3.9	2.1
45-49	10.0	4.5	4.3	5.7	2.6
Education:[a]					
A	7.8	4.1	4.0	4.0	-
B	5.4	2.1	3.0	3.7	-
C	7.5	2.9	2.2	3.1	-
D	6.8	3.2	4.4	3.3	-
E	7.1	5.1	4.9	-	-
Marital status:					
Single	-	5.0	-	-	-
Married	6.6	2.4	3.4	3.6	2.0
Other	9.9	5.4	6.4	7.4	9.8
Relationship to HH head:					
Head	6.6	3.0	4.3	-	-
Wife	5.6	2.6	3.1	-	-
Other	11.4	4.8	6.2	-	-
Area:					
Major urban	7.6	4.4	2.2		3.9
Other urban	6.0	3.3	3.5	4.0	2.7
Rural	7.6	2.4	4.0	3.6	1.9

[a] In Peru, the categories are 0,1-2, 3-4, 5,≥6; Costa Rica: <3, 3-5, 6, 7-10, ≥11; Philippines: ≤3, 4-6, 7-11, 12, ≥13; Syria 0, 1-5, 6,≥7.

- Data not available or not applicable.

Table A5 Response rates by call number, households and individuals

	Call number				Final result	Pay-off*
	1	2	3	4+		
Peru - Households[a]	8949	1351	555	268	8949	
Interviewed	70.6 %	49.0 %	40.4 %	70.1 %	82.6 %	
Definitive non-interview	14.3	9.9	11.3	29.9	17.4	
Stock for next visit	15.1	41.1	48.3	-	-	
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	79.5 %
Peru - Individuals	6062	618	254	109	6062	
Interviewed	84.3 %	54.0 %	50.0 %	61.5 %	93.0 %	
Definitive non-interview	5.5	4.9	7.1	38.5	7.0	
Stock for next visit	10.2	41.1	42.9	-	-	
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	85.3 %
Costa Rica - Households[a]	4724	344	124	72	4724	
Interviewed	84.2 %	59.3 %	33.9 %	26.4 %	89.8 %	
Definitive non-interview	8.5	4.7	8.0	73.6	10.2	
Stock for next visit	7.3	36.0	58.1	-	-	
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	76.7 %
Costa Rica - Individuals	4070	312	154	115	4070	
Interviewed	91.6 %	48.1 %	19.5 %	21.7 %	96.7 %	
Definitive non-interview	0.7	2.6	5.8	78.3	3.3	
Stock for next visit	7.7	49.3	74.7	-	-	
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	66.2 %
Ghana - Households[a]	6791	1154	540	248	6791	
Interviewed	75.6 %	46.5 %	44.6 %	43.1 %	88.7 %	
Definitive non-interview	7.4	6.7	9.4	56.9	11.3	
Stock for next visit	17.0	46.8	46.0	-	-	
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	77.1 %
Ecuador - Households[a]	7091	790	236	92	7091	
Interviewed	73.9 %	62.6 %	50.8 %	60.9 %	83.4 %	
Definitive non-interview	15.0	7.5	10.2	39.1	16.6	
Stock for next visit	11.1	29.9	39.0	-	-	
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	84.9 %
Ecuador - Individuals	7241	789	197	90	7241	
Interviewed	85.8 %	69.6 %	43.6 %	60.0 %	95.3 %	
Definitive non-interviews	3.3	5.4	10.7	40.0	4.7	
Stock for next visit	10.9	25.0	45.7	-	-	
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	87.3 %

[a] The denominator in the case of Peru, Costa Rica and Ecuador is addresses; in the case of Ghana, selected households.

* Per cent of units which were eventually interviewed among those to which a second or subsequent visit was made.

Table A6 Differential household composition by callbacks

Country	Total household population					Women 15-49		
	% under age 10	% single	% widowed divorced separated	% in top educ. category	Mean household size	% over age 30	% not currently married	% in top educ. category
Peru								
Urban, 1st call	29.2	63.4	5.2	29.0	5.79	37.3	46.2	46.4
Urban, later call	25.9	61.7	6.1	38.2	5.05	38.5	50.7	56.9
Rural, 1st call	33.3	60.5	5.3	4.6	5.48	46.1	34.2	6.2
Rural, later call	35.5	60.2	7.0	5.1	5.32	50.3	28.2	3.5
Costa Rica								
Urban, 1st call	22.2	61.5	6.2	14.0	5.12	38.9	52.7	23.5
Urban, later call	20.6	58.9	7.5	23.6	4.35	43.3	52.7	35.7
Rural, 1st call	29.1	66.1	4.3	2.1	5.88	39.3	42.7	4.1
Rural, later call	33.4	64.5	2.7	4.8	5.14	30.0	30.0	13.3
Pakistan								
Urban, 1st call	30.1	59.1	5.1	16.6	6.84	39.6	29.1	19.8
Urban, later call	28.8	57.2	6.3	19.5	5.03	37.8	23.5	25.2
Rural, 1st call	30.5	56.0	5.2	4.5	6.46	42.5	22.5	1.4
Rural, later call	30.2	57.3	5.9	3.7	5.08	43.4	27.8	1.7
Philippines								
Urban, 1st call	26.6	64.8	3.3	15.0	5.41	37.0	52.3	27.3
Urban, later call	25.2	65.5	3.2	21.1	6.30	36.1	55.9	33.2
Rural, 1st call	31.3	64.4	3.3	4.2	5.84	41.5	41.4	10.0
Rural, later call	32.5	65.7	2.7	4.8	6.15	40.5	41.3	11.9

[a] in Peru, $\geq$6 years; in Costa Rica, $\geq$11 years; in Pakistan, $\geq$6 years; in the Philippines, $\geq$13 years.

Table A7 Characteristics of respondents by individual interview callbacks

Country	N	% over age 35	% of EM not now married	% in top educ. category[a]	% who work away from home	Mean births in past 5 years
Peru[b]						
Urban, 1st call	3157	39.8	10.6	34.3	19.5	1.11
Urban, later call	475	47.5	14.6	42.8	40.8	0.88
Rural, 1st call	1922	48.0	8.6	2.8	2.4	1.32
Rural, later call	86	54.8	10.4	5.2	11.8	1.26
Costa Rica[b]						
Urban, 1st call	1486	42.6	14.1	23.6	28.9	0.69
Urban, later call	89	42.7	16.8	36.0	46.1	0.63
Rural, 1st call	1436	42.4	8.7	3.6	9.5	0.97
Rural, later call	26	46.1	11.5	7.7	30.8	0.96
Haiti						
Urban, 1st call	624	21.7 [c]	25.2	23.7 [c]	26.1	0.81
Urban, later call	143	19.4	18.2	28.1	47.6	0.75
Rural, 1st call	1318	29.3	12.7	2.6	11.5	1.15
Rural, later call	167	31.8	12.5	3.1	22.8	1.15
Ghana						
Urban, 1st call	1467	24.1 [c]	10.2	9.6 [c]	55.5	1.03
Urban, later call	122	27.5	13.9	10.2	67.2	0.99
Rural, 1st call	3157	29.7	10.3	1.6	19.1	1.12
Rural, later call	197	35.9	9.1	1.8	20.8	1.09
Pakistan[b]						
Urban, 1st call	1836	34.6	5.7	13.5	4.5	1.17
Urban, later call	73	41.0	5.5	25.7	17.8	0.97
Rural, 1st call	2975	35.1	5.5	0.7	9.3	1.12
Rural, later call	68	47.0	16.2	0.0	13.2	0.72
Philippines[b]						
Urban, 1st call	3702	46.5	4.4	21.7	24.7	1.05
Urban, later call	892	45.9	6.8	37.3	47.7	0.94
Rural, 1st call	4286	46.9	4.2	6.4	10.5	1.26
Rural, later call	388	45.9	5.2	9.0	16.5	1.27
Thailand[b]						
Urban, 1st call	438	41.6	9.6	25.8	29.0	0.83
Urban, later call	141	52.5	7.1	36.9	64.5	0.60
Rural, 1st call	2834	42.0	7.7	3.3	6.7	0.99
Rural, later call	406	46.3	8.1	6.0	15.5	0.83
Malaysia[b]						
Urban, 1st call	1283	47.4	7.2	20.5	13.8	0.99
Urban, later call	690	53.2	9.4	24.8	30.1	0.82
Rural, 1st call	3245	43.1	8.2	7.0	4.5	1.03
Rural, later call	1103	45.4	8.1	10.7	14.1	0.90

[a] For Peru, ≥6 years; Costa Rica, ≥ 11 years; Haiti, secondary or more; Ghana, ≥11 years; Pakistan, secondary or more; Philippines, college; Thailand, ≥ 5 years; Malaysia, ≥7 years.

[b] All figures are for ever-married women.

[c] Figures include single women for this variable for this country. Otherwise the figures are for ever-married women only.

Table A8 Distribution of interviews and proportion callbacks by serial day in cluster

	Per cent distribution of interviews					Per cent of int's achieved on callback				
	PE	CR	GH	PH	TH	PE	CR	GH	PH	TH
Two-day Clusters										
Day 1	-	37	63	63	74	-	0	0	3	7
Day 2	-	63	37	37	26	-	9	3	16	44
Total	-	100	100	100	100	-	6	1	7	17
n		198	1248	1761	1516					
Three-day Clusters										
Day 1	36	44	39	41	47	7	0	1	3	7
Day 2	35	33	45	36	39	5	3	3	11	23
Day 3	29	23	16	23	14	18	10	8	19	70
Total	100	100	100	100	100	10	3	3	9	22
n	171	215	1320	3467	625					
Four-day Clusters										
Day 1	28	44	26	26	39	1	0	0	4	5
Day 2	32	10	36	31	35	3	0	5	10	9
Day 3	26	23	25	26	16	3	3	6	20	38
Day 4	14	23	13	17	10	14	10	12	26	58
Total	100	100	100	100	100	4	3	5	14	17
n	332	151	884	2006	385					
Five-day Clusters										
Day 1	22	10	18	27	-	0	2	2	7	-
Day 2	24	25	24	23	-	16	4	2	12	-
Day 3	32	30	25	23	-	6	2	7	29	-
Day 4	16	28	18	15	-	10	3	10	36	-
Day 5	6	7	15	12	-	17	31	8	44	-
Total	100	100	100	100	-	8	5	6	22	-
n	202	423	712	943						
Six-day Clusters										
Day 1	31	-	16	23	-	2	-	12	22	-
Day 2	29	-	17	22	-	0	-	6	30	-
Day 3	9	-	14	16	-	6	-	12	36	-
Day 4	15	-	25	16	-	4	-	5	38	-
Day 5	10	-	20	11	-	19	-	13	49	-
Day 6	6	-	8	12	-	21	-	16	61	-
Total	100	-	100	100	-	6	-	10	36	-
n	165	-	392	387						

Table A9 Expected[a] and achieved sample sizes

Country	Households		Women	
	Expected	Achieved	Expected	Achieved
Africa				
Benin	-	20 030	5 000	4 018
Cameroon	32 000	37 870	10 000	8 219
Ghana	7 500	6 016	7 500	6 125
Ivory Coast	-	3 754	6 000	5 179
Kenya	-	8 891	10 000	8 100
Lesotho	20 000	18 244	5 000	3 603
Nigeria	11 000	8 624	10 000	9 727
Senegal	-	-	5 000	3 985
Egypt	10 000	10 079	-	8 788
Mauritania	24 000	14 827	6 000	3 504
Morocco	24 000	17 126	7 000	5 801
Sudan (N)	14 000	12 028	4 000	3 115
Tunisia	5 500	5 735	5 000	4 123
Asia and Pacific				
Jordan	20 000	14 490	4 240	3 612
Syria	-	14 670	5 000	4 487
Turkey	6 000	5 142	6 000	4 431
Yemen	15 000	13 255	5 000	2 605
Bangladesh	6 400	5 853	5 700	6 513
Iran	-	5 677	-	4 878
Nepal	5 000	5 655	5 000	5 940
Pakistan	-	4 901	6 000	4 996
Sri Lanka	9 000	8 149	-	6 812
Fiji	5 000	4 901	5 000	4 928
Indonesia	11 000	10 156	-	9 155
Korea, Rep. of	18 000	20 932	5 000	5 430
Malaysia	7 800	7 755	7 600	6 316
Philippines[b]	15 000	12 742	16 000	9 268
Thailand	4 000+	4 301	4 000	3 778
Americas				
Colombia	10 000	9 796	5 000	5 378
Ecuador[b]	7 000	5 825	6 000	6 797
Paraguay	3 800	4 030	4 500	4 622
Peru[b]	7 000	7 395	7 000	5 640
Venezuela[b]	5 000	8 560	4 000	4 361
Costa Rica[b]	3 300	4 244	4 100	3 935
Dominican Rep.	10 000	10 921	3 250	3 115
Mexico	-	13 080	8 000	7 310
Panama[b]	5 000	4 771	4 000	3 701
Guyana	5 000	4 433	6 000	4 642
Haiti	3 400	3 008	3 000	3 365
Jamaica	6 300	4 613	5 500	3 096
Trinidad & Tobago	5 000	4 583	5 500	4 359
Europe				
Portugal	14 500	10 888	6 500	5 148

a Expected sample size as specified in each project document (see text).

b A change occurred in either the definition of eligible women or in the sample 'take' for household interviews subsequent to the preparation of the project document. (There may be some others not footnoted.)

Table A10 Reasons for shortfall in obtaining individual interviews, selected countries

	Kenya	Lesotho	Sudan	Fiji	Pakistan	Jamaica	Trinidad & Tobago	Costa Rica	Panama
Planned interviews	10 000	4 993	4 000	5 000	6 095	3 500	5 500	4 021	4 500
Achieved interviews	8 100	3 603	3 115	4 928	4 996	3 096	4 359	3 935	3 701
Planned less achieved	1 900	1 390	885	72	1 099	404	1 141	86	799
Percentage deficit	-19.0	-27.8	-22.1	-1.4	-18.0	-11.5	-20.7	-2.1	-17.8
Reasons for deficit[a]									
Number of households	-4.2	-10.4	-11.8	-5.4	-20.9	-8.0	-11.1	-5.4	-13.4
Household non-response	+9.9	-1.5	+5.2	+1.0	+7.8	+10.5	+2.8	+2.3	+4.0
Number of women	-24.7	-14.3	-17.1	+3.0	-7.0	-11.7	-10.1	+0.3	-6.3
Non-response of women	*	-1.6	+1.6	*	+2.1	-2.3	-2.3	+0.7	-2.1

[a] As percentage of planned interviews.

* Allowance for non-response not projected separately for women; it is included in the household figures.