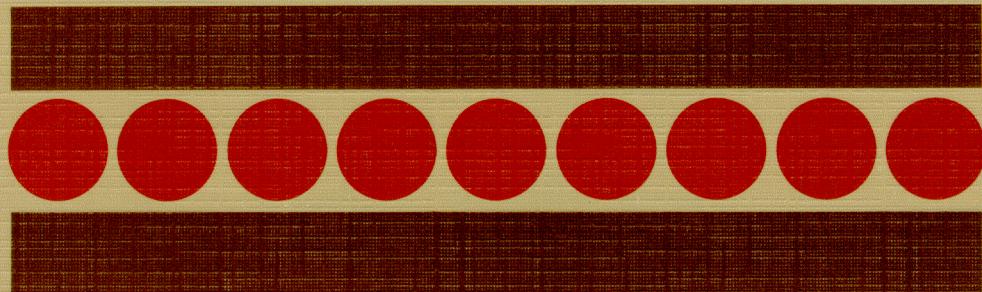


NAZIOARTEKO ESTATISTIKA
MINTEGIA EUSKADIN

1986

SEMINARIO INTERNACIONAL
DE ESTADÍSTICA EN EUSKADI



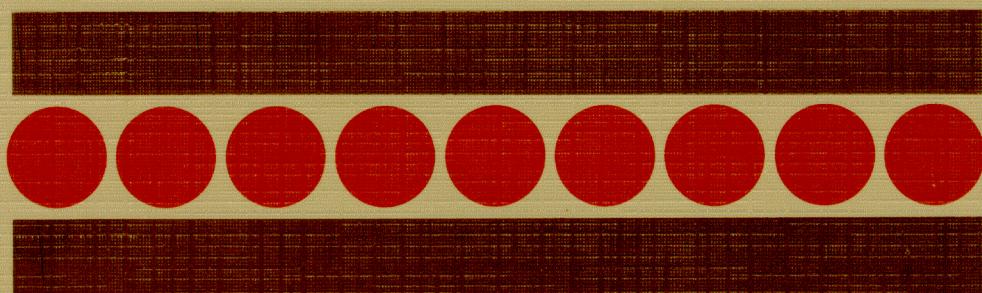
ERANTZUN EZAREN METODOLOGIA
ETA TRATAMENDUA

METODOLOGIA Y TRATAMIENTO
DE LA NO-RESPUESTA

METHODOLOGIE ET TRAITEMENT
DE LA NON-REONSE

METHODOLOGY AND TREATMENT
FOR NON-RESPONSE

R. PLATEK



NAZIOARTEKO ESTATISTIKA
MINTEGIA EUSKADIN

1986

SEMINARIO INTERNACIONAL
DE ESTADISTICA EN EUSKADI

**ERANTZUN EZAREN METODOLOGIA
ETA TRATAMENDUA**

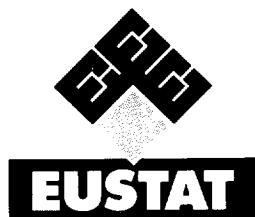
•
**METODOLOGIA Y TRATAMIENTO
DE LA NO-RESPUESTA**

•
**METHODOLOGIE ET TRAITEMENT
DE LA NON-REPONSE**

•
**METHODOLOGY AND TREATMENT
FOR NON-RESPONSE**

R. PLATEK

KOADERNOA 10 CUADERNO



Lanketa / Elaboración:

Euskal Estatistika-Erakundea /
Instituto Vasco de Estadística

Argitalpena / Edición:

Euskal Estatistika-Erakundea /
Instituto Vasco de Estadística
C/ Dato 14-16 - 01005 Vitoria-Gasteiz

Botaldia / Tirada:

1.500 ejemplares
X-1987

© **Euskadiako K.A.ko Administrazioa**
Administración de la C.A. de Euskadi

Inprimaketa eta Koadernaketa:

Impresión y Encuadernación:

Itxaropena, S.A.
Araba kalea, 45 - Zarautz

Lege-gordailua / Depósito legal: S.S. 128/88

ISBN: 84-7542-127-10 Obra completa
ISBN: 84-7749-018-x

AURKEZPENA

Estatistikako Mintegi Internazionalak sustatzean, hainbat xederekin bete nahi luke Euskal Estatistika-Erakundea, hala nola:

- Unibertsitatearekiko eta, Estatistika Sailarekiko lankidetza bultzatu.
- Funtzionari, irakasle, ikasle eta estatistikaren alorrean interesaturik leudekeen guztien birziklapen profesionala erraztu.
- Estatistikako alorrean eta mundu-mailan irakasle prestu eta abangoardiako ikerlari diren pertsonaiak Euskadira ekarri, guzti horrek zuzeneko harremanei eta esperientzien ezagupenei dagokienez suposatzen duen ondorio positiboarekin.

Iharduketa osagarri bezala eta interesaturik leudekeen ahalik eta pertsona eta Erakunde gehienetara iristearren, Ikastaro hauetako txostenak argitaratzea erabaki da, beti ere txostenemailearen jatorrizko hizkuntza errespetatuz, horrela gure Herrian gai honi buruzko ezagutza zabaltzen laguntzeko asmoarekin.

Vitoria-Gasteiz, 1988ko Otsaila
JOSE LUIS NARVAIZA SOLIS
Zuzendari Orokorra

PRESENTACION

Al promover los Seminarios Internacionales de Estadística, el Instituto Vasco de Estadística pretende cubrir varios objetivos:

- Fomentar la colaboración con la Universidad y en especial con los Departamentos de Estadística.
- Facilitar el reciclaje profesional de funcionarios, profesores, alumnos y cuantos puedan estar interesados en el campo estadístico.
- Traer a Euskadi a ilustres profesores e investigadores de vanguardia en materia estadística, a nivel mundial, con el consiguiente efecto positivo en cuanto a la relación directa y conocimiento de experiencias.

Como actuación complementaria y para llegar al mayor número posible de personas e Instituciones interesadas, se ha decidido publicar las ponencias de estos Cursos, respetando en todo caso la lengua original del ponente, para contribuir así a acrecentar el conocimiento sobre esta materia en nuestro País.

Vitoria-Gasteiz, Febrero de 1988
JOSE LUIS NARVAIZA SOLIS
Director General

PRESENTATION

L'Institut Basque de la Statistique se propose d'atteindre plusieurs objectifs par la promotion des Séminaires Internationaux de la Statistique:

- Encourager la collaboration avec l'Université et spécialement avec les départements de la statistique.
- Faciliter le recyclage professionnel des fonctionnaires, professeurs, élèves, et tous ceux qui pourraient être intéressés par la statistique.
- Inviter en Euskadi des professeurs mondialement renommés et des chercheurs de premier ordre en matière de Statistique avec tout ce que cela pourrait entraîner comme avantage dans les rapports et l'échange d'expériences.

En outre, il a été décidé de publier les exposés de ces rencontres afin d'atteindre le plus grand nombre de personnes et d'institutions intéressées, et pour contribuer ainsi à développer dans notre pays les connaissances sur cette matière. Dans chaque cas la langue d'origine du conférencier sera respectée.

Vitoria-Gasteiz, Février, 1988

JOSE LUIS NARVAIZA SOLIS
Directeur Général

PRESENTATION

In promoting the International Seminars on Statistics, the Basque Statistical Office is attempting to achieve a number of objectives:

- Encourage joint working with the Basque University and, in particular, with its Department of Statistics.
- Facilitate the in-training of civil servants, teachers and students and of all those interested in the field of statistics.
- Bring to Euskadi distinguished academics and researchers in the front line of statistics work, at a world-wide level, with all the benefits that this will bring through direct contacts and the interchange of experiences and ideas.

As an additional step, it has been decided to publish in advance the papers to be presented at these courses, respecting the native language of the speaker, in each case. This is in order that as many interested people and institutions as possible are made aware. In this way we hope to contribute to the growth and awareness concerning this topic in our country.

Vitoria-Gasteiz, February, 1988

JOSE LUIS NARVAIZA SOLIS
Director General

SARRERA

Euskal Estatistika Erakundea eta Euskal Herriko Unibertsitateak antolatutako IV Nazioarteko Estatistika Mintegia Euskadin barru, R. Platek emandako "Erantzun ezaren Metodologia eta Tratamendua" buruzko kurtsoa, liburu honek laburzen du. L. Kish, I.S.I.-ko ohorezko bazkidea, Michiganeko Unibertsitatekoa eta I.A.S.S.-eko (1983-85) Lehendakaria "Estatistika eragiketa laginketaz" kurtsoarekin, eta I. Gallastegi, Euskal Herriko Unibertsitateko Ekonometri Katedraticoa "Denborazko serieen analisia. Aurre-sate teknika batzu" kurtsoarekin IV Mintegi honetako partaideak izan ziren.

INTRODUCCION

Este libro resume el curso que sobre "Metodología y Tratamiento de la no-respuesta" ha impartido R. Platek dentro del IV Seminario Internacional de Estadística en Euskadi, organizado por el Instituto Vasco de Estadística y la Universidad del País Vasco. Este IV Seminario cuenta además con la participación de L. Kish, miembro honorario del I.S.I. y Presidente del I.A.S.S. (1983-85), de la Universidad de Michigan, con un curso sobre "Operaciones estadísticas por muestreo", e I. Gallastegi, catedrática de Econometría, de la Universidad del País Vasco, con un curso sobre "Análisis de series temporales. Algunas técnicas de Predicción".

INTRODUCTION

Ce livre résume le cours sur "Méthodologie et Traitement de la non réponse" qui a été fait par R. Platek dans le IV Séminaire International de la Statistique au Pays Basque, organisé par l'Institut Basque de la Statistique et l'Université du Pays Basque. Ce IV Séminaire compte aussi avec la participation de L. Kish membre honoraire de l'ISI et Président du IASS (1983-85), de l'Université de Michigan, avec un cours sur "Opérations statistiques par échantillonage" et I. Gallastegui, professeur d'Econométrie à l'Université du Pays Basque avec un cours sur "L'Analyse des séries temporelles. Quelques techniques de prédiction".

INTRODUCTION

This book summarizes the course on "Methodology and Treatment of the non-response", given by R. Platek during the 4th International Statistics Seminar, organized by the Basque Statistical Institute and the University of the Basque Country. This 4th Seminar counts on the participation of L. Kish, honorary member of ISI, and President of the IASS (1983-85), from Michigan University, with a course on "Statistical operations by sampling", and I. Gallastegi, University Professor of Econometry from University of the Basque Country with a course on "Analysis of time series. Some prediction techniques".

BIOGRAFIA

Richard Platek, 1928an Canadian jaio zen. Filosofian Doktorea, Londres eta Carleton (Canada-ko) Unibertsitateetan ikasi zuen (Matematika, Estatistika, Ekonomikak). Kanadako Estatistika Bulegoko Metologo eta Garapen Zentsu eta Etxeetako Inkestaren Zuzendaria da, A.S.A. (American Statistical Association)-ko ohorezko bazkidea, I.S.I. (International Statistical Association)-ko bazkide aukeratua, I.A.S.S. (International Association of Survey Statisticians)-eko Kontseiluaren Bazkidea. Berrogeitamar artikulu baino gehiago idatzi ditu, haien artean askok liburu eta aldizkarietan jasota izan dira. Canada estatistiketan "Journal of Survey Methodology"rekin nazioarteko fama nabaria garatu du. Aldizkari honen hasieratik Komite Argitaratzailearen Zuzendari izan da. "Current Topics in Survey Sampling" 1981en Academic Press-ek argitaratua, "Imputation Methods" John Wiley and Sons argitaratuko dute, "Development and design of Survey Questionnaire" 1985 an Statistics Canada argitaratua "Estimation for small area statistics" 1986ko Urrian John Wiley and Sons-ek argitaratua, "Methodology for small area Statistics" Carleton-eko Unibertsitateak argitaratua, liburu horien egilea izan da.

Richard Platek, nació en Canadá en 1928. Doctor en Filosofía, se educó en la Universidad de Londres y la Universidad Carleton de Canadá (Matemáticas, Estadística, Económicas). Es Director de la División de Metodología y Desarrollo en el Área de Censos y Encuestas en Hogares en la Oficina Estadística de Canadá, Miembro Honorario de la A.S.A. (American Statistical Association), Miembro Elegido del I.S.I. (International Statistical Institute), Miembro del Consejo del I.A.S.S. (International Association of Survey Statisticians). Ha escrito casi cincuenta artículos, muchos de los cuales han sido incluidos en libros y varias revistas. Ha desarrollado una revista de Metodología de Encuestas en la Oficina Estadística de Canadá, que ha logrado prominencia internacional y ha sido Director del Comité Editorial de la Revista desde su iniciación. Ha sido co-autor de libros como: "Current topics in survey sampling" publicado por Academic Press en 1981, "Imputation methods" que será publicado por John Wiley and Sons, "Development and design of survey questionnaire" publicado por Statistics Canada 1985, "Estimation for small area statistics" publicado por John Wiley and Sons en octubre 1986, "Methodology for small area statistics" publicado por la Universidad Carleton en 1986.

Richard Platek, est né au Canada en 1928. Docteur en Philosophie, il fit ses études à l'Université de Londres et à l'Université de Carleton au Canada (Mathématiques, Statistiques, Économiques), il est Directeur de la Division de Méthodologie et Développement au Service de Recensements et Enquêtes aux foyers au Bureau de la Statistique du Canada, membre honoraire de la A.S.A. (American Statistical Association) membre élu de l'I.S.I. (Institut International de la Statistique), membre du Conseil de la I.A.S.S. (Association Internationale des Statisticiens d'enquêtes). Il a écrit presque cinquante articles, plusieurs d'entre eux ont été inclus dans des livres et des revues. Il a développé un "Journal of Survey Methodology in Statistics Canada" (journal de méthodologie d'enquêtes) qui a obtenu une rélevance internationale, il a été Directeur du Comité Editorial de la revue depuis son commencement. Il a été co-auteur de livres comme "Current topics in Survey Sampling" publié par l'Académie Press en 1981, "Imputation Methods" qui sera publié par John Wiley et Fils, "Development and Design of Survey Questionnaire" publié par Statistics Canada 1985, "Estimation for small area statistics" publié par John Wiley et Fils en Octobre 1986, "Metodologie for small area statistics" publié par l'Université de Carleton en 1986.

Richard Platek, was born in Canada in 1928. Doctor of Philosophy, educated at the University of London and Carleton University Canada. (Mathematics, Statistics, Economics). He is Director Census and Household Surveys Field, Household Surveys Development Division, Statistics Canada, honorary Rank of Fellow of the A.S.A. (American Statistical Association), elected member of the I.S.I. (International Statistical Institute), member of the council of I.A.S.S. (International Association of Survey Statisticians). He has written close to fifty papers, many of which have been published in books and various journals. He has developed a Journal of Survey methodology in Statistics Canada which has achieved and international prominence and he has been chairman of the editorial board of the Journal from its inception. He has co-authored books on: Current Topics in Survey Sampling published by Academic Press 1981, Imputation Methods to be published by John Wiley and Sons, Development and Design of Survey Questionnaire published by Statistics Canada 1985, Estimation for Small Area Statistics published by John Wiley and Sons in Oc. 1986, Methodology for Small Area Statistics published by Carleton University 1986.



AURKIBIDEA

INDICE

Aurkezpena/Presentación/Présentation	
Presentation	5
Sarrera/Introducción/Introduction	
Introduction	7
Biografía	9
Laburpena	13
I. Introduction to Nonresponse	15
II. The Definition of Response/Nonresponse Rates	21
a. Response/Nonresponse Components	21
b. Definitions of Rates	22
1. Eligibility Rate	22
2. Response and Completion Rates	22
3. Contact Rates	24
4. Refusal Rates	24
5. Item Response/Nonresponse Rates	25
6. Weighted Rates and Characteristic Rates	25
7. Size of Response/Nonresponse Rates	26
III. Nonresnse in the Labour Force Survey	27
1. Behaviour of Nonresponse over time	27
2. Nonresponse and Tenure in the Survey	29
3. Incidence of Nonresponse for Interviewer Assignments	31
IV. Nonresponse by Type of Survey	32
V. The Role of Nonresponse in Surveys	33
1. Design	33
2. Questionnaire	35
3. Relative Bias due to Nonresponse	36
4. Data Collection	37
5. Motivation of Respondent	38
6. Respondent Relations	39
7. Analysis	39
8. Survey Management	40

I. Introducción a la no-respuesta	63
II. La definición de la tasa de respuesta/no-respuesta	69
a. Componentes respuesta/no-respuesta	69
b. Definición de tasas	70
1. Tasa de elegibilidad	70
2. Tasa de respuesta y cumplimentación	70
3. Tasa de contacto	72
4. Tasa de rechazo	73
5. Tasa de respuesta/no-respuesta a un item	73
6. Tasas ponderadas y tasas características	74
7. Tamaño de las tasas de respuesta/no-respuesta	74
III. No-respuesta en la encuesta de fuerza de trabajo	75
1. Comportamiento de la no-respuesta en el tiempo	75
2. No-respuesta y tiempo de permanencia en la encuesta	77
3. Incidencia de la no-respuesta en las asignaciones del entrevistador	79
IV. No-respuesta por el tipo de encuesta	81
V. El papel de la no-respuesta en las encuestas	82
1. Diseño	82
2. Cuestionario	84
3. Sesgo relativo debido a la no-respuesta	85
4. Recogida de datos	86
5. Motivación del que responde	87
6. Relaciones con el que responde	88
7. Análisis	88
8. Dirección de la encuesta	90

VI. Dealing with Nonresponse	42	VI. Tratamiento de la no-respuesta	91
1. Callbacks	42	1. Encuestación repetida	91
2. Proxy	43	2. Delegadas (Proxy)	92
3. Refusals	43	3. Rechazos	92
4. Telephone	43	4. Teléfono	92
5. Quota	43	5. Cuotas	92
6. Substitution in the Field	44	6. Sustitución en el campo	93
7. The use of incentives	45	7. El uso de incentivos	94
VII. Types and Conceptual Issues of Nonresponse	46	VII. Cuestiones conceptuales y tipos de no-respuesta	95
VIII. Imputation Methods	48	VIII. Métodos de imputación	97
1. Deductive Imputation	48	1. Imputación deductiva	97
2. Hot Deck	48	2. Fichero caliente	97
3. Modified Hot-Deck	48	3. Fichero caliente modificado	97
4. Regresion	49	4. Regresión	98
5. Weighting	49	5. Ponderación	98
6. Duplication	50	6. Duplicación	99
7. Substitution	50	7. Sustitución	99
8. Raking Ratio	51	8. Tasa RAD	100
IX. Response Probability Approach	53	IX. Aproximación probabilística a la respuesta	103
1. The Concept of Response Probability	53	1. El concepto de probabilidad de respuesta	103
2. Various Approaches to Response Probability	54	2. Diversas aproximaciones a la probabilidad de respuesta	104
3. Theoretical Implications of Response Probability Approach	54	3. Implicaciones teóricas de la aproximación a la probabilidad de respuesta	104
4. Practical Implications of Response Probability Approach	55	4. Implicaciones prácticas de la aproximación a la probabilidad de respuesta	105
X. Application of Adjustments for Nonresponse in the Canadian Labour Force Survey	56	X. Aplicaciones de ajustes por no-respuesta en la Encuesta sobre Fuerza de Trabajo Canadiense	106
XI. Reporting Nonresponse	59	XI. Informando sobre la no respuesta	109
References	61		

ERANTZUN EZAREN METODOLOGIA ETA TRATAMENDUA

LABURPENA

Informazio estatistikoa, bere parterik handienean behintzat, zentsu eta inkestetatik lortzen da. Zein ere den bere jatorria, informazioak beti jasango ditu halako hutsune batzuk ez-erantzunaren kariaz. Onartzen da, oro har, ez-erantzuna datuen kalitatearen neurri garrantzitsu bat dela, nahiz eta garrantzi hori gero maila desberdinetan tratatua izan. Erantzunak sar dezake alborapenik estimazioak egiteko orduan, eta laginketa-bariantza areagotu, laginaren tamaina murriztu egiten duelako. Jakinekoa den bezala, laginaren tamainaren kontrako proportziokoa da laginketaren bariantza. Ez-erantzunaren tamaina-ren eta alderapenaren arteko erlazioak, bestalde, garrantzi handiagoa badu ere, ez da, agian, hain bistakoa, zeren eta ez-erantzunaren magnitudearekiko duen adinakoxe menpekotasuna bait du erantzuten dutenen eta ez dutenen arteko ezaugarri-diferentziarekiko ere.

Eta hain zuzen bi osagiaok duten garrantziari buruzko gogoetak egingo dira ikastaro honetan, bai inuesta beteko emaitzen interpretazioari dagokionez eta bai presupostua diseinuko fasean nola asignatu erabakitzeari dagokionez.

Zentzu praktiko batean inuesta bateko etapa desderdinetan, ez-erantzuna edo honen osagai batzuk izan daitezke oso baliagarriak arazo operatiboak juzgatu eta datuen balioa baloratzeko orduan.

Ez-erantzunaren estudioiko oinarri bezala, inuesta edo zentsu bat egin aurretik kontutan eduki behar diren funtsezko hainbat etapa laburbilduko ditugu. Lehenik, aztergai bezala hartu den biztanleria

identifikatu eta definitu behar da. Eta bigarren, biztanleria hori errepresentatuko duen laginketa-edo zentsu-marko bat definitu.

Behin lagina hautatuz gero, datuak denboraldi jakin batean (esate baterako astebete, hilabete, etab. batean) lortu behar dira, estatistikak erreferentzi denbora espezifiko horrekin bat etorizeko moduan alegia. Zenbat eta denbora gehiago sartu inkesta egiten, hainbat eta handiagoa izango da inkesta-ko datuak lortzen diren unearen eta erreferentziako denboraren arteko aldea. Eta ondorioa izan daiteke erantzunetako erroreak gehitrea eta, beraz, nola erantzunaren alderapena hala bariantza bera ere areagotzea. Jakina, erroreak eman daitezke erantzunetan aipatu horretaz bestelako kausengatik ere, zeren eta gerta bait daiteke erroreak egotea: galderak ondo ez ulerizeagatik, inkestagilea nahasteagatik, inkestatuek datu gutxi jaso dituztelako (esate baterako, zergei buruzko datu gutxi), inkestatuek ezin dutelako edo nahi ez dutelako argi eta garbi mintzatu inkesta-gardian, inkestatuek propio erantzun okerrak ematen dituztelako (adibidez errenta, arazo delikatuei buruzko jarrera, ...).

Erantzunetako errore-arazo hauek beti azaltzen dira, zein ere den datuak erdiesteko metodoa, hots, elkarrizketa personala, telefonozko elkarrizketa, posta edo metodo hauen bestelako edozein konbinazio edo aldakuntza. Elkarrizketarien sentimendu personalak ere transmiti dakizkieke inkestatuei, horrela errore koerlazionatuak sortzeko bidea emanez. Telefonozko elkarrizketek arindu dezakete baina ez erabat ezabatu errore koerlazionatu horietarako joera.

Eta erantzunetako erroreez gainera falta daitezke datuak ez-erantzunarengatik ere. Ze ez-erantzuna edozein datu-bilketa-motatan ematen bait da. Elkarrizketa pertsonalak, hala ere, telefonozkoak baino ez-erantzun-tasa baxuagoak ematen ditu. Galdera partikular bati buruzko ez-erantzuna, gehiago da inkestatuen konplexutasunarengatik edo galde sortaren diseinuarengatik, datu-bilketakako metodologiarengatik baino. Ez-erantzun-mota honek baloreak egotzi beharra suposatzen du, alde batetik, eta beren hedadura noraino iristen den zehaztea nekez lor daitekeen egorpen-erroreak ekartzen, bestetik, ondorio bezala.

Ez-erantzuna tratatzeko modu desberdinak ikusiko dira, bada eta berdin falta diren datuak inputatzeko prozedura diferenteak ere, hauek inkestaren baldintzen edo lortu nahi diren helburuen arabera erabiliz.

METHODOLOGY AND TREATMENT FOR NON-RESPONSE

I. INTRODUCTION TO NONRESPONSE

a) Statistical information is usually derived from such sources as surveys and censuses. Whatever its origin the information will suffer from missing observations due to nonresponse. Nonresponse has been generally but to a varying degree recognized as an important measure of data quality. It introduces a possible bias in the estimate and an increase in sampling variance due to a reduction in sample size. We know, that sampling variance is inversely proportional to a sample size. Thus, as estimate based on a simple random sample with 80% response rate will have a sampling variance that is 12.5% higher than the variance of the corresponding estimate with 90% response rate. The relationship between bias and the size of nonresponse, while perhaps more important is less obvious since it depends on both the magnitude of nonresponse and the difference in the characteristics between respondents and nonrespondents. Let us assume, however, that nonresponse bias is proportional to nonresponse rate. Then, independently of sample size the percentage bias is the same for a given nonresponse rate. Sampling variance, on the other hand will be affected by changes in the sample size. One can, therefore speculate about the relative

importance of the two components in both the interpretation of survey results and in the allocation of funds at sample design stage. To illustrate this let us look at the coefficient of variation which is an important measure of data quality. When sample size increases, the coefficient of variation of an estimate will diminish. The percentage bias, however, will remain the same as long as there is no change in the size of nonresponse rate. Thus, relative to the size of coefficient of variation, the bias assume a more important and dominating role as far as the estimate is concerned. This would imply that in large samples, usually designed for estimates on a national level, survey designers should concern themselves more with the size of bias than with the size of variance. On the other hand the bias due to nonresponse may be not nearly as serious relative to sampling variance for small samples. In this case, the variance will assume a dominating role due to an increase in its size relative to nonresponse bias. It should be noted that the confidence interval of the estimate may be affected accordingly. In small samples, the confidence interval may cover the true value due to the large sampling variance but it may not do so

in the case of large samples when the sampling variance is relatively small.

This simple analysis may have important practical implications. For example, if the purpose of a survey is to obtain estimates mainly on a national level, more attention should be paid to the reduction of biases. In other words, more resources should be diverted to the sources of biases i.e. of interviewers training and control of interviewing process. Of course the opposite would apply if the estimates were required for small areas. Unfortunately, in practice survey results are used at both levels and the above observations, although helpful, should be used judiciously. But, in a still more practical way and at different stages of surveys, nonresponse or some different components of it may be very helpful in assessing operational problems and provide an insight into the reliability of survey data.

b) As a background to nonresponse let us illustrate various fundamental requirements that have to be met in order to take a survey or a census. First, it is advisable that a target population be identified and carefully defined. A simple example of a target population may be city dwellers in some metropolitan area and examples of list frames to represent that population are city directories (with addresses) and telephone directories. Second, a sampling or census frame must be defined to represent the target population. The frame may be a list frame, an area frame or a mixture of both. An example of an area frame to represent the population in a city may be a city map with the streets and blocks. Telephone lists and city directories do not cover everyone in a city and it might be necessary to utilize an area frame to derive a sample of the residual population excluded from the telephone or city directory.

Once the sample has been selected, the survey data must be obtained in a given length of time (e.g., interview week, interview month) to derive statistics that pertain to a specific reference period. The longer the time taken for the survey, the wider the gap between the time of obtaining the survey data and the reference period. The result may be an increase in response errors, resulting in an increase of either or both response bias and response variance. Response errors may of course arise for reasons other than time lag and some examples include (i) misinterpretation of questions (ii) misunderstanding the interviewer, (iii) poor records on the part of respon-

dents (e.g., poor tax records), (iv) inability or unwillingness of respondents to express themselves clearly during a survey and (v) purposely incorrect responses by respondents (e.g., income, attitude toward sensitive issues). These response error problems may arise, regardless of the method of obtaining the survey data, i.e., (i) face-to-face interview (ii) telephone interview or (iii) mail survey or a combination or a variant of these methods. However, the personal feelings of an interviewer may be transmitted to the respondents during the survey so that correlated response errors may occur. Telephone interviewing may alleviate but not entirely eliminate this tendency for correlated errors. Mail surveys are useful for simple straightforward surveys where self-interview is undertaken. The type of mail surveys include drop-off, pick-up or mail-back questionnaires and, depending upon the intervention of the interviewer or the lack thereof, some or no correlated response errors will occur. An interviewer may have some opportunity to aid the respondents in case of difficulty, thereby removing some response errors.

In addition to response errors, there occur missing data due to either item or unit nonresponse. The nonresponse may occur, under any method of data collection. However, face-to-face interviewing usually results in a lower nonresponse rate than telephone interviewing. The mail surveys tend to result in the highest nonresponse rate among the methods of data collection. The problems of item nonresponse are related more to the complexity of the survey data and questionnaire design than to the actual method of data collection. Item and unit nonresponse usually imply a need for imputation for missing data which may result in imputation errors, the extent of which is frequently difficult to determine.

There are many ways of dealing with nonresponse and many procedures to impute for missing data due to nonresponse, each serving a somewhat different purpose and different survey conditions.

A flow chart (Table 1) summarizes the stages of obtaining survey data, beginning with the target population and culminating with the estimation and tabulation. In between these stages, there are many problems which result in errors of omission of units (undercoverage), missing data (unit and item nonresponse) and response errors. These are by no means the only steps or the only problems in obtaining survey data. However, most

of the problems of data quality pertain to the steps and problems presented in the flow chart.

As noted in Table 1 the immediate potential or an actual problem of an incomplete list and/or area frame occurs prior to the sample selection. This leads to coverage errors in the sampling frame when it does not adequately represent the target population at the time of the survey. This occurs for the following reasons: (i) units in the target population are missing from the sampling frame (undercoverage), (ii) units not in the target population are included in the sampling frame (overcoverage) and (iii) units are included more than once in the sampling frame (duplication, resulting in overcoverage). There will also be undercoverage in the sample based on the sampling frame since there is no chance of selecting some units in the target population omitted from the sampling frame. This is indicated by the connecting lines from errors in list frame (box 5) and errors in area frame (box 6) to sample selection, (box 7). In general, missing sampled units as a result of undercoverage are not known during the data collection operation. It is usually only after completion of the data collection operation (box 8) and data processing, which include the steps from questionnaire edit to estimation that undercoverage or overcoverage in one or more post-strata cells is revealed; hence, the connecting line from imputation for under/overcoverage (box 17) to estimation (box 21) only.

Another branch in the flow chart proceeds from data collection operation (box 8), when unit response or unit nonresponse may occur (boxes 10 and 11 respectively), or the unit may be non-existent or incorrectly included (box 9), contributing to over-coverage (box 13). Unit response is followed by a questionnaire edit (box 12), which may reveal (i) complete and consistent entries (item nonresponse) or (iii) a completely unusable questionnaire (boxes 14, 15, 16). While item nonresponse, including inconsistent entries are usually detectable during the edit questionnaire procedure, it is usually impossible to detect response errors of an item if there is no logical reason for the item to appear faulty. Thus, one would proceed directly to estimation despite a possible response error (box 18). Imputation errors

(box 20) may result from imputation for item or unit nonresponse (box 19) and may be also undetectable. Consequently, one proceeds directly to estimation (box 21).

Unit and item nonresponse (boxes 11, 15 and 16) are immediately followed by imputation for missing data (box 19). It should be noted that, depending upon the imputation procedure and availability of external sources of data, some imputation may be undertaken immediately after the detection of unit and item nonresponse (e.g., hot deck or other duplication methods). The other imputation methods must await the estimation stage (e.g., weight adjustment to enlarge the sample, substitution of last month's data), hence, connecting dotted line between imputation for nonresponse and estimation.

Any imputation procedure (box 19) will likely result in imputation error (box 21), which is present in both micro-data pertaining to sampled units and possibly in the estimates at macro-levels if the imputation errors after weighting up the data do not cancel out, hence the connecting line in both directions between imputation error and estimation.

The estimation (box 21) and tabulation (box 22) culminate in an analysis of mean square errors (not included in flow chart), which includes bias, arising from both response and imputation errors, sampling variance (unless the data are obtained in a census), and non-sampling variance, arising from variations in response and imputation errors.

Survey data invariably require adjustments necessitated by various deficiencies due to missing values or response errors, if observed. In Table 2 an imputation/response error model is given to illustrate the stages at which a particular unit will contain specific values. It can be seen that various types of errors of observed or imputed values of a survey characteristic "y" will result, depending upon what occurs at unit level. For example, if the unit responds (box 2) but fails to respond with respect to characteristic "y" (item nonresponse, box 5) an imputation error, NR_{iy}^{eI} , may result in the imputed value z_{iy}^I (box 7).

TABLE 1
FLOW CHART FOR SURVEY DATA

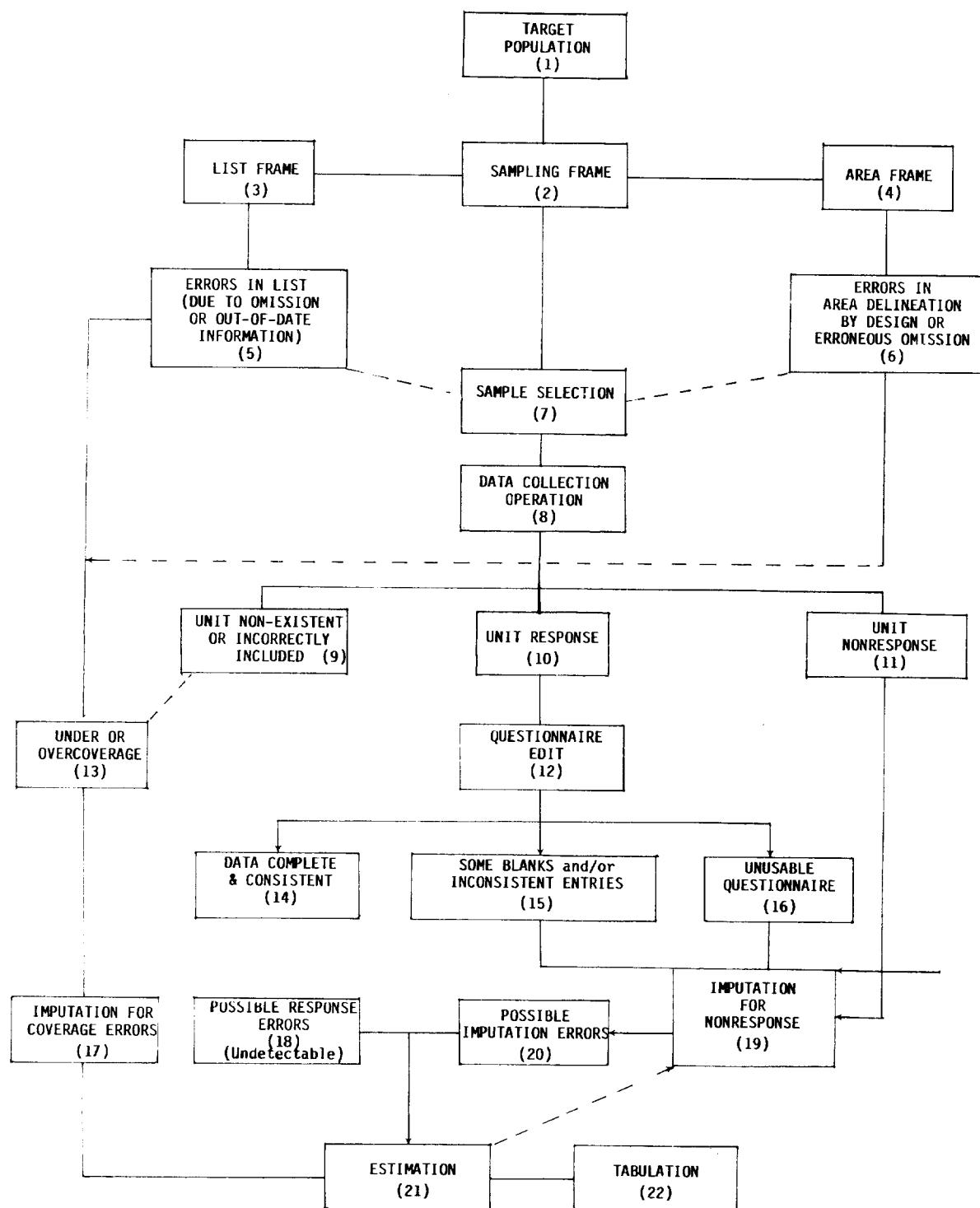
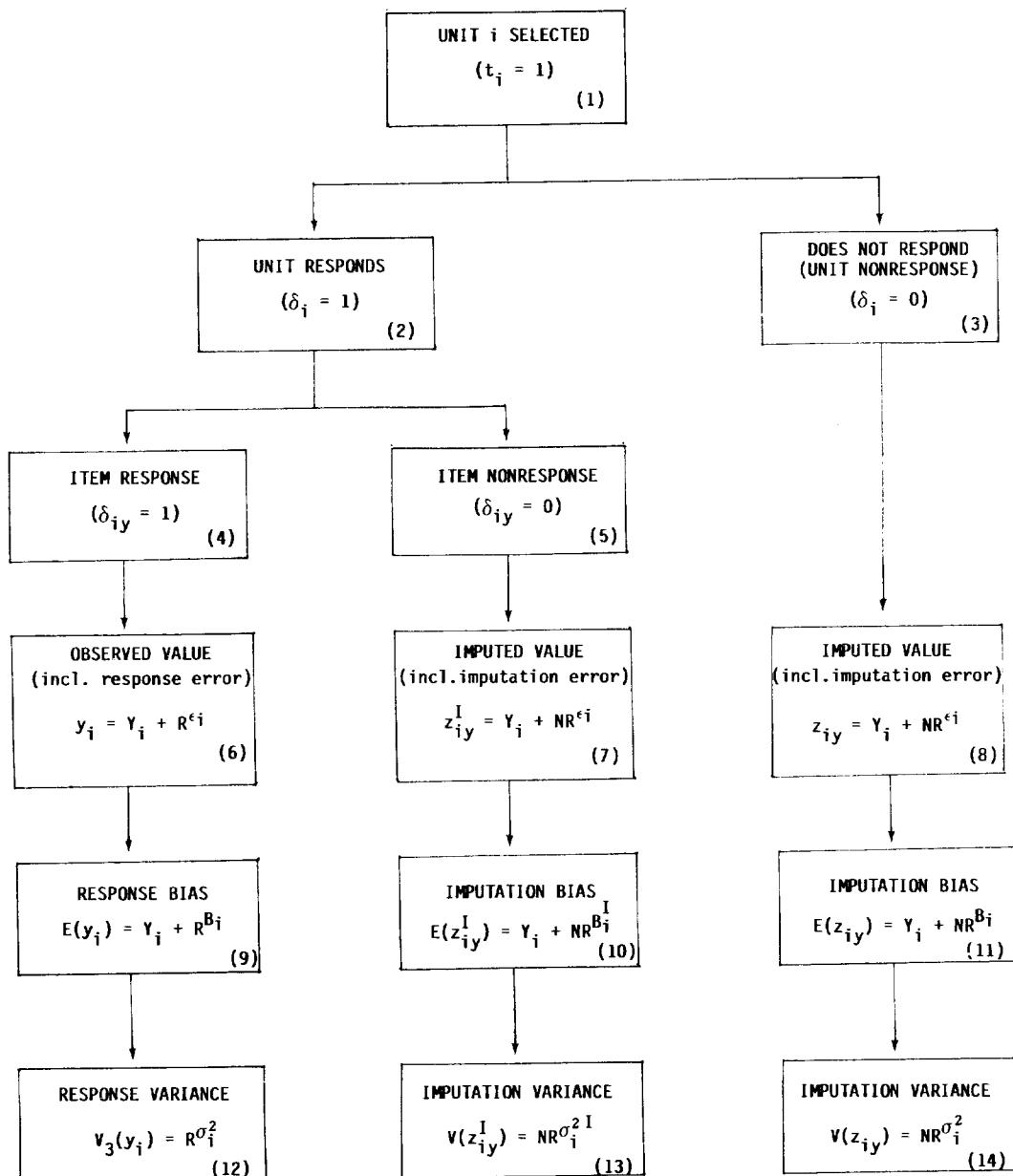


TABLE 2
IMPUTATION/RESPONSE ERROR MODEL (Unit Level)



$$u_{iy} = \delta_i [\delta_{iy} y_i (1 - \delta_{iy}) z_{iy}^I] + (1 - \delta_i) z_{iy}$$

= observed or imputed value for missing data to item or unit nonresponse.

where t_i = 1 or 0 according to whether a unit is selected or not selected.
 δ_i = 1 or 0 according to whether a unit responded or did not respond.
 δ_{iy} = 1 or 0 according to whether item response was obtained or not obtained.
 y_i = true value
 Y_i = observed value
 R_i^ϵ = response error
 Z_{iy} = imputed value
 NR_i^ϵ = imputation error
 R_i^B = response bias
 MR_i^{BI} = imputation bias
 $R_i^{\sigma^2}$ = response variance
 $NR_i^{\sigma^2}$ = imputation variance

II. THE DEFINITION OF RESPONSE/NONRESPONSE RATES

Recently, nonresponse has been increasing in many surveys all over the world. Consequently, there is a greater need than ever before to monitor nonresponse rates, to make comparisons between surveys, countries, survey organizations, and to ensure some degree of comparability. There have been attempts to standardize the definition of response rate and its complement, the nonresponse rate; see for example, Kviz (1977) Cannell (1978). Problems of inconsistent definitions of response rates related to telephone surveys are described by Wiseman and McDonald (1980).

There are also problems of inconsistent terminology with regard to response/nonresponse in surveys. Terms such as completion rate, contact rate, and under-coverage rate have been used in different contexts in reports and articles dealing with data collection. While these terms may be readily distinguished in an individual report, they may be confusing and subject to conflicting interpretations, when studying different reports.

To consider response/nonresponse problems, a distinction must be made between unit and item nonresponse rates. Unit nonresponse rates generally pertain to the level at which survey data are gathered during the first contact. Examples of the level could be a dwelling, individual, store or establishment. However, in the case of multi-stage sampling, there may be nonresponse of all units within clusters or even primary sampling units (psu) so that unit nonresponse could apply to a selected cluster or psu as well as a dwelling or individual.

Item nonresponse usually pertains to the questionnaires, where information has been provided for some questions but not to all that should have been provided. However, if a unit fails to respond, it automatically fails to respond to any item. Hence unit nonresponse and item nonresponse are distinct events that should be dealt with separately.

The response rates pertain to the whole sample and part of a sample such as design-dependent areas or they may apply to administrative areas such as an interviewer assignment, or a group of assignments overseen by a supervisor or field office.

a) Response/Nonresponse Components

In order to define various response rates and discuss their uses and applications, it is necessary to split up the target population for the sample or census into

the various components, by type of response/nonresponse. Table 3 accomplishes this very purpose, indicating most of the important components of the whole survey that will be used in the rates. Once a target population (Box 1) is defined for a survey, a survey frame of N units (Box 2) is then determined.

For the survey to be taken, a data gathering procedure (Box 3) and appropriate design are decided upon.

$t_i = 1$ or 0 according as unit i is selected or not,

Σ = summation over all N units in the survey frame.

Out of the sample of n units, $\Sigma t_i e_i$ are eligible (Box 8) and $\Sigma t_i (1 - e_i)$ are ineligible (Box 5) for the survey, where

$e_i = 1$ or 0 according as unit i is eligible or not.

Sometimes the eligibility criterion may not be determined if the unit cannot be contacted while at other times the eligibility criterion is obvious from the physical appearance, such as vacant/non vacant dwellings in a household survey.

The $\Sigma t_i (1 - e_i)$ ineligible units of (Box 5) may be split up between $\Sigma t_i (1 - e_i) (1 - \delta_i)$ units not interviewed just as they should not have been (Box 6) and $\Sigma t_i (1 - e_i) \delta_i$ units incorrectly interviewed (Box 7). One hopes that the number of such units in Box 7 is non-existent or at least very small. However, if such units are discovered, they should be deleted from the sample. In the above and in the breakdowns that follow, $\delta_i = 1$ or 0 according as unit i responded or did not respond.

The $\Sigma t_i e_i$ eligible units (Box 8) may be split up between $\Sigma t_i e_i \delta_i$ unit respondents (Box 9 + Box 10) and $\Sigma t_i e_i (1 - \delta_i)$ unit nonrespondents (Box 11), i.e., they provided unusable survey data and little, if anything, is known about the units, except perhaps their geographic location.

The $\Sigma t_i e_i \delta_i$ units respondents may be split up first between $\Sigma t_i e_i \delta_i \pi(\delta_{iy})$ units, free of item nonresponse, but with possible response errors (Box 9) and $\Sigma t_i e_i \delta_i [1 - \pi(\delta_{iy})]$ units with item nonresponse in at least one characteristic but not in all characteristics (Box 10). Here $\delta_{iy} = 1$ or 0 according as responding unit i responds or does not respond to item or character-

ristic y . In (Box 9), $\delta_{iy} = 1$ for unit i and all items while in (Box 10), $\delta_{iy} = 0$ for one or more items but not for all of them. For a particular item y some of the $t_i e_i \delta_i \delta_{iy}$ item y respondents (Box 12) come from those unit respondents, free of item nonresponse in (Box 9) while the remainder come from those unit respondents with some item nonresponse among one or more items other than item y . The $t_i e_i (1 - \delta_{iy})$ item y nonrespondents of (Box 13) come from those unit respondents with some item nonresponse of (Box 10) that include item y .

The item y respondents of (Box 12) may be decomposed into three components, (i) those units with item y free of response error, (ii) those with a detected response error for item y , and (iii) those with an undetected response error for item y , in Boxes 15, 16A and 16B respectively.

b) Definition of Various Rates

The sample of $n = \sum t_i$ units decomposed in Table 1 in section (2) into eligible units, unit respondents/nonrespondents, refusals, item respondents/nonrespondents, etc. leads to many different types of rates which are defined below. For each rate, the numerator is a particular subset of the denominator. Wherever possible, the rate is defined in terms of the counts of units as broken down in Table 3.

(i) Eligibility Rate

The eligibility rate is given by:

$$\bar{e} = \frac{\sum t_i e_i}{\sum t_i}, = (\text{Box 8}) / (\text{Box 4})$$

Wiseman and McDonald (1980) used the term *incidence rate* but applied the term only to selected persons of telephone samples that actually answered (responded) at the screening phase to determine their eligibility for the survey.

The eligibility rate, demonstrates the quality of the survey design in selecting eligible units from a frame, where the eligibility may not be readily determinable without some cursory contact or observation. The rate provides, at the screening stage, information to determine how many eligible units will result at the survey data gathering stage. Thus, the rate may be employed at the design stage if data on eligibility are available from earlier studies. Depending upon the nature and procedure of the survey, the eligibility of units may not be determinable among non-contact

or even among units that refused. There are two alternatives to the definition of eligibility rate and response rates (which will be defined later) pertaining to eligible units. One can assume, for conservative estimates of data quality and the quality of the procedure for gathering survey data that all non-contacts and refusals would be eligible even though realistically the proportion of eligible units among such nonrespondents is often lower than among respondents and non-respondents for which the eligibility criteria are known. Under the above assumption a lower bound for the response rate and an upper bound for eligibility rate would be obtained. Alternatively, one can assume the same proportion of eligible units among units whose eligibility cannot be determined as among those whose eligibility are known. Under that assumption we would likely have a slight over-estimate of eligibility rate and some of the other rates.

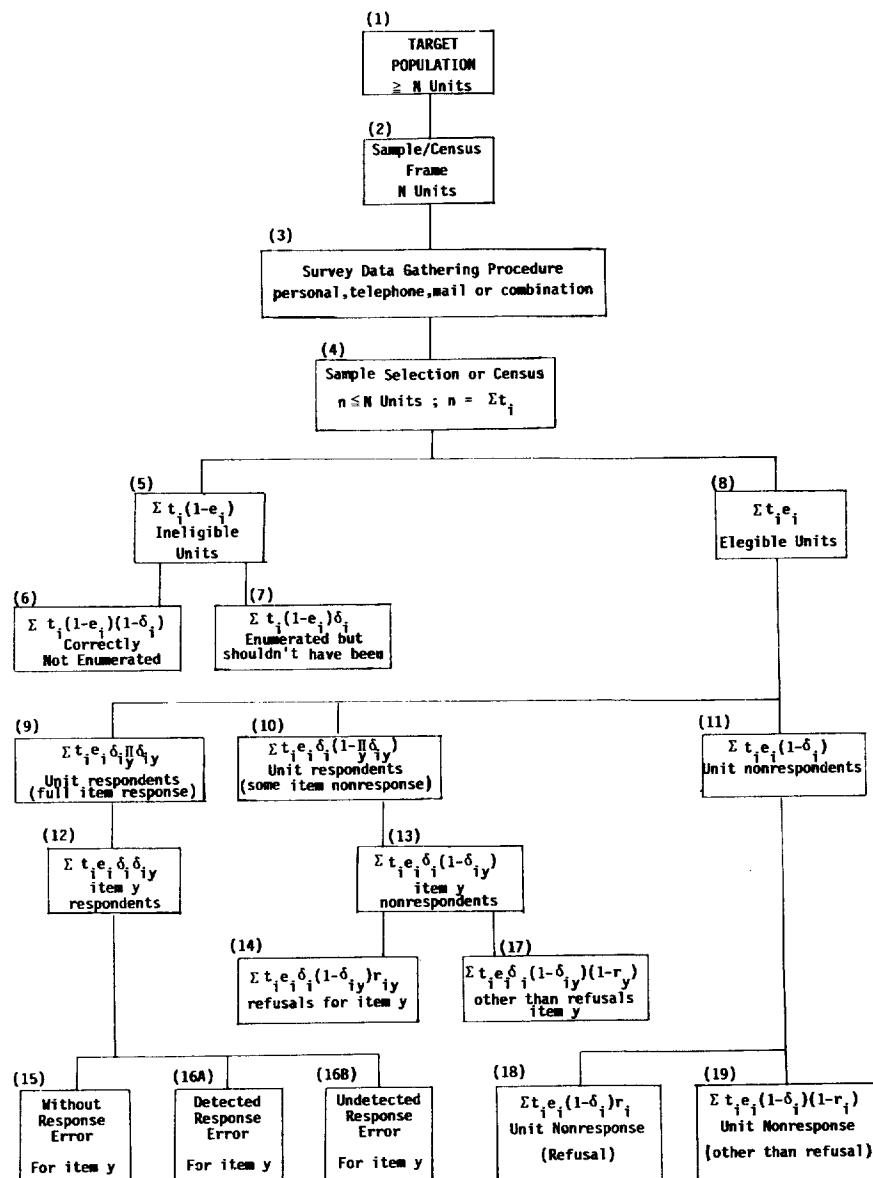
(ii) Response and Completion Rates

a) According to one of the two alternative definitions provided by the U.S. Federal Committee on Statistical Methodology (1978), the response rate is the percentage of the eligible sample for which information (survey data) is obtained. Thus the response rate is defined as:

$$R_{(1)} = \frac{\sum t_i e_i \delta_i}{\sum t_i e_i} \\ = [(\text{Box 9}) + (\text{Box 10})] / (\text{Box 8})$$

The above is the most commonly employed response rate in practice as it yields the percent of the sample for which some useful survey data are obtained once the ineligible units are deleted. All types of non-respondents of eligible units are included in the denominator. The inverse of the above rate at an adjustment cell is frequently used as a weight adjustment to compensate for missing data of nonresponding units, for example, such rates are frequently used in the Canadian LFS for weight adjustments (see Platek and Gray 1985). The above rate or its complement the nonresponse rate, is frequently used for administrative and operational assessments of survey organizations. The rates are also used to assess interviewer's ability to contact respondents and to elect this co-operation to provide usable survey data, e.g., response/nonresponse rates by interview assignment. The nonresponse rate includes both refusals, which may be controlled by good public relations and diplomacy, and non-contacts, which may be beyond the control of the interviewer. Hence, whenever possible, the nonresponse rates are frequently split up by reasons.

TABLE 3
RESPONSE/NONRESPONSE COMPONENTS



$$e_i = 1,0 \text{ (unit eligible/ineligible)}$$

$$\delta_{iy} = 1,0 \text{ (item } y \text{ response/nonresponse)}$$

$$t_i = 1,0 \text{ (selected/not selected)}$$

$$r_i = 1,0 \text{ according as unit refused or not}$$

$$\delta_i = 1,0 \text{ (unit response/nonresponse)}$$

$$\text{For } r_i = 0, \text{ mainly "Not at Home" or "Temporarily Absent"}$$

A similar rate to the above was defined as a completion rate by Kviz (1977), who included the whole sample in the denominator. Such a rate may provide a more conservative estimate of quality in that eligible units such as vacants are included in the denominator.

Another definition which is given below is the percentage of times an interviewer obtains interviews at sample addresses, where contacts are made.

$$R_{(2)} = \sum_i t_i \delta_i / \sum_i t_i [\delta_i + (1 - \delta_i)r_i]$$

where unit i refused or did not refuse according as $r_i = 1$ or 0 respectively. The above was defined as a completion rate by O'Neill, Groves and Cannel (1979). If the eligibility of all units that are contacted can be determined, then another and perhaps superior definition of the above rate pertaining to eligible units can be given by

$$R_{(3)} = \sum_i t_i \delta_i e_i / \sum_i t_i e_i [\delta_i + (1 - \delta_i)r_i] \\ = [(Box\ 9) + (Box\ 10)] / [(Box\ 9) + (Box\ 10) + (Box\ 18)]$$

where e_i , the eligibility criterion is defined after Table 1.

The above rates may be useful in personal and telephone surveys where nonrespondents may include non-contacts and refusals. The rates are not practical in mail surveys unless there is a telephone or a personal follow-up on nonrespondents since in most pure mail surveys, the survey organization is faced with either response or nonresponse with unknown reasons. Where the above rates are meaningful, however, they measure the ability of a data collection method to elicit co-operation of responsible respondents at selected units, given that they are contacted. The non-contacts that may be beyond the control of interviewers and in some survey procedures are removed from the rates entirely.

The response rate was also defined as completion rates by Klecka and Tuchfarber (1979), who assumed, perhaps unrealistically, that all refusals were eligible for the survey. The completion rate would then have been a conservative estimate for the measure of performance of the data collection method in eliciting the co-operation of eligible units. Alternatively, one may assume the eligibility among refusals to be the

same proportion among refusals as among completed and other units whose eligibility criteria is known.

(iii) Contact Rates

A contact rate, defined by Hauck (1974) is the percentage of sample units that are contacted and is given by:

$$R_{(4)} = \frac{\text{Completed interviews} + \text{Refusals (contacted)}}{\text{Completed interv.} + \text{Refusals (cont. + Noncont.)}}$$

where the Noncontacts were assumed to be eligible for a conservative estimate of the success in contacting sampled units. The Refusals may include Terminations or Incomplete Interviews that are essentially Refusals for some items as in (Box 10) of Table 3.

The algebraic expression for the contact rate is given by:

$$R_{(4)} = \frac{\sum_i t_i \delta_i e_i + \sum_i t_i (1 - \delta_i) r_i \hat{e}_i}{\sum_i t_i \delta_i \hat{e}_i + t_i (1 - \delta_i) r_i \hat{e}_i + \sum_i t_i (1 - \delta_i) (1 - r_i) \hat{e}_i}$$

$$= \frac{(Box\ 9) + (Box\ 10) + (Box\ 18)}{(Box\ 9) + (Box\ 10) + (Box\ 18) + (Box\ 19)}, \text{ where}$$

$\hat{e}_i = e_i = 1$ or 0 if eligibility criterion is known, and, for non-contacts,

$\hat{e}_i = 1$ according to Hauck definition,

$\hat{e}_i = \bar{e}$, the average eligibility rate among those units whose eligibility criteria are known.

The contact rate measures the ability of the survey organization or interviewers to contact respondents whether or not they succeeded in eliciting their co-operation.

(iv) Refusal Rate (Non-refusal Rate)

Two definitions of refusal rates are given by Hauck (1974) and Wiseman and McDonald (1980) respectively as:

$$F_1 = \frac{\text{number of refusals}}{\text{number of completed interviews and refusals}}$$

$$= \sum_i t_i \hat{e}_i (1 - \delta_i) / [\sum_i t_i \delta_i + \sum_i t_i \hat{e}_i (1 - \delta_i) r_i]$$

$$= (\text{Box 18}) / [(\text{Box 9}) + (\text{Box 10}) + (\text{Box 18})] = 1 - R_{(3)}$$

and

$$F_2 = \frac{\text{number of refusals}}{\text{number of all selected units}}$$

$$= \sum_i t_i (1 - \delta_i) r_i / \sum_i t_i$$

$$= (\text{Box 18}) / (\text{Box 4})$$

With the eligibility criteria taken into account, the refusal rate may be given by:

$$F_3 = \sum_i t_i \hat{e}_i (1 - \delta_i) / \sum_i t_i \hat{e}_i$$

$$= (\text{Box 18}) / (\text{Box 8})$$

The refusal rate measures the extent of the inability of the survey organization or the interviewer to elicit the co-operation of units to provide usable survey data, relative to the contacted units, relative to the whole sample or relative to the eligible sample. One may also wish to determine a pure refusal rate without non-contacts that are often beyond the interviewers' control in order to study the efficiency of a questionnaire or effect of the survey topic on the co-operation of contacted units. Alternatively, one may prefer to examine the refusals rate as one of several components of overall nonresponse.

(v) Item Response/Nonresponse Rates

Complex questionnaire design may result in item nonresponse of specific questions for reasons other than refusals, as noted in Box 17. A controversial or personal question or termination of the interview may result in a refusal to provide data for a specific item as in (Box 14).

Thus, one may measure the overall item nonresponse rate for item y , relative to all responding units, given by:

$$R_y = \frac{(\text{Box 13})}{(\text{Box 9}) + (\text{Box 10})}$$

or if item y is relevant only for some units (questionnaires) but not for all of them, one may measure the item nonresponse relative to only those responding units for which item y is relevant (eligible). Consequently, one may define a whole set of item response/non-response/eligibility rates, analogous to the unit rates replacing in the rates the number of units (eligible/ineligible)/(responding/refusing, etc.) with the number of responding units (eligible or relevant for item y , irrelevant, responding for item y /refusing for item y etc.) respectively. Most of the rates pertaining to units other than contact rates should have their item y counterparts readily defined by making the proper substitutions in the expressions. However, it may be more difficult to record the reasons for item nonresponse, compared with unit nonresponse, as frequently the item nonresponse is detected only through an edit and imputation routine.

(vi) Weighted Rates and Characteristic Rates

In the case of sample with different sample weights n_i^{-1} 's for the units as in probability proportional to size (pps) sampling, all of the above rates may be defined as weighted rates by applying the sample weight n_i^{-1} with the sample selection indicator variable t_i in all the expressions. In the case of self-weighting samples in an area or class for which the rates are calculated the sample weights are redundant. In pps sampling at the final stage, however, the usual tendency is for large units to respond more readily than small ones so that weighted response rates, with smaller sample weights applied to the large units than for small units, tend to be smaller than unweighted rates based on the counts of units as in Table 3.

The weighted response rates estimate the proportion of the population that would have responded to the survey under similar survey conditions, while the unweighted response rates provide a measure of data collection performance only for the sample or sub-sample pertaining to a specified area or class.

By estimating the nonresponse rate for the entire population rather than for the sample as the unweighted rates do, the weighted rate may provide misleading information on the quality of the data since it may distort the distribution of characteristics in the sample. The advantage of the weighted rates, however, is that the units are added to population levels rather than sample levels so that one obtains an estimate of the rate that would prevail at census levels under similar conditions of gathering survey data.

(vii) Size of Response/Nonresponse Rates

Nonresponse rates vary from survey to survey. Some surveys have nonresponse rates above 50% and others only 4% to 6%. It is important to realize that it is the purpose of the survey that determines whether nonresponse rates are too high or too low. If the objective of a survey is to estimate a 10% item in the population, then nonresponse rate of 5% would affect it considerably. On the other hand, one can think of a situation when high nonresponse is not necessarily useless. Like all statistics, response rates are subject to sampling and nonsampling variances and under the response probability model (Platek, Gray 1983) the expression for the variance of the rate may be derived.

An approximation of the variance of a response rate R , according to any definition, may be given by:

$$V(\hat{R}) = \frac{\hat{R}(1-\hat{R})}{n} F$$

where $R = E_1 E_2(\hat{R})$; and E defines expectation

E_1 = expected value taken over all possible samples

E_2 = expected value over missingness patterns of selected units

n = sample size

F = design effect

"F" may be taken between 1 and 2 depending upon survey design and interview procedure. Thus, "F" could be close to 1 for mail surveys based on a simple random sample. For telephone surveys it may be close to 1.5 due to clustering effect of interviewer assignment. For personal interviews it may be about 2 due to combined effect of design clusters and interview procedure.

III. NONRESPONSE IN THE LABOUR FORCE SURVEY

Even though the Canadian Labour Force Survey is a well established and controlled survey, nonresponse is a continuing concern. A detailed record is kept of total nonresponse which may be broken down into a number of components, each of which have different causes and requires a different treatment. One can recognize the following components: (1) households temporarily absent, (ii) no one at home (iii) refusal (iv) no interviewer available (v) bad weather conditions and (vi) other miscellaneous (e.g. language, illness).

(i) Behaviour of Nonresponse over time

In Table 4, a few trends in the behaviour of nonresponse rates in the LFS should be pointed out. The overall nonresponse rate always increases sharply during the

months of July and to a lesser degree in August. This is mostly due to respondents being *Temporarily absent* on vacation. The size of nonresponse is mainly determined by *Temporarily absent* and *No one at home*. The refusal rates have been fairly steady over a number of years with some downward trend recently and they appear to be slightly higher in summer months. One significant trend is that the Revised Labour Force Survey in 1975 had a relatively high nonresponse rate. This was probably due to the hiring of an almost completely new staff of interviewers along with heavy burdens on the field supervisors due to new procedures, new samples and very heavy training loads. As the survey settled down and supervisors were able to devote more attention to response rate, the nonresponse was reduced to more acceptable levels.

TABLE 4
LFS NONRESPONSE RATES BY COMPONENT AT THE CANADA LEVEL

YEAR	MONTH	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Overall													
	1973	7.3	7.2	6.8	7.9	7.0	8.4	15.1	10.9	6.5	5.7	5.2	6.6
A	1974	6.0	6.0	6.4	8.3	7.0	6.8	10.4	8.8	5.6	5.5	4.3	4.6
-	<u>1975</u>	4.3	4.7	4.6	4.7	4.7	5.8	7.6	6.3	4.3	4.5	4.3	5.3
B	1976	10.2	9.3	8.7	7.5	7.9	8.9	13.2	11.0	7.5	7.5	7.0	7.6
	1973	7.9	7.7	7.3	8.8	9.2	8.0	11.7	9.2	5.8	5.9	6.0	5.3
Temporarily Absent													
	1973	1.8	2.2	1.9	2.4	1.8	3.8	9.1	5.6	1.6	1.3	1.2	1.7
A	1974	1.7	1.8	1.9	2.0	1.5	2.0	6.1	4.7	2.0	1.7	1.0	1.4
-	<u>1975</u>	1.4	1.6	1.6	1.2	1.2	2.2	4.2	3.0	1.1	0.9	0.7	1.2
B	1976	2.0	2.2	2.1	1.5	1.6	2.6	7.1	5.0	1.7	1.4	1.2	1.4
	1973	1.9	2.1	2.3	2.3	1.8	2.4	6.2	4.1	1.7	1.4	1.3	1.3
No one at Home													
	1973	2.5	2.1	2.0	2.6	2.5	2.7	3.2	2.3	2.1	1.9	1.6	2.0
A	1974	1.5	1.7	1.8	2.8	1.9	1.8	1.7	1.7	1.4	1.7	1.4	1.2
-	<u>1975</u>	1.0	0.9	1.0	1.2	1.1	1.3	1.2	1.2	1.1	1.4	1.2	1.6
B	1976	2.7	2.2	1.8	1.5	1.9	1.8	2.2	2.1	1.8	1.7	1.7	1.8
	1973	2.2	1.9	1.5	2.8	2.8	2.4	2.4	2.2	1.7	2.0	2.2	1.8
Refusal													
	1973	1.7	1.9	1.9	2.0	2.0	1.9	1.9	2.3	2.1	2.0	1.9	1.7
A	1974	1.6	1.6	1.7	2.1	2.4	2.3	2.1	1.9	1.6	1.4	1.3	1.2
-	<u>1975</u>	1.2	1.2	1.2	1.4	1.6	1.4	1.4	1.3	1.3	1.2	1.4	1.3
B	1976	2.1	2.2	2.1	2.0	1.9	1.8	1.7	1.6	1.6	1.5	1.5	1.4
	1973	1.5	1.6	1.6	1.9	2.5	2.4	2.0	1.9	1.6	1.6	1.7	1.6
Other													
	1973	1.3	1.0	1.0	0.9	0.7	0.5	0.9	0.7	0.7	0.5	0.5	1.2
A	1974	1.2	1.0	0.9	1.4	1.2	0.7	0.5	0.5	0.6	0.7	0.6	0.8
-	<u>1975</u>	0.7	1.0	0.8	0.9	0.8	0.9	0.8	0.8	0.8	1.0	1.0	1.2
B	1976	3.4	2.7	2.7	2.5	2.5	2.5	2.2	2.3	2.4	2.9	2.6	3.0
	1973	2.3	2.1	1.9	1.8	2.1	1.6	1.1	1.0	0.8	0.9	0.8	0.6

A rates are those rates taken from the old Labour Force Survey

B rates are those rates taken from the new Labour Force Survey

Nonresponse rate is calculated as the percentage of nonrespondent households out of all sampled households

TABLE 5
LFS NONRESPONSE RATES BY COMPONENT AT THE CANADA LEVEL

YEAR	MONTH	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Overall													
1977		6.1	5.9	5.8	4.5	5.8	5.7	8.4	5.6	4.6	4.3	4.3	4.2
1978		5.0	5.2	5.7	5.3	5.7	5.1	7.9	5.9	4.7	4.6	5.0	4.5
1979		5.4	5.8	5.8	5.1	5.1	5.2	7.5	5.9	4.9	4.5	4.2	4.8
1980		5.3	5.4	5.2	5.3	5.9	6.1	7.6	5.9	4.7	4.3	4.2	4.4
Temporarily Absent													
1977		1.9	1.9	2.2	1.4	1.9	2.3	5.1	2.5	1.5	1.3	1.1	1.1
1978		1.7	2.1	2.5	1.7	1.7	1.9	4.6	2.9	1.6	1.4	1.2	1.2
1979		1.8	2.1	2.1	1.5	1.4	1.7	4.1	2.6	1.6	1.2	1.1	1.5
1980		1.6	1.7	1.7	1.7	1.6	2.2	4.0	2.7	1.5	1.2	1.0	1.3
No one at Home													
1977		1.8	1.9	1.6	1.4	2.0	1.7	1.6	1.4	1.4	1.4	1.3	1.2
1978		1.4	1.3	1.4	1.5	1.8	1.4	1.4	1.3	1.4	1.5	1.8	1.5
1979		1.6	1.6	1.5	1.5	1.7	1.6	1.6	1.6	1.7	1.8	1.6	1.6
1980		1.8	1.5	1.5	1.7	1.9	1.7	1.6	1.4	1.4	1.5	1.5	1.4
Refusal													
1977		1.6	1.5	1.4	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
1978		1.7	1.3	1.4	1.6	1.7	1.5	1.5	1.4	1.4	1.4	1.4	1.4
1979		1.4	1.5	1.5	1.5	1.6	1.5	1.5	1.3	1.3	1.2	1.2	1.3
1980		1.3	1.5	1.4	1.5	1.9	1.7	1.6	1.5	1.3	1.3	1.3	1.4
Other													
1977		0.8	0.6	0.6	0.4	0.6	0.4	0.4	0.4	0.4	0.3	0.6	0.5
1978		0.6	0.5	0.4	0.5	0.5	0.3	0.4	0.3	0.3	0.3	0.6	0.4
1979		0.6	0.6	0.7	0.6	0.4	0.4	0.3	0.4	0.3	0.3	0.3	0.4
1980		0.6	0.7	0.6	0.5	0.5	0.5	0.4	0.3	0.5	0.3	0.4	0.3

(ii) Nonresponse and Tenure in the Survey

Table 6 provides an example of nonresponse rates

in the LFS according to number of times the household has been in the survey. This is followed by a brief discussion of various components of total nonresponse rates.

TABLE 6
NONRESPONSE RATES (%) ACCORDING TO TENURE OF HOUSEHOLD
IN THE LFS (1979-1980)

Number of Months in Survey	Nonresponse Rates (%)			
	Total Nonresponse	Refusal	No one at home	Temporarily Absent
1	8.04	1.43	2.96	2.94
2	5.09	1.21	1.44	1.99
3	4.71	1.32	1.10	1.90
4	4.65	1.46	1.09	1.79
5	4.62	1.51	0.99	1.77
6	4.45	1.52	0.78	1.73

On the basis of the results shown in Table 6 the following comments can be made:

- (a) The total nonresponse rate was highest during the first month, clearly because interviewers had more difficulty in finding people at home having not yet determined the best time to call as one may observe in the higher No one at home rate, for example. The rate then decreased sharply in the second month and continued to decrease through the third and fourth months.
- (b) The Refusal rate decreased in the second month, increased gradually through the third, fourth and fifth months and levelled off in the sixth month. The initial decrease probably represents the effect of supervisory intervention; the subsequent increase is likely due to the effect of cumulative response burden.
- (c) The No one at home rate decreased sharply from the first month to the second month by roughly 50 percent. It continued to decrease from the second month to the third month but decreased very gradually through the fourth and fifth months. A larger decrease then occurred in the sixth month. The behaviour of the No one at home rate over the six month tenure of households in the survey is most probably due to the fact that the longer a household is in the survey the more familiar the interviewer becomes with knowing when the respondent is most likely to be at home.
- (d) The Temporarily absent rate decreased through all six months, particularly from the first to second month. The Temporarily absent rate should not be expected to depend on how long a household remains in the survey, and it seems almost certain that interviewers may have confused No one at home and Temporarily absent types of nonresponse.

TABLE 7
NONRESPONSE RATES (%) BY TYPE OF AREAS (MONTHLY AVERAGE: 1985)

Type of Area	Proportion of sample	Total Nonresponse	Refusal	No one at home	Temporarily absent
NSRU ¹	48.03 %	4.9 %	1.2 %	1.4 %	1.8 %
- urban ²	17.58 %	5.2 %	1.1 %	1.4 %	2.2 %
- rural ²	30.45 %	4.8 %	1.3 %	1.4 %	1.6 %
SRU ³	51.08 %	5.7 %	1.8 %	1.8 %	1.8 %
- built-up ⁴	36.34 %	5.6 %	1.7 %	1.7 %	1.8 %
- fringe ⁴	10.47 %	4.8 %	1.6 %	1.2 %	1.6 %
- apartment ⁵	4.27 %	9.9 %	2.5 %	4.0 %	2.6 %

¹ Non-Self Representing Units are the areas outside SRU's and contain rural and small urban centres.

² Every primary sampling unit in an NSRU is divided into an urban and a rural portion.

³ Self-Representing Units are cities whose population exceeds 15,000 persons or whose unique characteristics demanded their establishment as SRU's. Every SRU is selected with certainty.

⁴ SRU's are stratified into sub-units and sub-units are classified as built-up or fringe on the basis of their potential for future growth. Generally speaking, SRU fringe households belong to the fringe or sub-urban areas.

⁵ In seventeen large cities across Canada there is a separate frame of apartments in buildings having at least five storeys and thirty or more units.

Within NSRU's the total nonresponse rate was higher in the urban portion due to higher temporarily absent rates among NSRU urban households. The No one at home rates in the urban and rural portions were the same, but the Refusal rates were 20 percent higher in NSRU rural areas than in NSRU urban areas.

Within SRU's built-up areas had a higher total nonresponse rate than fringe areas due to higher No one at home and Temporarily absent components.

Thus, it appears that people living in the core areas of cities tend to be more difficult to contact than people living in the fringe areas; the differences, however, were not large.

SRU apartments had a higher total nonresponse rate than any other area shown in Table 7. In fact, the total nonresponse rate in the SRU apartment sample

was almost twice the rate in the SRU non-apartment sample (consisting of both built-up and fringe areas). The Refusal, No one at home and Temporarily absent components were also highest among apartments.

The No one at home rate was almost three times higher in the apartment sample than in the non-apartment sample. This large difference may be due to the different life-styles of apartment and non-apartment dwellers. Apartment households usually consist of single persons or very small families who tend to be more mobile and difficult to find at home, while non-apartment households are more likely to contain larger families with children. Another problem is that interviewers often find it difficult to gain entrance into apartment buildings.

Whereas the Temporarily absent rate was usually twice as high for apartments as non-apartments, the difference was less noticeable during July and August

than in the other ten months. This probably resulted from the fact that it is easier for single persons and families without children to take their vacations during the fall, winter and spring than it is for families with school-age children.

The Refusal rate was almost always higher in the apartment sample than in the non-apartment sample, although the difference in the Refusal rates between the two samples was not as great as the differences observed for the No one at home and Temporarily absent rates. Recent results, however indicate that the gap in the level of Refusal rates between apartments and non-apartments is gradually widening, to the extent that the Refusal rate in the apartment sample is now almost double the corresponding rate in the non-apartment sample.

The significance of examining nonresponse rates according to breakdowns such as SRU and NSRU is that this approach helps establish relationships among the various types of area in items of the behaviour of

nonresponse rates. For instance, the overall nonresponse rate is always expected to be higher in SRU's than in NSRU's, and any deviation from this relationship is considered unusual. The same holds true for the No one at home and Refusal rates in SRU's and NSRU's. Another example is the two-to-one ratio of the total nonresponse rate in the apartment sample to the corresponding rate in the non-apartment sample. If nonresponse rates ever increase beyond average or expected levels, then knowledge of these realtionships is useful for the purpose of analyzing the situation and taking remedial action.

(iii) Incidence of Nonresponse for Interviewer Assignments

As expected, nonresponse rates varied among interviewers. Many interviewers, in fact, achieved 100 percent response rates, while a few interviewers did no better than 75 percent. It is interesting to look at the distribution of interviewers according to the level of their total nonresponse rates as shown on Table 8.

TABLE 8
DISTRIBUTION OF INTERVIEWERS ACCORDING TO NONRESPONSE RATES
(MONTHLY AVERAGE: 1980)

Total Nonresponse rate (%)	Number of Interviewers	Percentage of Total Interviewers
0.0	159	15.0
0.1 to 5.0	434	40.8
5.1 to 10.0	333	31.3
10.0 to 15.0	98	9.2
15.1 to 20.0	29	2.7
Over 20.0	10	1.2
TOTAL	1,063	100.0

The data in Table 8 are based on all interviewers who enumerated assignments with at least 20 households and represent an average over the twelve months of 1978. The Table indicates that 56 percent of interviewers achieved nonresponse rates of 5.0 or better.

The 13 percent of interviewers with nonresponse rates higher than 10.0 accounted for 31 percent or almost one-third of all nonresponses. It is also interes-

ting to observe that 59 percent of interviewers did not record any No one at home nonresponses and that 54 percent encountered no refusals. Furthermore, 71 percent of interviewers achieved no one at home rates and refusal rates respectively of 2.0 or better. These percentages reflect the success of the on-going training programs, monitoring and controls as well as interviewing techniques and procedures which are all aimed at maximizing response levels in the LFS.

IV. NONRESPONSE BY TYPE OF SURVEYS

The major factors influencing the nonresponse rates are the sensitivity and complexity of subject matter (e.g. income, victimization, information on education and nutrition, the latter imposing an additional respondent burden by requiring clinical tests). It has been our experience that the effort of the LFS to increase response rates also benefits supplementary surveys. However, a combined effect of sensitive questions, complexity and the length of questionnaire may impose too much burden on the respondent resulting in a higher nonresponse, even though the survey may be a supplement to the LFS. This is, in fact, the case with the Survey of Consumer Finance (Income). However, during the years when SCF is an independent sample, the burden is at least a *one-shot affair* and the respondent may cooperate resulting in lower nonresponse rates than when it is

a supplement.

A second but also important fact is the method of data collection. Usually personal interviews will have lower nonresponse rates than any other method of interviewing for approximately the same degree of sensitivity of the subject matter. For example, our Travel Surveys, when conducted using personal interview, resulted in a lower nonresponse rate (7.3%) than using telephone (10%) or drop off-mailback (16%) methods of data collection. Similarly, the survey on smoking habits had nonresponse rate of 8% for personal and 10% for telephone interviewing respectively. In the latter case it should be noticed that both surveys were conducted as supplements to the LFS.

TABLE 9
NONRESPONSE RATE BY TYPE OF SURVEY

Name of Survey	Year	Collection Method	Nonresponse (1) Rate	Supplement to LFS
Survey of Consumer Finance (Frame)	1972	Drop-off-Pick-up	28.3	Yes
" "	1973	"	18.1	No
" "	1974	"	25.7	Yes
" "	1975	"	20.2	No
" "	1976	"	27.1	Yes
" "	1977	Personal	20.3	No
" "	1978	Drop-off-pick-up	34.4	Yes
" "	1979	"	28.1	No
Selected Leisure Activities	1978	"	13.0	Yes
Nutrition Survey	1972	Personal	50.0(2)	No
Household Facilities	1978	"	9.1	Yes
" "	1979	"	11.5	Yes
" "	1980	"	15.9	Yes
Travel Survey	1979	"	7.3	Yes
Smoking Habits	1979	"	8.0	Yes
Study of Housing	1979	"	20.0	No
Travel Survey	1979	Telephone	10.0	Yes
Smoking Habits	1979	"	10.0	Yes
Victimization	1979	"	15.0	No
1976 Graduate	1979	"	14.5	No
Travel Survey (1977)	1977	Drop-off-mailback	16.0	Yes
Methodology Test	1977	Drop-in-mail-back	18.0	Yes

¹Non response rates = $\frac{n_2}{N} \times 100$ where N is total sample and n_2 is the number of units which responded.

²include 25% who refused to come to clinic

V. THE ROLE OF NONRESPONSE IN SURVEYS

Although, the measurement of response/nonresponse, as defined by various rates is rather simple, nevertheless, the importance, interpretation and control of them require some discussion.

Let us look at the role of nonresponse from the point of view of such survey operations as design, data collection, analysis and management of surveys.

(i) Design

At the planning stage it is important to make a decision as to the tolerance level of nonresponse in the sample. It can be argued that for surveys when only national estimates are required and if the respondent's and nonrespondent's characteristics are fairly similar a nonresponse rate (20%-30%) may be tolerated. In this case the bias would be relatively small though there will be some increase in variance. The same argument would apply if we were to estimate trends and preparations. However, if the estimates have to be precise and are also required at various sub-national levels, then the size of nonresponse should be kept as low as 5% to 7% and one should watch for pockets of high nonresponse in local areas.

Another important role that response/nonresponse play at this stage is the cost of survey. It is important to allocate the cost to various factors in such a way as to achieve nonresponse sufficiently low to serve the goals of the survey. It is frequently better to accept a smaller sample than originally planned and to allocate more money to data collection and a follow-up. This would be particularly appropriate if large differences are expected between respondents. Most of the above, experienced survey designers can estimate fairly accurately. In addition and in a more formal way survey designers may identify a number of important factors in survey design which will also affect nonresponse.

It is fairly self evident that the following factors will affect nonresponse: sample frame, method of interviewing, selection, training and control of staff, length of questionnaire and wording, sensitivity of questions, type of area, feasibility and the number of call backs, and finally publicity. Ideally, a survey designer would like to conduct a pilot study or studies to test the importance of at least some of these factors. Short of experimentation, a thorough discussion based on past experience and intuition is essential to arrive

at a final design.

In survey planning and development a number of factors should be taken into account in arriving at the final design. These factors can be classified into three groups:

Group I: These factors pertain to the sample design stage, i.e., factors which have an indirect effect on response rates.

- a) sample frame
- b) sample size
- c) stratification
- d) sample allocation by strata
- e) sampling procedure within strata
- f) sample allocation by stage, especially clustering of sample

Group II: The factors pertain to questionnaire design and data collection procedures, and public relations, items which may directly affect nonresponse rates.

- a) listing of units for sample selection
- b) subject matter and type of survey
- c) interviewing procedure
- d) questionnaire length and complexity (wording)
- e) sensitivity of questions to respondents and to interviewers
- f) selection, framing and control of field staff, including interviewers
- g) type of area for survey
- h) feasibility and cost of call-backs, follow-up procedures
- i) publicity/media

Group III: These factors pertain to the processing and analysis stage after the survey has been completed and there is no longer any opportunity to deal with missing data in the field.

- a) edit and imputation
- b) estimation procedures, including weight adjustments in cells
- c) variance and variance estimation
- d) publication and data analysis

All of these factors are inter-related and those of Groups I and II at least have a bearing on the incidence of nonresponse, which in turn affects the

mean square error of estimates.

On the basis of an error model for survey data which includes sampling response, and imputation error components, the following expression for the mean square error (MSE) of an estimate from a survey, based on imputation procedures, may be obtained.

$$\text{MSE} = V + B^2$$

$$V = SV + SRV + CRV + VRR + CVRR$$

$$B = SB + RB + IB$$

where SV = Sampling variance

SRV = Simple response variance

CRV = Correlated response variance

VRR = Variance arising from the variance of response/nonresponse status of selected units

$CVRR$ = Covariance arising from the covariance between response/nonresponse statuses of pairs of selected units.

B = bias

SB = Sampling bias

RB = Response bias

IB = Nonresponse or imputation bias.

The sampling variance SV is affected by all the factors of Group I and Group III (a) and (b) and most likely increases with the size of nonresponse, after the application of the edit and imputation procedures.

The sampling bias SB is mainly affected by two factors in group I, including (a) (incomplete sample frame) and (e), sampling procedures, within strata if non-probability or quota sampling is undertaken. It is also affected by a couple of factors in Group II, i.e., (a) (incomplete listing of units), (f), pertaining to the case of poorly trained field staff and interviewers who may not interview units as they have been designated in the sample. The sampling bias, however, is not likely to be affected by an increase in the nonresponse rates.

It is important to note that each survey determines its own requirements with respect to design, questionnaire and interviewing procedures. As far as the design is concerned, an important factor that results in non-response are factors (e) and (f), of Group I the sampling procedure within strata and the sample allocation by stages (if there are stages). For example, an unclustered sample may produce a higher nonresponse rate than a clustered design. This may be due to the requirements for extensive travelling in the case of personal interviewing where for reasons of cost, repeated callbacks must be restricted.

The questionnaire design and interviewing procedures and other factors of Group II all have an effect on the non-sampling variance components $SRV + CRV + VRR + CVRR$, on the response bias RB , on the nonresponse rate and consequently on the nonresponse bias IB which may increase with the nonresponse rate, after the application of the edit and imputation procedures. Furthermore all of the non-sampling variance components tend to increase as the responding sample sizes decrease.

Some examples of special cases of factors in Group II which affect directly the nonresponse rate will be mentioned. These include interviewing procedures, questionnaire length and wording, and type of area for survey.

For example it was demonstrated in section 2.5 and Table 3 how interviewing procedures in Group II (c), affect the nonresponse rates.

Group II (d), questionnaire length and wording: The length of the questionnaire not only may affect the nonresponse rate but also may dictate the interviewing procedure to be employed. Long complex questionnaires may result in the need for personal interviewing while mail surveys may suffice for short simple questionnaires. Since the questionnaire is especially important to the survey design, it is discussed more fully in Section 2.10.

Group II (g), type of area for survey: In close knit rural or small urban communities with mostly families, it will be usually easier to find someone at home than in large cities, especially in apartments, with one or two person families who tend to be absent quite frequently.

The above are only a few examples of consideration

of factors that must be taken into account at the planning and development stage of a survey design. Consequently if careful attention is paid to the factors in Group II at the stage, serious nonresponse problems may be avoided.

At the design stage it may be argued that the problems associated with nonresponse cannot be simply resolved by starting with an excess sample to allow for a potential nonresponse, since in the presence of initial nonresponse, the sample is no longer a probability sample. Furthermore, the respondents in some ways and to varying degrees are usually different from those who respond. It is true that if it is assumed that the response probability is the same for all respondents the bias is easily eliminated by adjusting (inflating) the sample weights of the respondents. However, the probability of response may depend on the characteristic of interest and the adjustment of respondent weights to account for the nonresponse will give rise to nonresponse bias. The magnitude of the nonresponse bias will depend on the relationship between the characteristic of interest and the response probability.

(ii) Questionnaire

Integral to a survey development is the questionnaire. Its length, clarity and structure greatly affect the magnitude of nonresponse. The simplest definition of a questionnaire or a schedule is that of a group or sequence of questions designed to elicit information upon a subject from a respondent. Within the range of techniques in interviewing, the questionnaire may range from a list of undefined topics to a highly structured set of questions with no options for response other than those listed.

In most surveys, the questionnaire is an important method of standardizing and controlling the data collection process. Without specific question wordings and instructions to follow, interviewers would inevitably change the meaning or emphasis of questions and quite possibly the responses. The questionnaire plays a central role in a complex process (the interview) in which information is transferred from those who have it (the respondents) to those who need it (the users). The questionnaire is the means through which the information needs of the users are expressed in operational terms which can be presented to a respondent in such a way that he or she will supply the required information. For this transfer of information to be effective, the questionnaire must meet the requirements of both users

and respondents. The questionnaire's total content, style of questioning, format, length, etc. are all closely linked to the method of data collection and the survey's subject matter. Each method of data collection, such as personal interviewing, telephone interviewing and mail surveys creates its own survey conditions and form an important background to which a well designed questionnaire must be sensitive, through an appropriate style of questioning, content, format, length and so on. In personal interviews, for example, it is often possible for the interviewer to collect certain data, such as the type of dwelling or the sex of the respondent, by direct observation rather than by questions. As well, the element of face-to-face communication is a powerful motivating factor for the respondent. A personal interview is often the only choice when a complex, long and demanding questionnaire is involved. In telephone interviews much of the social interaction between interviewer and respondent is lost and the respondent's co-operation may be affected. The questionnaire must rely entirely on verbal communication for its success and the subject matter may have to be less demanding. However, with certain sensitive surveys (e.g. criminal victimization surveys), the extra distance between interviewer and respondent may actually make it easier to answer questions. In mail surveys, the questionnaire itself assumes the role of the interviewer. It must introduce the survey, motivate the respondent to co-operate and guide the respondent in completing the interview.

The questionnaire is an important source of nonresponse errors which can basically occur for either of two reasons: (a) *no contact* (e.g. no one home, temporarily absent, bad weather, etc.) or (b) *refusal*. The latter may be either a unit nonresponse (complete unwillingness to participate in the survey) or an item nonresponse (refusal to answer one or more questions). The questionnaire can do little to eliminate the *no contact* type of nonresponse but it does play an important role in preventing the refusal. Unrealistic demands on the respondent's knowledge or memory, the use of overly difficult and technical language, or excessive demands on the respondent's patience are all sources of nonresponse which have their roots in the questionnaire.

The questionnaire is also an important source of response errors, which can occur because of complexity in the questions and questionnaire design. In such cases, the respondents reply to the questions but provide incorrect answers because of misunderstanding or lack of knowledge of relevant material.

The questionnaire as we have seen is inevitably a cause of nonsampling errors, but it must also go as far as possible in preventing them. The degree to which the questionnaire succeeds at this task depends on its proper design. Each new survey may present new problems and pitfalls and as such they must be anticipated and taken into account in designing the questionnaire.

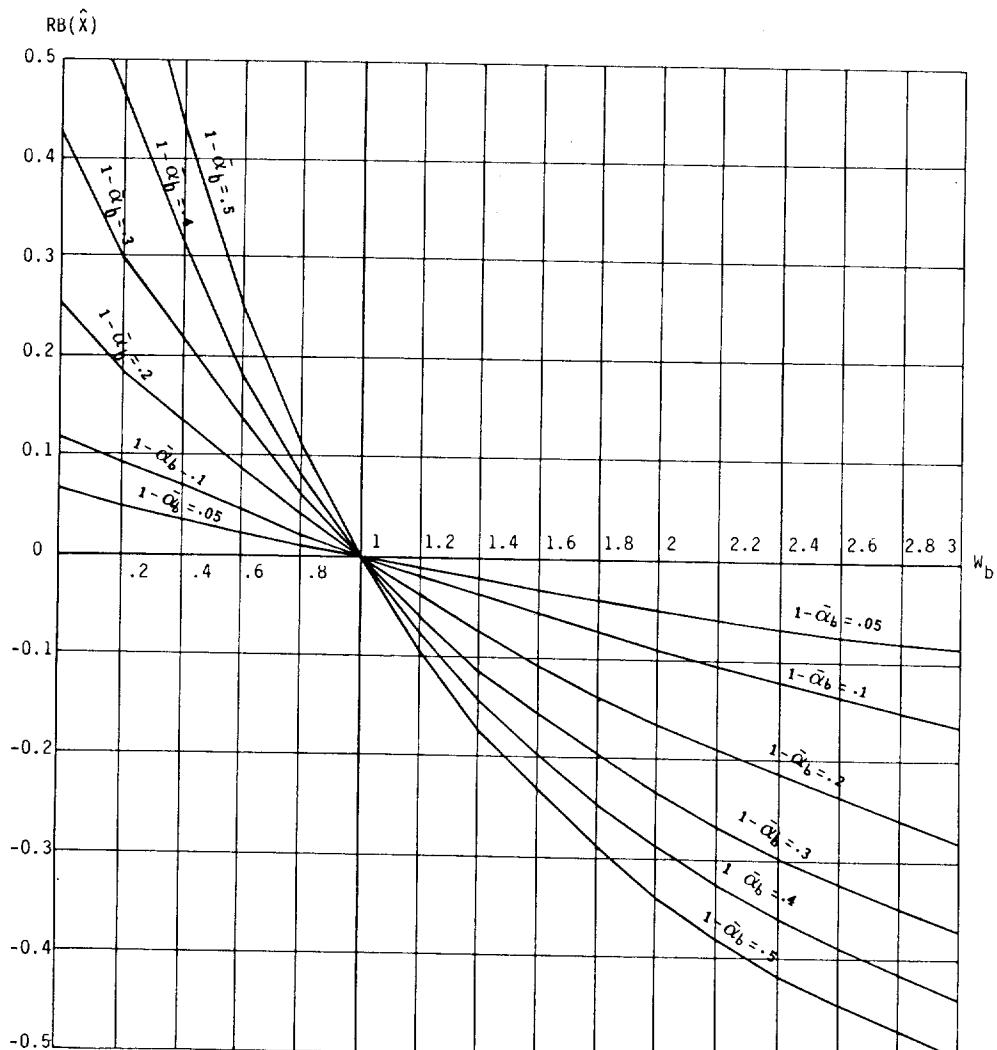
(iii) Relative Bias due to Nonresponse

Having examined the various factors which might possibly affect the size of nonresponse in a survey that is being designed, it can be deduced that a survey

statistician must have a notion of what to expect as far as nonresponse is concerned. Consequently, one of the important factors in planning a survey is a decision on the tolerance level of nonresponse and an experienced statistician can estimate fairly accurately the level of response for a particular survey that can be expected under various survey conditions. To help the survey statisticians to arrive at the tolerance level R. Platek, M.P. Singh and Tremblay in a paper published in a book on Survey Sampling and Measurement by Academic Press develop a formula and a graph which enables one to estimate the upper limit of relative bias for a given survey characteristic and for a given imputation method.

RELATIVE IMPUTATION BIAS $RB(\hat{x})$ OF ESTIMATE BY NONRESPONSE

RATE $(1 - \bar{\alpha}_b)$ & w_b



w_b = NONRESPONDENTS/RESPONDENTS

Let's consider the estimation of the total population X for some characteristic x . The entire population can be classified into respondents and nonrespondent such that

R = number of units that responded

X_R = number of respondent with characteristic x

x_{N_i} = number of nonrespondents with characteristic x_i

N = number of units that did not respond

$X = X_P + X_N$ = total population with characteristic x.

$$T = R + N = \text{number of units}$$

Let's assume that Z_N is imputed for X_N then $\hat{X} = X_R + Z_N$ will be the estimate for X . In general, the bias of X may be expressed as:

$$\text{Bias } (\hat{x}) = (x_R + z_N) - x = x_R + z_N - x_R - x_N = \\ = z_N - x_N = N(\bar{z}_N - \bar{x}_N)$$

a) If there is no adjustment made for nonrespondents, the bias is entirely due to the size of nonrespondent (X_N) i.e. the coverage of the survey is restricted to respondents only.

b) If a correction is applied then

$$B(\hat{X}) = N(\bar{Z}_N - \bar{X}_N)$$

In many practical situations, a survey statistician is not in a position to estimate nonresponse bias. However, it would be highly desirable to establish an upper bound (positive or negative) for the bias.

One can argue, with some justification that when \bar{z}_N is imputed for \bar{x}_N it is hoped that z_N will be as close or closer to \bar{x}_N than \bar{x}_R . Otherwise one would not attempt to impute for \bar{x}_N . Then in the mind of a statistician the following inequality holds:

$$|\bar{z}_N - \bar{x}_N| \leq |\bar{x}_R - \bar{x}_N|$$

In that case the upper limit of the bias may be obtained by $(\bar{x}_R - \bar{x}_N)$. If that was not the case, one would apply the least expensive and the simplest method of compensation for nonresponse (by estimating \bar{x}_N by \bar{x}_D).

In that case the upper limit of bias may be obtained by $(\bar{X}_p - \bar{X}_N)$ and the relative bias may be expressed as

$$RB(\hat{X}) = \frac{\hat{X} - X}{X} = \frac{(1 - \alpha_b) (1 - w_b)}{\alpha_b + w_b (1 - \alpha_b)} = \frac{(1 - \alpha_b) (1 - w_b)}{1 - (1 - \alpha_b)(1 - w_b)}$$

$$\text{where } w_b = \frac{x_N}{x_B} =$$

$1 - \alpha_b$ = Nonresponse rate

b = balancing unit (area)

It is seen from the graph how the relative bias is affected by changes in the response rate. For example, if $1 - \bar{\delta} = 0.2$ and $W_b = 1.5$ then $RB(\hat{X}) = -0.09$ (or 9%). If by some inappropriate procedures $(1-r)$ was reduced to 0.10 but W_b increased to 2 then $RB(\hat{X}) = -0.09$ and the added cost of reducing nonresponse results in no reduction in the nonresponse bias.

Thus, if the value of α_b is obtained from earlier surveys or from follow-up studies, then it would be possible to obtain the magnitude of relative bias with the help of the graph. In other words one can determine the target response rate to be achieved in a given survey so that the bias is within tolerable limits.

(iv) Data Collection

The importance of response/nonresponse rates at data collection stage is derived from their uses for different purposes. One can distinguish between contact and no contact types. One type such as no one at home or temporarily absent is in fact an example of no contact problem and is primarily operationally oriented. The other type is the true nonresponse problem, where contact has been made with respondent but no response or acceptable response is obtained. The no contact type of problem is usually attacked with operational solutions. For example, in a telephone or personal interview, the time and patterns of calling on respondent are important. The size or assignment and the time allotted to data collection must be adequate. Further examination of no contact may reveal that some units, for example households, may be ineligible for the inclusion in the sample. Ineligibility varies with the type of information required.

For example, a vacant household is ineligible for a

survey of all households. But it is eligible in a housing survey. Other households may be mostly apartments occupied usually by younger persons, single and employed. Still other units may have other characteristics. Thus the importance of knowing the origin of nonresponse is that it may determine the strategy appropriate for a particular type of nonresponse.

The problem of refusals is somewhat different. It should be conceded at the outset that refusal rates are not always as straightforward as one might expect. An interviewer may prefer to record a refusal as *no one at home* or a respondent may simply not answer the door as a means of refusing and yet is being recorded as *no one at home*.

In a mail survey one is not always certain whether a respondent has received the questionnaire and having received it simply neglected to mail it. In an interviewer process itself an interviewer may find units that should not be there or units with questionnaire fully or partially completed. Surveys dealing with sensitive subject may not only affect refusals of the survey itself but may have a longer effect on the respondents with respect to other surveys. To avoid pockets of known nonresponse it may be desirable to define survey population in such a very way so that such pockets are eliminated. But one must be aware of the difficulty to make inferences from such a survey population to the characteristics of the target population.

One important role that nonresponse plays at the stage of data collection concerns the size of response in certain situations or areas. For example, if nonresponse rates for an interviewer or by interview areas are produced before the final data collection, they may identify interviewers and areas that need support in order to achieve satisfactory response rates. Also if preliminary tabulations are made they may indicate that the responding sample is too small or too unequally distributed for satisfactory estimates of important population characteristics. A major data collection effort may then be made satisfactory for interviewers areas, and important population characteristics.

Also, the interviewer may carelessly code the response in an incorrect location on the questionnaire resulting in invalid data which must be discarded. The problem of nonresponse seems to be also how to motivate the appropriate respondent to produce a valid response.

With respect to motivation, let us look upon the res-

ponding unit as being neutral towards the survey and consider the influence which may motivate him/her either to respond or not to respond. Such factors as difficulty in understanding questions, use of respondent time, privacy, difference, difficulties in recalling information, embarrassing or personal questions are all examples of motivation not to respond. On the other hand, examples of motivation to respond are: an interest in the survey, willingness to help out, duty, understanding of the importance of survey results, etc.

The problem becomes, how to accentuate the positive motivation and reduce the negative motivation until the balance swings in favour of response. The key element is the respondent and anything which affects his/her ability and motivation to respond must be of interest and concern to a survey designer.

(v) Motivation of Respondent

It is a matter of common experience that in every day life when one asks a question, one normally receives an answer. What then, motivates a respondent not to respond to a survey? Invasion of privacy, respondent burden and general hostility or distrust of government are the three major reasons. In dealing with these problems, it is important to consider them from the respondents' point of view and not from some preconceived notions on the part of the sponsor or survey designer. Certain questionnaires may be perceived by the respondents as burdensome if they do not understand why they are being asked or how the survey relates to them. Under different conditions, however, the same questions may be perceived as very interesting and the respondent is motivated to participate in the survey. The general means by which respondent motivation can be sought may be divided into two parts, public relations and respondent relations.

Public relations are activities directed to the general public, and can take many forms. One long range objective is to create a climate in the general public which will tend to motivate it in the direction of co-operation with surveys. The image of the survey organization as perceived by the respondent can be an important factor in his/her motivation to respond. The extent and manner of data dissemination can be used to impress upon the public the importance of the role of the statistical agency. The statistical agency must be alert to possibilities for favourable comment via the media and also to the necessity for timely and appropriate handling of criticism. A public relations approach, in addition to maintaining a favourable image,

may be used to publicize specific activities and request co-operation in carrying them out. Badly conceived publicity may, in fact have a negative effect. As a general rule, publicity campaigns are most suitable for population census operations where everyone is affected, and there is a need to motivate the population as a whole. Where the ratio of sample size to population is small, it is usually more cost effective (in the context of a single survey) to avoid direct publicity and to concentrate on respondent relations.

(vi) Respondent Relations

A working definition of respondent relations might be that it comprises any action directed toward the individual respondent which may affect his/her attitude and motivation with regard to the survey. As was previously stated, the problems of most concern are invasion of privacy, respondent burden, and general hostility or distrust. In regard to hostility or distrust, public relations and the agency's image have already been mentioned. Identification of the interviewer and sponsor or agency conducting the survey is very important. Unwelcome callers have been known to use the pretext of a survey to establish contact or gain entry. It is important that the interviewer present official identification and that the survey materials convey an appearance of being official.

Introductory letters, examples of the uses of the data, and brochures describing the objectives and authority for the survey, are often excellent means of avoiding hostility and distrust.

Invasion of privacy is usually related to the content of the questionnaire although the reaction of different respondents is quite variable. Many procedures exist for minimizing the effect on the respondent and the specific procedure should be tailored to the specific situation. In some cases, it may be best to allow the respondent to reply in a completely anonymous fashion. This can be accomplished by self-interviewing with no identification whatsoever on the questionnaire. Quite often, though, it is essential to have some area code or sample designation for weighting and estimation purposes and in that, care must be taken that respondent does not perceive this as a means of identifying his/her replies. Of course, it should be clear that whereas anonymous surveys might be beneficial to response from the point of view of privacy concerns, these benefits might be more than offset by the fact that such surveys permit no follow-up of non-respondents.

In addition to the assurance of privacy, some forms of compensating the respondents for their time and effort have been practiced by some survey taking organizations.

(vii) Analysis

When looking at survey results and we find that nonresponse rates are low we feel quite confident about the data from the survey. On the other hand if nonresponse rates are high they are frequently taken as an important evidence of poor quality data. The interpretation of measures of nonresponse are even more difficult when one deals with complex designs since the concentration of nonresponse may be higher in one area than in other. Still response rates have been used as proxies for data quality by almost all survey statisticians. That is why the interest in collecting data on nonresponse and their evaluation has usually been part of survey taking. Yet, by themselves nonresponse may or may not affect data quality. Strictly speaking only the Mean Square Error i.e. bias² and variance can provide an uniformed basis for survey results. But nonresponse can certainly affect its magnitude as can be seen at the Design Stage. Also its affect may depend on the type of statistics that is being analysed. Intuitively, it seems that when we wish to estimate a means of a characteristic the size of the bias due to nonresponse will be determined by both the difference in the characteristics between the respondents and nonrespondents as well as by the size of nonresponse. But, if we have reason to believe, that respondents and nonrespondents do not differ very much than the bias due to nonresponse will be small and independent of the size of nonresponse. In such a situation we may not be concerned about a follow-up of nonrespondents. On the other hand, if the difference is large then the bias will be large, unless the size of nonresponse is very small. In this case a follow-up or some other way of reducing the size of nonresponse should be undertaken. In practice such a clear cut situation never exist. An analyst is usually interested in estimating several means of characteristics. Therefore, a safe procedure is to have a low nonresponse.

The above intuitive reasoning can be easily seen from a simple expression for the mean.

Mean

Suppose

$$R = \text{number of respondents}$$

M = number of nonrespondents
 N = population size

Therefore $N = R + M$

Similarly, r , m and n refer to sample values.

If no compensation is made for nonresponse the respondent mean \bar{y}_r is used to estimate \bar{Y} .

Y is the number of unit having a particular characteristic.

$$E(y_r) = \bar{Y}$$

and the bias $B(\bar{y}_r) = \bar{y}_r - \bar{Y} = \bar{y}_r - (\bar{R}\bar{Y}_r + \bar{M}\bar{Y}_n)$

$$= \bar{M}(\bar{Y}_m) + \bar{Y}_r(1 - \bar{R})$$

$$\text{but } 1 - \bar{R} = \bar{M}$$

$$\bar{M}(\bar{Y}_r - \bar{Y}_m)$$

which demonstrates what was discussed intuitively above for the mean. It also shows the conditions under which \bar{y}_r is an unbiased estimate for \bar{Y} .

Total

If we were going to estimate the total population Y then it seems natural to assume that the bias in the presence of nonresponse will be small if only the number of nonrespondents is small. In this case we must attempt to make nonresponse as small as possible. This can be seen from the following:

$$E(\hat{Y}) = y_r$$

$$\text{therefore } B(\hat{Y}) = y_r - Y = Y_r - (Y_r + Y_m) = -M\bar{Y}_m$$

The above expression is zero if $M = 0$ or $\bar{Y}_m = 0$

Correlation and Variances

Similarly one can examine the effect of nonresponse on other statistics such as variances and correlations. In such cases, however, one's intuition is only of a limited value. One should resort to examining mathematical expressions in relation to nonresponse bias. Since we are dealing with more than one variable the

error mechanisms are more complex and may operate in the opposite directions. On the whole, variance and covariance estimates should account for the additional sampling and non-sampling errors that arise from nonresponse. When weight adjustments by inverse response rates are employed to inflate the deficient sample, the variance and covariance estimates may be subject to small biases, relative to their true values in the population as long as appropriate replication methods are applied. The bias depends upon the way in which replicates are formed for variance/covariance analysis as well as the way adjustment cells for weight adjustments are delineated. One would have to set up a sampling/nonsampling error model along the lines of Platek and Gray (1983) or J. Lessler (1983) and examine the bias of the variance/covariance estimate.

When hot deck or historical survey data are substituted for missing values, the tendency is to treat the substituted data as real data. The variance estimates will likely be under-estimates under these circumstances, especially in the case of hot deck substitution, which is merely duplication of survey results from similar type respondents. In the case of historical data substitution if that data are highly correlated with current data, then one may be justified in treating that data as observed responses in variance/covariance analyses.

Thus, it may be observed that the treatment for compensation for nonresponse as well as nonresponse itself can affect various statistics differently. This fact may be helpful in the choice of the appropriate method for reducing controlling, or dealing with nonresponse.

(viii) Survey Management

Survey management is defined here as a series of control measures which would ensure that a survey produces statistics of acceptable quality. Some of the measures have a preventative character. Others attempt to measure the effectiveness of such programs in maintaining the quality. Thus, a careful analysis of nonresponse rates may help identify weaknesses in various programs. For example, any weakness in training of interviewers or the lack of proper understanding of interviewing procedures and conceptual issues will result in a higher bias. Also, the lack of adherence to interviewing procedures would increase the correlated response variance.

Two programs which would help analysis some of the nonresponse rates and their possible effect on data quality are observation and re-interview. Observation generally occurs at the time of the actual interviewing. It provides an opportunity for training (or retraining and motivating the interviewer) and identifying field problems. Some interviewers whose nonresponse rates are relatively high may be subjected to more frequent observations. Thus, the measures of nonresponse would determine intensity and frequency of control programs.

A reinterview is an interview with the units that have already been interviewed during survey week. The difference observed between the two sets of responses are attributed to several sources such as respondent, interview and re-interview. It is the interviewer performance that concerns us here with respect to nonresponse rates in certain areas and in certain categories of population. Any changes in response rates may be examined by attempting to establish a reason for them. Here again, the measures of nonresponse rates determine the intensity and frequency of reinterviews.

A more direct use of nonresponse would manifest itself in looking and examining interviewers work assignment, training, payments as well as sample design. For example, for the same surveys and under similar survey conditions one would expect a fairly uniform level of nonresponse due to no one at home. But if there is a large change between the same kind of surveys on two different occasions, then a number of causes can be identified. One is that there has been a change in interviewing procedures. Another possibility is that interviewers did not plan their work properly. From the design point of view it may be that the target and sample populations may have been changed at the design stage. All of these possibilities should be examined. It should also be emphasized that survey management should be aware that nonresponse and response errors can be related. For example, if a potential respondent is pressured to provide information that he/she would prefer not to provide, the respondent may give information that is not correct. This is sometimes referred to as the phenomenon of the reluctant respondent.

VI. DEALING WITH NONRESPONSE

It has been mentioned that at the design stage an understanding of the origin of nonresponse and its effect on survey data will undoubtedly lead to minimizing the size of nonresponse. The main thrust of dealing with nonresponse however, takes place in the actual collection of data and at the processing or estimation stage.

(i) Callbacks

Nonresponse can be reduced by persistent efforts of interviewers and by motivating nonrespondents to become respondents. The persistent efforts are usually in the form of callbacks which are an essential part of any survey. It is intuitively clear that are at home on the first call are likely to differ from households that are not at home. Thus, reliance on those at home on the first call only will likely result in a survey with characteristics which will be different from those obtained on subsequent calls. Therefore, in the design of survey data gathering procedure, there must be provision for callbacks in order to attempt to reduce the nonresponse and the resulting biases. The callbacks may take a variety of forms, depending on the type of survey and method of gathering survey data. Callbacks are the cleanest form of reducing nonresponse. No assumptions are needed. Other methods are based on the willingness of a survey designer to accept the risk of bias due to a particular method against the cost of making callbacks if they are feasible.

In mail surveys the callbacks may be a letter reminding the individual at the selected unit of the importance of the survey, usually with another copy of the questionnaire to be completed in case the nonrespondent discarded the first copy. The reminder must tell the recipient that the notice should be disregarded if he/she had mailed in the first copy of the questionnaire in the meantime. In mail surveys there is usually a limited number of repeated reminders (up to three) before some cut-off date for the data processing is decided upon. Less frequently, callbacks are made in the form of telephone or personal interview. These methods, especially the personal interview may defeat the purpose of the mail survey, which, after all, is to minimize the cost and to avoid the need to cluster the sample which tends to increase the sampling variance.

In the case of telephone surveys repeated attempts at callbacks by the telephone may be less costly than by personal interviews or by repeated mailout of remin-

ders. Thus, it may be possible to allow for more telephone attempts than by any other procedures. There may, however, be individuals who refuse to be interviewed by telephone or the interviewer fails to obtain a successful call. In such cases the final attempt(s) to obtain successful interview may have to be a personal interview. As in the case of mail surveys, the potential need for personal interview in follow-up of nonrespondents may dictate a more clustered sample design than a pure telephone survey.

Finally, in the case of personal interviews, callbacks may be made by repeated attempts by knocking at the doors or by telephone. A successful telephone call may be employed either to obtain the survey data in place of the personal interview or to obtain the best time to call for the interview. The number of callbacks by repeated calls personally or by telephone is usually limited because of the short duration of survey data gathering inherent in personal interview procedure compared with mail and sometimes telephone surveys.

Whatever the type of interview procedure of a survey, callbacks should be monitored at the selected units in order to provide input to an imputation system, to study the effect on the quality of data, and to determine sample design strategies for future callback procedures. At the design stage one can estimate the number of callbacks on the basis of past experience. It is important to have the variance of information obtained from callbacks and the cost associated with successive callbacks.

A number of factors such as demographic and socioeconomic characteristics of the interviewed, the strategies of the interviewer, like scheduling the callback, and the cost of obtaining the needed information affect the success of callbacks.

A number of studies conducted by various individuals and organizations confirm that:

- i. All surveys show that the number of response increase substantially as the number of calls increase.
- ii. In most surveys, the estimate of the mean after the first call differs substantially from the estimate after a several calls.

A number of experiments were conducted with respect to the methodology of interviewing at callbacks. For

example, one can interview any person, random adult or a specified person. Each demonstrated differences in the estimates between the calls. In order to establish the number of callbacks to be carried out it is important to know the cost of various operations associated with the calls. For more details see Durbin Stuart 1954, Drew, Fuller, ASS proceedings, 1980, Cochran, Deming, etc.

(ii) Proxy

A second major technique in collecting data is from an alternative respondent. For example, the survey instructions may state that any persons age 20 and over may be a respondent to provide information about other members of the household. Obviously, such responses are less likely to be correct than the responses of the designated persons. Still for certain types of surveys this may not be dangerous. But if the requested information is confidential or personal then a proxy response may be quite misleading. Studies of proxy interviewing should be conducted before its application in a particular case.

However, because of time and cost constraints it is virtually impossible to obtain non-proxy responses from every individual. For this reason proxy interviews are accepted in the Canadian Labour Force Survey. Generally only one member of a household is interviewed, and this member responds on behalf of all other members. Occasionally, separate interviews are required for household members such as roomers or boarders. On the average it has been found that proxy interviews account for approximately 50 percent of all respondents. Furthermore, because proxy interviews are accepted, it is possible to obtain complete responses for all household members in virtually 100 percent of all responding households. In fact, there are less than 0.2 percent of all households where interviews (proxy, non-proxy) are obtained for some, but not all members of the household. The acceptance of suitable proxy respondents is, therefore, an effective means of reducing nonresponse.

(iii) Refusal Follow-ups

Refusal households should be followed-up whenever feasible. In most cases this involves a personal visit by a senior interviewer or a regional office representative. In areas where this type of follow-up is not possible, a letter may be sent in which households should be provided with additional information about the survey and how the data will be used. The importance

of the survey and the cooperation of the respondent should be emphasized. The result based on Canadian experiences is that 20 to 30 percent of the refusal households can be successfully interviewed the following month. In the case of households which cannot be persuaded to respond interviewers are told not to visit them again unless there is a complete change in household composition.

(iv) Telephone

The telephone interviewing procedure which is used in the LFS involves a combination of personal visits and telephone calls and is carried out only in large urban areas. Interviewers must conduct all first month interviews in person, and telephone interviews can only be carried out in subsequent months if the respondent agrees to be interviewed by telephone. Experience has shown that approximately 75 percent of households are interviewed by telephone in assignments where the telephone interviewing procedure is allowed.

Although the primary reason for using telephone interviewing in the LFA is to reduce costs, telephone interviewing seems worthwhile from the point of view of nonresponse since (a) it allows interviewing to be completed on time regardless of weather conditions, (b) it is especially suitable for single persons, small family households and apartment dwellers who are difficult to find at home and who can often be reached only during the evening (it is easier for an interviewer to phone at night than to make a personal visit), (c) it allows interviewing to be conducted more readily at the convenience of the respondent (if one time is not suitable, then another can be easily arranged, and (d) it has the potential of reducing non-response by allowing more opportunity and time for callbacks.

(v) Quota

In quota sampling, the population is stratified usually according to variables expected to be closely related to the response that may be given. For example, a sample of households is selected and the interviewer is given instructions to obtain a given number of interviews from persons falling into certain categories of age, sex and income. The danger in such procedures is that the actual selection is made by the interviewer from among those who present themselves in these categories. The question is how much of a risk (bias) survey designers are willing to accept as against the costs of making callbacks. Quota samples will essentially have no response because the interviewer does not keep

a record of interviewing. Furthermore, quota sampling suffers from a ill defined selection mechanism and therefore it is impossible to estimate sampling error. Biases are due to differences in respondent availability. The advantage of quota sampling is speed and relatively small cost.

(vi) Substitution in the Field

One method of reducing the size of nonresponse at the data collection stage is to substitute other units that have not been selected for the current sample. There are two basic types of substitutions that are used.

- (a) Selection of other units at random
- (b) Selection of specifically designated substitutes

Selection at random- an additional population unit is selected on a probability basis to replace each nonrespondent. Usually the substitute for a particular nonrespondent is selected from a restricted population sub-group (e.g. the same block, are stratum or a group of strata where nonrespondent occurred). In each case it is assumed that the characteristics of a substitute chosen from a population subgroup will be more like the nonrespondent than would be the characteristics of a substitute chosen from the entire population. For many random procedures, potential substitutes are selected in advance of actual data collection in order to save data collection time. There are usually two operational choices i) each interview has a potential substitute with him/her, ii) they are communicated to interviewers on request. Each has advantages and disadvantages.

In practical, it is better to undertake the random selection of substituted units at the survey centre rather than rely on an interviewer to do so in order to ensure that he/she does not make a biased selection. In some procedures potential substitutes are selected prior to the data collection so as to avoid delays or problems such as selection bias if the substitutes were undertaken during or after data collection. It should be emphasized however that this procedure does not eliminate bias due to initial nonresponse. It only substitutes in an unbiased manner one unit for another that failed to respond, which may or may not have the same characteristics.

Selection of a specifically designated substitute
- Identifies one or more units which are assumed to

have similar characteristics of the nonrespondent. Sometimes this may be a neighbouring unit but does not have to be so.

In the procedure that uses specifically designated substitutes, the nearest neighbour is often selected because of the tendency for a similarity of characteristics among neighbouring units compared with randomly selected substitutes. If the nearest neighbour fails to respond, e.g. refuses to be interviewed or is never available for interview, the next unit is selected and the procedure is repeated until a unit has responded. In addition to the distance criteria for a substitute, an attempt may be made to search for a substitute unit with characteristics similar to the originally selected unit. Care must be taken to prevent the re-selection of a unit that has already been selected and has already responded to the survey. The key question regarding the worth of substitution procedures is whether or not the use of substitution provides better proxy values for nonrespondents than those provided by alternative imputation procedures. In examining the two methods of substitution, it is evident that the first substitution rule removes the sampling bias while it does not remove the nonresponse bias. The second one tends to remove the nonresponse bias but fails to remove the sampling bias. Undoubtedly, there are some advantages and disadvantages to the use of substitution procedures. The first advantage is that it is a convenient way of balancing the sample with respect to sample size. The other is the reduction of the sampling variance due to an increase in the effective sample size.

In further examining substitution methods one must look at them from the operation and Mean Square Error point of view. For example, it is more than likely that proximity may provide a better substitute than a weighting procedure in which an average is substituted. This may not be the case, however, if some additional information is available about nonrespondents in a particular group. When one deals with a small survey and whose design is deeply stratified it seems that substitution would be more appropriate than other methods. The more you know about nonrespondent, the better substitution you can make. However, in large surveys and with very little or no information about nonrespondents substitution would be inferior to say weighting procedures. One major disadvantage of substitution is that interviewers efforts are somewhat minimized in obtaining initial interviews.

(vii) The use of incentives

The method is well established and often used. An incentive may be less costly than an additional call. The planning of incentives should be based on the cost of additional calls that might otherwise be made in an effort to complete the interview. But

the use of financial incentives has many concerns. Firstly, this may become an accepted practice and the cost of surveys may increases. Secondly, to qualify for one incentive, erroneous information may be given by the respondent. It may be argued, however, that to provide information may be a burden for the respondents and they should receive some compensation for it.

VII. TYPES AND CONCEPTUAL ISSUES OF NONRESPONSE

As information flows from data collection to tabulation, the various types of responses can be identified and are presented as follows in Table 10.

From Table 10, it can be seen that two of the three groups of questionnaire following the edit stage require further action prior to estimation. These are (7) the unusable questionnaire, (6) those containing some blanks and/or inconsistent entries and (3) containing nonresponse. An unusable questionnaire could be classified as total nonresponse or it can be associated with a respondent household with some blank or inconsistent entries. In either case, however, further action denoted by (8) Imputation would be required. Complete nonrespondents are usually weighted up in some manner with the exception of the Census. The deficient questionnaires, (Item Nonresponse) on the other hand, fall into two categories such as (6) inconsistent entries of illegitimate blanks.

The inconsistent entries can be either logical impossibilities or they can be plausible but highly unlikely. It seems natural that if the entries are logical impossibilities and they can be detected as such, they ought to be adjusted even though they may not affect the data to any great extent. The adjustment would eliminate a great deal of embarrassment to subject matter analysts associated with the published reports.

In the case of plausible but highly unlikely entries, one is faced with a difficult choice between leaving what might seem to be an unnatural distribution or removing some of the extreme values of the distribution which may actually represent the real life situation. Ideally, one ought to opt for one or the other choice on the basis of experience with error mechanisms and the nature of the substantive distribution based on the knowledge of subject matter. In any case, one has to be able to identify the problem cases, i.e. one has to have suitable edit rules whenever one encounters impossible or highly unlikely events and a method of dealing them (i.e. imputation).

There is a fundamental distinction between editing and imputation. Let us consider the set of all possible code combinations on a questionnaire. Editing can be defined as the division of this set into two manually exclusive subsets: Those combinations which are judged acceptable and those which are unacceptable, the latter including questionnaires with invalid blanks and inconsistent entries. Thus, editing is basically a diagnosis and operationally it must be defined by a set of rules. Imputation, on the other hand, is more in the nature

of treatment of data.

Imputation is, as we have defined it, the assignment of data to empty fields (including total nonresponse) or a replacement of invalid data in fields following a certain set of rules. There is no known unbiased method of imputing but some methods may be more suitable than others.

It is possible that, rather than imputing for nonresponse at the time when survey tabulations are prepared, tabulations could be presented with the amount of nonresponse reported. In this case, the users would have a choice among various methods of imputation from tabulated data. At the first glance, this approach would appear to have some advantages giving the users the opportunity of selecting their own method of imputation. There are however, some serious disadvantages. Conflicting estimates would be produced by various users due to the different imputation methods employed, and problems in the consistency and integration of data would be created. As well the data collection agency is usually in a better position, due to its proximity to the sources of the data, to make imputation decisions. For these reasons, imputation is normally carried out by data collecting agencies rather than by individual users. The whole philosophy of imputation is based upon the expectation that an appropriate procedure, whether for nonresponse or for blanks resulting from edit failure, will provide a more logical relationship between cross-classified data and will also lower the mean square error of estimates.

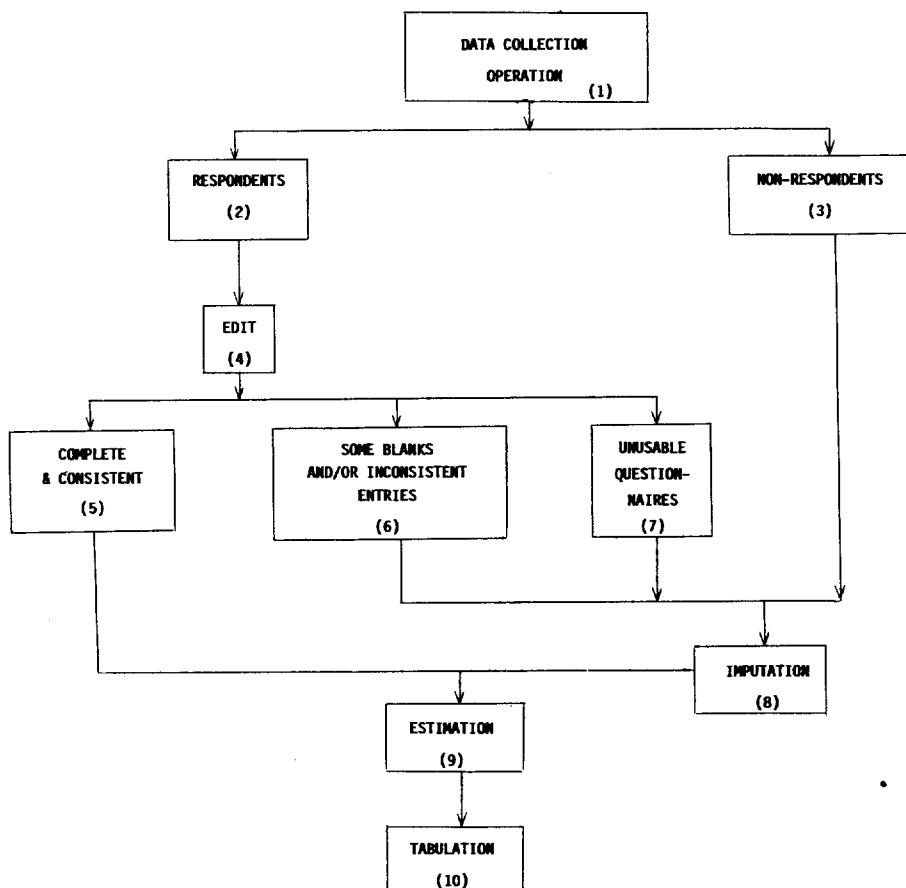
The simplest situation occurs when there is only one possible value which can be imputed for a field in such a way that after the imputation the record will be consistent. This is what is called deterministic imputation. For example, if wife is coded *male* then there is only one possible value to impute for sex to make it consistent with the other information. Sometimes, there may be more than one value which would make the record consistent. If this is the case, one would choose a particular value which is more predominant in proportion to total frequency or more plausible. A good example of this kind can be found in the Labour Force Survey where in the fall to spring months, for 15 and 16 years old persons, if there is no Labour Force characteristics entered, one imputes that they are attending school, although it is not at all impossible they do not attend school. So long

as the proportion of such cases is sufficiently small, the effect of this imputation will be a slight increase in bias, but there will be some reduction in variance.

In other situations where one could reasonably impute a whole range of values, one needs some other criteria. One possible criterion would be to minimize the mean square error of the resulting estimates. The question, however, arises, the mean square error of which estimates? With the continuously increasing demand for data tabulated in a number of different and unforeseen ways one really does not know which mean square error one ought to minimize. Furthermore, one would not know all the kinds of aggregates to which a particular record may con-

tribute in different kinds of tabulations. Consequently, one might prefer to use some other criterion which would produce the most appropriate entry for a field in a particular record in relation to the other information in the record. In other words, how can one best predict the value of one field on the basis of knowing the other fields on the record? A good example of this kind of imputation is the use of previous month's data in the Labour Force Survey: for a particular person, one could hardly find a better imputed value, particularly in those cases where demographic characteristics change slowly. If one does not have information based on the past, one would have to resort to other methods of imputation.

TABLE 10
FLOW CHART PERTAINING TO EACH SAMPLED UNIT



This is, of course, a highly simplified diagram of the process and it is produced only for the purpose of the discussion.

VIII. IMPUTATION METHODS

For a number of years various procedures of imputation for missing data due to nonresponse have been used in household surveys and censuses. The use of a particular procedure has been, to my knowledge, mostly justified on the ground of experience, intuition and experience. It was often assumed that the probabilities of units responding were uniform and frequently the nonresponse bias was largely ignored. Imputation is the process of assigning values for missing values to produce a complete data set. The aim of imputation is, like weighting adjustment to reduce bias due to nonresponse. Imputation ensures that the results obtained from different analysis applied to the data sets are bound to be consistent with one another. A feature which may not apply when analysis is performed on incomplete data set. Whatever the methods of imputation, there is no guarantee that the results obtained after imputation will be less bias than those based on incomplete data set. Since, after all, imputation merely fabricates data. It does not increase the amount of information available and indeed the use of random subsample of donors (respondents) serves to reduce rather than to increase the effective sample size. Like sampling with replacement has a higher variance. However, it is hoped that the knowledge of various relationships and their application to missing data will reduce bias due to nonresponse.

There are a number of imputation procedures which are applied and a few will be briefly mentioned below.

(i) Deductive Imputation

In its present form this method is applied in situations in which a missing response can be deduced from other information on the record. Thus, for example, if a respondent's sex is missing, the person has a female name and is known to be married to a male, the sex of the respondent may be deduced to be a female. This is a very popular method and it is frequently used.

(ii) Hot Deck

It is a common method of adjusting data sets for missing values and there are many variations of hot-deck procedures. In general, a hot-deck procedure is a duplication process. When a value is missing from a sample a reported value in the sample is duplicated to replace it.

The primary purpose of a hot-deck is to reduce nonresponse bias. To reduce this bias, a hot-deck procedure

has some classification process associated with it. All of the sample units are classified into disjoint groups so that the units are homogeneous as much as possible within each group. For each missing value, a reported value is imputed which is in the same group. Thus the assumption is that within each classification group the nonresponse follows the same distribution as the respondent. This assumption places a heavy burden on the classification variables. These variables must be correlated with the reported values and also correlated with missing values. If this does not hold, Hot-deck is a part from the reduction of nonresponse bias, (i) production of a clean data set, i.e. a data set which appears complete and clean, (ii) reduce bias while preserving joint and marginal distributions. For example, if one substitutes the mean for marginal distributions. For example, if one substitutes the mean for each missing value then the distribution of sample values is affected. If one chooses a reported value at random, one may reduce the distortion in the distribution but the bias may not be reduced.

As a method of imputation hot-deck procedures have some attractive features, including the following:
 (a) the procedures result in a relatively easy way of constructing post-strata, (See Fellegi, Holt), (b) matching of records does not present any special problems
 (c) no strong model assumption need be made in order to estimate the individual values for missing items.

However, Hot-deck procedures have certain disadvantages. They lack probability mechanisms, therefore it is impossible to calculate its reliability without some modelling procedure. There is a possibility, of the multiple use of the same donor.

In evaluating hot-deck procedures one would like to know how the bias and reliability of the principal estimates are affected by the size of classification groups (often referred to as weighting classes), the frequency of missing data, the choice of matching items, etc. Some theoretical work relating to hot-decks has been done by (Fellegi and Holt), (Bailar and Bailar 1978) and (Cox B. and Folsom R.E. 1978).

(iii) Modified Hot-deck

The essence of it is the sorting and matching of potential donors and recipients on a sizeable number of variables. The matching is done on hierarchical basis in the sense that if no donor is found to match a recipient on all the control variables, some of the

variables are regarded as less important and are dropped so that the match is made at a lower level. The hierarchical levels are defined for this purpose, with the lowest being such that a donor can always be found.

(iv) Regression

With this process the item y for which imputations are needed is regressed on control variables (x_1, x_2, \dots, x_p) for units providing a response on y . The control variables may be both qualitative and quantitative. The qualitative variable is incorporated into the regression model by means of dummy variables. Though careful development of an explicit model, regression imputation has the potential to produce values closer to the true value than other imputation methods. The construction of a good regression model is costly, time consuming and it seems unrealistic to consider its application for all the items with missing values in a survey. It is used at all, it seems best to reserve it for a few major variables.

Weighting

One of the most common procedures for accounting for non-response at the processing and estimation stage is that the design-dependent balancing area, in which the weights are further inflated by the inverse of the response rate. In a balancing area b , an estimate of the characteristic total is given by $\hat{x}_b = \sum_{i=1}^{n_b} x_i / \pi_i$, where x_i is the response, π_i = the inclusion probability, and n_b is the sample size in the balancing areas. If only m_b units respond, then the weight π_i^{-1} is further inflated by the inverse of the response rate, m_b/n_b , i.e. by the fact n_b/m_b .

The balancing areas should, preferably, be determined at the planning stage rather than at the processing stage and they could be individual strata, groups of strata, a province, primary sampling unit, or a cluster. The choice of balancing area is quite crucial since the nonresponse rates and the bias may differ from area to area.

An important methodological problem, for example, is to determine an optimum or in some way appropriate size of balancing area. The appropriate refers to a proper size which would ensure both a sufficient response rate and the advantages due to the measures of homogeneity to help reduce the nonresponse bias. It can readily be shown that the weight inflation of all the records in a balancing area to compensate for nonrespondents is equivalent to the substitution of the mean values of

all the weighted respondents to each nonrespondent in the area. If a characteristic has a high measure of homogeneity (increasing with decreasing size of area), then weighting (or substitution of mean value) in small areas would tend to result in mean values that are more similar to the actual characteristic value of the nonrespondent in those areas than would be the case in larger areas. Thus, the nonresponse bias would tend to be lower in the case of small balancing areas than in the case of large balancing units. What about the variance? As balancing areas become smaller, the weight inflation becomes more unstable as the variation in response rates becomes more unstable among many small balancing areas as opposed to a few large balancing areas and the instability of the weight would tend to increase the variance. Clearly, there is some trade-off on the size of the balancing area between small areas to take advantage of the measure of homogeneity and large areas to ensure stability in the weight adjustments. The possible extreme values of the sizes of the balancing areas are the whole sample at the upper end and a size of 1 at the lower end. However, in the latter case, one is faced with the problem of what should be done if that unit fails to respond. Instead of substitution of the mean value, one would have to resort to regression analysis or superpopulation models (an entirely different approach to substitution) or else employ historic values. The choice of the size of balancing area depends not only on the measure of homogeneity but also on the sample design, the sample size and the response rate. Surveys with low response rates would require larger balancing units than those with high response rates. One could utilize small balancing areas and adopt some procedure of collapsing them until the response rate reaches some respectable level (not too much below the overall response rate) so as to minimize the instability of the weight. The collapsing of balancing areas however adds a complex dimension to the variance estimation since one would have to consider the probabilities of collapsing 1, 2, 3, 4, etc., balancing areas and the choice of 1, 2, 3, or 4 balancing areas. While such a procedure is undertaken in the Labour Force Survey, the need to collapse is frequent enough not to warrant a special treatment for variance estimation purposes. Consequently, if any variance calculations or analysis other than mere averages or totals are contemplated, balancing areas that are expected to be stable without much collapsing should be incorporated into the sample design. That is, the response rates should be sufficiently large with high probability to avoid the necessity of collapsing balancing areas if variance estimation is contemplated. This would

discourage one from using small areas to balance for non-response.

(vi) Duplication

Instead of weighting by the inverse of the response rate in a balancing area, one could duplicate a sufficient number of units among the m_b nonrespondents to bring the apparent sample size up to the original level of n_b units. However, it can be shown that an additional variance component occurs over that incurred when simple weighting is applied. The main advantage of duplication versus weighting lies in ensuring that integral rather than fractional weights are applied to each record. In certain types of published data, e.g., the number of persons with some characteristics, integral weights would tend to avoid rounding errors when sub-classifying data. When one estimates means or proportions or certain types of quantitative totals such as gross national product, the use of integers rather than fractional weights are of no advantage.

Apart from the comments in the above paragraph, the methodological problems concerned with weighting in balancing areas also apply to duplication in balancing areas.

(vii) Substitution

Another important method of imputation for nonresponse is one of substitution of historical (previous month's data) or external source data (administrative files, other surveys, Census data). Once the substitution of historical or external source data has been undertaken to the extent possible for nonrespondents, the weighting or duplication may be affected within balancing areas. In the case of weighting, one would inflate the weight π_i^{-1} by the factor $n_b(m_b+m'_b)$, where m_b respondents were obtained as before and for m'_b of the (n_b-m_b) nonrespondents, historical records were substituted for the missing data. In such a method of imputation the sampling variance is reduced from that which occurs in the weighting scheme since we have increased the effective sample size from n_b to somewhere between m_b and $(m_b+m'_b)$ units. The increased sample, including those records imputed from historical records will not be as good as $(m_b+m'_b)$ since historical or external source data are not as good as current response information unless there has been no change in the characteristics of the units for which substitution of historical data was undertaken.

Alternatively, one may wish to duplicate respondents;

i.e. take a sample from respondents equal in size to the number of respondents and apply a weight of 2 instead of inflate the weight for all the respondents. However, one may wish to avoid duplication of those nonrespondents for which substitution of historical information had been undertaken but one would rather subsample $n_b - (m_b+m'_b)$, say m'_b units from the m_b respondents to duplicate in order to bring the apparent sample size from $(m_b+m'_b)$ to n_b units in balancing area b. The estimated total for balancing area b would be

$$\hat{x}_b = \sum_{i=1}^{m_b} w_i x_i / \pi_i + \sum_{j=m_b+1}^{m_b+m'_b} x'_j / \pi_j$$

where x'_j is the imputed value for unit j and $w_i = 1$ or 2 (2 for a random subsample of $n_b - m_b - m'_b$ units among the m_b respondents). The expected value of \hat{x}_b over all possible ways of duplicating is

$$\hat{x}_b^* = (n_b - m'_b) / m_b \cdot \sum_{i=1}^{m_b} x_i / \pi_i + \sum_{j=m_b+1}^{m_b+m'_b} x'_j / \pi_j$$

Consequently, $V(\hat{x}_b) = V(\hat{x}_b^*) + (\text{additional component of variance as a result of subsampling among the respondents})$. \hat{x}_b^* is not the same as the estimate

$$n_b / (m_b + m'_b) \left[\sum_{i=1}^{m_b} x_i / \pi_i + \sum_{j=m_b+1}^{m_b+m'_b} x'_j / \pi_j \right]$$

and the variance of \hat{x}_b^* is also different from that of the estimate where the weight inflation of $n_b / (m_b + m'_b)$ is applied.

The estimation procedure dealt with above include weighting or duplicating in design-dependent balancing areas. If historical or external source data are available for some of the nonrespondents, these could be employed for imputation purposes prior to weighting or duplication in balancing areas. Instead of balancing areas, one could utilize weighting classes for imputation purposes and these are discussed in the next paragraph.

Weighting classes are distinguished from balancing areas in that they generally comprise characteristics of ultimate units (e.g., dwelling types, special income groups, etc.) as opposed to geographic areas, though one could conceivably group areas according to some distinct characteristics that are not related to the sample design. Usually, one defines weighting classes as well as balancing areas prior to the survey gathering procedure and makes adjustments through collapsing

if the response rates are unacceptably low or the sample too small to employ any type of adjustment of the weights. In some imputation procedures, however, weighting classes are defined after the survey data have been gathered where factor analysis or other analytical tools are employed to determine the most efficient set of weighting classes. After the weighting classes have been determined, the estimation procedures are essentially identical to those used in balancing areas. The biases and the variances (at least in terms of individual and joint inclusion probabilities of the ultimate units and other parameters not related to the sample design) are identical. However, upon further expansion of the variance to take into account the particular sample design, the variances of the estimates pertaining to balancing areas and weighting classes will be quite different. In order to utilize weighting classes for imputation purposes some knowledge about the non-respondents (such as income class, size of household dwelling type) must be available. In practice, when such information is not available, the procedure cannot be used.

The estimation formula for the methods of imputation discussed here may be written as below.

$$X_b = \sum_b X_b \text{ estimates the total of some characteristic}$$

where b is either the balancing area or weighting class. The estimate for a given balancing area or weighting class is in turn given by

$$X_b = \sum_{i=1}^{m_b} w_i x_i / \pi_i + \sum_{j=m_b+1}^{m_b+m'_b} w_j x'_j / \pi'_j, \text{ where } \pi_i \text{ or } \pi'_j$$

is the inclusion probability and m_b is the number of units that responded out of n_b units in balancing area b .

$$x_i = \text{response value for unit } i = x_i \text{ (true value)} + R_{\varepsilon_i} \text{ (response error)}$$

$x'_j = \text{historical value for unit } j \text{ (if available) given that unit } j \text{ failed to respond. Among the } (n_b - m_b) \text{ nonresponding units } m_b \text{ possess historical records in balancing area } b.$

$$x'_j = x_j \text{ (true value)} + R_{\varepsilon_j} \text{ (response error or historical value relative to } x_j)$$

w_i and w_j are the weights, appropriate to a particular method.

(viii) Raking Ratio

This procedure can be a useful weighting adjustment when weighting classes are defined in terms of the cross-tabulation of a number of characteristics. The underlying basis of the technique is to make the weighted marginal distributions conform to those in the population. There are two main reasons for Raking Ratio. 1) when many classes are formed by the cross-tabulations of characteristics, the resulting separate class weighting adjustment can be extremely variable and unstable because of the small size in the classes. Secondly, separate class weighting adjustment require knowledge of joint distribution of the weighting classes characteristics for the population, whereas Raking Ratio only require knowledge of the population marginal distribution.

Population				Sample					
Age				Age					
1	2K	Total	1	2	..K	Total		
1	w_{11}	w_{12}	w_{1K}	$w_{1.}$	1	q_{11}	q_{12}	q_{1K}	q_1
2	w_{21}	w_{22}	w_{2K}	$w_{2.}$	2	q_{21}	q_{22}	q_{2K}	q_2
	$w_{.1}$	$w_{.2}$	$w_{.K}$	1		$q_{.1}$	$q_{.2..}$	$q_{.K}$	1

The W 's are the population distribution of the two characteristics, while q 's are the sample distribution of the two characteristics.

Raking Ratio employs an iterative procedure to make the weighting sample marginal distributions ($q_{1.} \dots$)
($q_{1.} \dots$)

equal to population marginal distribution
($w_{1.} \dots$)
($w_{1.} \dots$)

The bias in the Raking Ratio depends on the closeness of the cell distributions between the two populations and the difference of the means between respondents and nonrespondents in the cell.

TABLE 11
PROCEDURES FOR DEALING WITH NONRESPONSE AND IMPLICATIONS

<u>Stage of Survey</u>	<u>Dealing with Missing Data</u>	<u>Implications</u>
Planning	<ul style="list-style-type: none"> - Unit nonresponse rates anticipated from earlier experience of similar type of survey according to method of gathering of survey data. - Sample increased to allow for nonresponse (oversampling). - Pilot studies to test questionnaires and field procedures. 	<ul style="list-style-type: none"> - Various strategies must be determined for the field and processing stage / (callbacks, imputation, etc.) - Sampling variance would likely remain as expected; however nonresponse bias remains. - One attempts to reduce response error and item nonresponse.
Interviewer Training	<ul style="list-style-type: none"> - Aptitude testing, scenarios for survey taking, clustering of assignments to permit callbacks. 	<ul style="list-style-type: none"> - Experienced interviewers would be able to minimize nonresponse and hence, // bias especially in continuous surveys; for ad hoc surveys there may be little one can do but make repeated callbacks, resulting in higher costs per selected unit.
Survey Data Gathering	<ul style="list-style-type: none"> - Attempts callbacks, mailbacks, repeated phone attempts until as many successful responses are obtained as possible; attempts to convert refusals through diplomacy, media, etc. - Substitution of other units in the field not / originally selected. - Substitution of other units from a pool of reserved units (increase of sample to compensate for nonresponse). - Subsampling of nonrespondents. - Ignoring nonresponse in the field. 	<ul style="list-style-type: none"> - More callbacks increase the cost per / unit, especially if proper time for interview cannot be determined from neighbours; reduced level of nonresponse reduces sampling variance and nonresponse bias. - Reduced sampling variance but there may exist sampling and nonresponse biases. - Reduced sampling variance and no sampling bias but nonresponse bias may remain. - Increased cost because of extra effort required and nonresponse bias reduced, though not entirely eliminated. - Reduced sampling variance, possible under-estimate of totals unless weight adjusted at processing stage. - Reduction of nonresponse error not elimination of it by appropriate imputation / procedure. - Some reduction in sampling variance // through ratio estimation or use of independent sources of data for imputation. - Slight increase in processing cost because of possible complex procedure to deal with missing data. - Some difficulty of analysis of complex / survey data beyond that arising from / using other than SRS sample, because of missing data and imputation for them. - Difficulties of warning users if missing data when they are being analysed.
Processing and estimation Stage	<ul style="list-style-type: none"> - Imputation for nonresponse. - Formation of adjustment cells. - Weight adjustment or explicit substitution / for missing data? 	

IX RESPONSE PROBABILITY APPROACH

Although variations in response rates have been detected among units according to their characteristics, the effect of individual units responding or not responding upon the bias and variance of the estimates has usually been insufficiently examined.

To facilitate a detailed examination of the effect, Platek and Gray (1978) have developed methodology with respect to the bias and variance pertaining to several imputation procedures. The development of the expression of bias and variance of the estimates is based on a fundamental concept that a unit, if selected, responds or does not respond with a certain response probability attached to that unit. This is an extension of the approach taken by Platek, Singh and Tremblay in 1977 with respect to censuses.

(i) The Concept of Response Probability

Close examination of the behaviour of response rates for a variety of surveys under different survey conditions leads to the natural assumption that in the population, each unit has some probability of responding (if selected), given the survey conditions. In other words, for most situations, the population is not formed of a dichotomy of respondents and nonrespondents, but rather of a collection of potential respondents some of which are respondents and the others nonrespondents depending on the conditions under which the survey is conducted. The concept is implicit in the application of the Politz-Simmons Technique (1949).

So we have defined response probability as the probability that a selected unit provides a response completely or partially to a survey questionnaire. The response probability may be related with the time of day that an interview takes place. It may be also correlated with such characteristics as age, sex, income level of respondent, size of dwelling (i.e., number of persons residing within) interviewing procedure, the subject and interest of the survey. Response probabilities are usually higher in the case of personal interviews than for either telephone or mail surveys but not in all cases.

As implied in the definition of unit response probability in the preceding paragraph, when a unit responds, it may or may not respond to a particular question. One may therefore define an item response probability for a particular question as the probability that a selected and responding unit responds to that question.

This probability depends largely on the complexity of the question, and in the case of personal and telephone surveys at least, the ability of the interviewer to explain or clarify the questions to the respondent.

In developing further the concept of response probabilities, let us examine the possible correspondence between the response rates which depend upon the characteristic and the differences in response probabilities for a particular characteristic. Consider a situation of a complete enumeration (census) of a population of N units of which Y have characteristic y and $N-Y$ do not. The total number of units with the characteristic y is to be estimated. Suppose however, at the time of enumeration, we have $R < N$ respondents and $M = (N-R)$ nonrespondents. Suppose further that the R respondents are divided into Y_R with the characteristic y and $(R-Y_R)$ not having characteristic y . Suppose also that the M non-respondents are divided into Y_M with the characteristic y and $(M-Y_M)$ not having the characteristic. The population of N units may be split up into the 4 categories as in Table 12 below.

TABLE 12
CENSUS OF POPULATION TO ESTIMATE Y
(Simple situation)

	Respondents	Nonrespondent	Total
Having the Characteristic	Y_R	Y_M	Y
Not Having the Characteristic	$(R - Y_R)$	$(M - Y_M)$	$(N - Y)$
TOTAL	R	M	$N = (R + M)$

The parameter Y , which is of course unknown except under full response may be expressed in the following two ways:

$$(i) N(Y/N) \text{ or } (ii) Y_R(Y/Y_R)$$

From the above table, it is seen that the overall response rate = R/N which is also the average response probability of all N units. Similarly, the response rate among persons with characteristic y equals Y_R/Y , which equals the average response probability among all units characteristic y . Among persons not having

the characteristic, a similar expression for the nonresponse rate or the average response probability equals $(R - Y_R)/(N - Y)$. These rates may or may not equal the overall response rate of N/N . The extent to which the rate deviate from R/N determines the nonresponse bias of the estimate.

The true total Y is unknown except under full response but it may be estimated by either of two equivalent expressions:

$$(a) \bar{Y} = N(Y_R/R)$$

or (b) $\bar{Y} = Y_R(N/R)$

In case (a), (Y_R/R) estimates the proportion (Y/N) of persons with characteristic y , assuming that respondents and nonrespondents have similar characteristics, i.e., one assume that $(Y_R/R) = (Y/N) = (Y_M/M)$.

Alternatively, in (b) (N/R) the overall inverse response rate estimates the inverse response rate (Y/Y_R) for the persons having the characteristic on the assumption that the response rates (or average response probabilities) are identical among persons having or not having the characteristic. The estimate of the parameter Y could be written as $\bar{Y} = (Y_R/R) \cdot N$ under (a) which is obviously the same as that arrived under (b) namely $(N/R)Y_R$.

(ii) Various Approaches to Response Probability

The use of a particular approach within the Response Probability concept has practical and theoretical consequences which will be touched on briefly below.

(a) Dichotomy: Assumes that population divided into strata of respondents with probability 1 for respondents and 0 for nonrespondents. Under this assumption there would be a bias if the characteristic of interest were correlated with the response probabilities of 1 and 0. The nonresponse bias that occurs under the assumption reduces to the difference in the means of the characteristic of interest between the respondents and nonrespondents. The difference may be detected by subsampling nonrespondents; however, in so doing, there is immediately an implied non-zero probability of responding among the so-called nonrespondents, a contradiction to the dichotomy assumption.

(b) Constant Probability: Assumes that each unit in the target population has equal probability of response. Thus a subsample of respondents is identical to

the sample of respondents.

(c) Variable Probabilities: Assumes that the target population is made up of respondents, each having certain response probabilities under given survey conditions. Under this assumption, the bias in the estimate would depend on the magnitude of the correlation between the characteristic of interest and the response probabilities of the units.

(iii) Theoretical Implications of Response Probability Approach

(a) Nonresponse Bias: The nonresponse bias frequently referred to in the literature usually pertains to the estimates of such statistics as means, totals, etc. after compensating for missing data due to nonresponse. A common method of imputation is one of weight adjustment by inverse response rate in an adjustment cell. It can be readily shown that nonresponse bias of an estimate with the weight adjustment procedure is proportional to the correlation between magnitudes of the response probabilities and the values of the characteristic whose statistics are being estimated. Under a dichotomous approach, the response probabilities are 0 and 1 and the correlation exists when the characteristics tend to differ between respondents and nonrespondents. In the literature where nonresponse bias is discussed, the dichotomy approach is usually assumed, whereby respondents and nonrespondents form separate strata and the differences in their characteristics are analyzed theoretically using other data sources. In the approach of Platek, Singh and Tremblay (1977), Platek and Gray (1983), the variable response probability approach results in a more general approach to the derivation of nonresponse bias, with the dichotomy approach as a particular case. In the case of equal response probabilities, when weight adjustments are used to inflate the responding subsample to a full sample there is no bias due to nonresponse.

(b) Mean Square Error (MSE): The response probability approach makes it possible to develop the general theory of sampling and non-sampling variance components in the context of nonresponse and imputation for it. The theory may be derived from a sampling, response and imputation error model pertaining to different imputation procedures for missing data. The development is more general under the variable response probability than under either the dichotomy or equal response probability approach. These two cases may be derived as special cases but one cannot proceed the other way by assuming response/nonresponse strata or equal response probabili-

ties and obtain the expressions based on variable response probabilities.

(c) Allocation Studies and Formation of Adjustment Cells: Adjustment cells are very important in the area of imputation. Instead of employing, for example, one weight adjustment to inflate the respondents to the full sample, the weight adjustments may be undertaken in separate cells defined by separate subsamples in as post-strata cells of sampled units with specific characteristics or in separate design strata or primary sampling units. The post-strata cells are referred to as weighting classes while the design-dependent cells are referred to as balancing areas. The imputation procedures are essentially identical for the two types of cells. However, the mean square error expressions, particularly the sampling variance for estimates based on the two types of cells, may be quite different.

(iv) Practical Implications of Response Probability Approach

The variable response probability approach is more realistic than the dichotomy or constant response probability approaches. Only a small portion of the population in the sample selected from it would have no chance of responding (hard core nonrespondents) or a full chance of responding (someone always available and fully co-operative to the survey taker). Many hard core refusals may be converted into willing respondents through response incentives (Gower, 1977), tact and sourcefulness of an interviewer or public relations. Equal response probabilities for all units pertain to another unrealistic assumption since in practice, it is difficult to delineate adjustment cells containing units with equal response probabilities.

However, the application of a fixed weight adjustment in a cell by an inverse response rate implies the assumption of constant response probabilities for all the units in the cell. The consequence is a nonresponse bias, depending upon the correlation between the actual response probabilities and the value of the characteristic in each cell.

The practical implications of response probability approach are inherent in the use of hot deck imputation procedures, particularly for item nonresponse but occasionally employed for unit nonresponse. In a hot deck procedure, the units are delineated into similar type records in a cell, applying some criteria of similarity. When missing data occur because of either

item or unit nonresponse, a so-called donor record in the cell is used to impute the missing data. The selection of a donor record is either systematic or random, often on the basis of a distance function pertaining to pairs of units. Clustering, which may be measured by a distance function for geographic proximity may be considered for the choice of donor records. However, through the formation of a cell of potential donor records, the response rate, which is usually the item response rate, is an estimate of the average item response probability in the cell of potential donor record.

Using the Horvitz-Thompson estimation, Platek and Gray have developed a general imputation formula for \tilde{Y} given below.

$$\tilde{Y}^{(G)} = \sum_{i=1}^N t_i w_i w_{ai} \{ \delta_i | \delta_{iy} y_i + (1-\delta_{iy}) z_{iy}^1 \} + (1-\delta_i) \cdot z_{iy} \}$$

t_i = 1 or 0 according as unit i is selected or not selected,

w_i = sample weight applied to unit i , when selected,

δ_i = 1 or 0 according as selected unit i , responds or does not respond, $\Pr(\delta_i = 1 | t_i = 1) = \alpha_{iy}$,

δ_{iy} = 1 or 0 according as responding unit i responds to characteristic y or no, $\Pr(\delta_{iy} = 1 | t_i = 1, \delta_i = 1) = \alpha_{iy}$,

w_{ai} = weight adjustment to compensate for unit nonresponse,

y_i = observed value of characteristic y for unit i , as defined earlier in (4.6),

z_{iy}^1 = imputed value of characteristic y for item nonresponse,

z_{iy} = imputed value of characteristic y for unit nonresponse.

X. APPLICATION OF ADJUSTMENTS FOR NONRESPONSE IN THE CANADIAN LABOUR FORCE SURVEY

The adjustment procedures that are in common use in sample surveys and census include weight adjustment, duplication, substitution of historical or external data and Hot Decks.

The application of these various procedures will be demonstrated in relation to such important surveys as the Canadian Labour Force Survey and other major surveys. Specific examples will be given and traced for partially completed questionnaires.

One of the major continuous surveys conducted by Statistics Canada is the Labour Force Survey from which monthly estimates of unemployment, employment, and many other characteristics are obtained. The data are published less than three weeks after survey about 56,000 households in approximately the third week of each month. It is impossible to contact every household that should have been contacted because of the stringent schedule in collecting and processing the data. Some households are away for the entire week, absent each time the interviewers call, or else they refuse to be interviewed. There are of course other reasons for nonresponse but they make only a small contribution to the total nonresponse which is maintained around five percent most of the time except for increases to seven or eight percent in the summer months because of temporary absences due to vacations.

Adjustment for complete household nonresponse is carried out according to the following criteria: (i) for about one-third of the nonrespondents substitution of last month's values where they are applicable (with suitable transformations in some fields to update last month's data or (ii) inflation by inverse response rate in balancing units (response in this case including substituted values for those nonrespondents who actually responded in the preceding survey with applicable data). In the case of (ii), the imputation for the remaining 2/3 of the nonrespondents is implicitly the mean of all respondents, in the balancing unit (primary sampling units, small urban and rural portions of primary sampling units, or subunits in large cities).

Adjustment for item nonresponse or edit rejects is carried out in one of three ways, depending upon the item or items with missing or faulty data and depending upon the response status and characteristics of the unit in the previous survey:

(i) The proper item response that has been omitted

can be unambiguously deducted from the remainder of the questionnaire (a decision table would ensure a unique and consistent response).

(ii) The substitution of the item response of the previous survey if it is available and if it is appropriate according to a decision table.

(iii) The application of a Hot Deck procedure whereby a similar record is obtained in the same primary sampling unit, same path taken (one of six possible) in the sequence of questions and same age-sex group. Here, collapsing of weighting classes may be required to find a similar type record; usually age-sex categories rather than psu's or paths taken are grouped together for adjustment purposes when necessary.

To illustrate the imputation procedure for partially completed questionnaire, consider the following examples, pertaining to the Canadian Labour Force Survey (LFS).

Example 1. (Canadian Labour Force)

Every person 15 years of age and over, within the households selected for the LFS is asked, Question 80 (Q80) Last week, was he/she attending a school, college or university? If he/she is, then the response is coded 1, if not, then the response is coded 2. For those not going to school no related questions are asked. However, for those going to school, there are follow-up questions: Question 81 (Q81), Was he/she enrolled as a full-time or a part-time student? If yes, then Question 82 (Q82) What kind of school was this? Thus, the following situation may arise.

$Q80 \neq 1$ or 2 which is in error.

The required relationship between questions must be of the following kind:

- a. If $Q80 = 1$, i.e. going to school,
then $Q81 = 1$ or 2, i.e. full-time or part-time
and $Q82 = 1$ or 2 or 3 or 4, i.e. the type of school
- b. If $Q80 = 2$ i.e. not going to school
then $Q81 = \text{blank}$ (not applicable)
and $Q82 = \text{blank}$ (not applicable).

There are also entries in other questions that may be

related to Q80 and Q82 and these are: Q14 and Q36: Attending school as a reason for working < 30 hours, Q58 Going to school as an answer to what the respondent was doing immediately before looking for work and Going to school as a reason in Q64 for not being able to take a job last week. In each case, the code is 3.

As an example of a specific set of conditions**, let us suppose the answer to Q80 is not available or is inconsistent with the logic of subsequent questions. The following decision table summarizes the steps followed in imputing the appropriate value for Q80.

	Imputation Rule						
	1	2	3	4	5	6	7
Q81 = 1 or 2	Y	N	N	N	N	N	N
Q82 = 1, 1, 3, or 4	Y	N	N	N	N	N	N
Q14, 36, 58 or 64 = 3		Y	N	N	N	N	N
Age = 15 or 16 yrs.			Y	Y	N	N	N
Survey Month = July or August*				Y	N	N	N
Q80 in Previous Month = 1					Y	N	N
= 2						Y	N
= 1 or 2							Y
Then Q80 in Current Month	1	1	2	1	1	2	Hot Deck

Here, Hot Deck means a search within the same primary sampling unit, path and age-sex group as stated earlier.

Each column represents a separate sequence of steps that must be followed in order to arrive at the imputed value of whether a person is or is not going to school (1 or 2). The Y = Yes and N = No in the main body of the decision table corresponding to the condition statements.

The detailed discussion for each imputation rule is given below:

1. The imputed value is based on the internal logic of the related questions. Thus if Q81 and Q82 have valid responses, then only the value of Q80 consistent with other information is assigned i.e. if Q81 = 1 or 2 and Q82 = 1, 2, 3 or 4, then impute 1 for Q80, i.e. going to school.

2. If the information for Q81 and Q82 are not available, then the next step is to seek relevant information elsewhere within the Questionnaire. The other questions are asked depending upon the path a person may have followed. Should he/she have indicated earlier that he/she was going to school, then Q80 is also coded as such, i.e. if Q81 ≠ 1, 2 and Q82 ≠ 1, 2, 3, 4 but Q14, 36, 58 or 64 = 3, then Q80 = 1, i.e. going to school.

3-4. If there is no directly related information available, then use is made of the information that is indirectly related to the question in edit conflict. In this case, it is the age and the month of the survey. If a person is 15 or 16 years old, then during the months of July and August he most likely was not going to school and thus Q80 is coded 2, i.e. not going to school. Whereas for all other months he is coded as going to school, i.e. Q81 = 1, i.e. going to school.

* For persons not 15 or 16 years of age, it does not matter whether the survey month is July or August for imputation rules 5, 6 and 7.

** It excludes other conditions, for example if Q81 is not 1 or 2 but Q82 is 1, 2, 3 or 4 vice versa.

These imputations are examples of rare cases so that the assignment of the *most probable value* is justified, even though hot deck might be theoretically better but possibly more expensive.

5-6. If no information in the current month's data is available and the person is not 15 nor 16 years

old, the next recourse is to the previous month's data. Whatever the response is from the previous month, if it is available, it is transferred to the current month's questionnaire.

The use of last month's data is justified largely on the ground that the month to month correlation of characteristics of certain estimates is quite high.

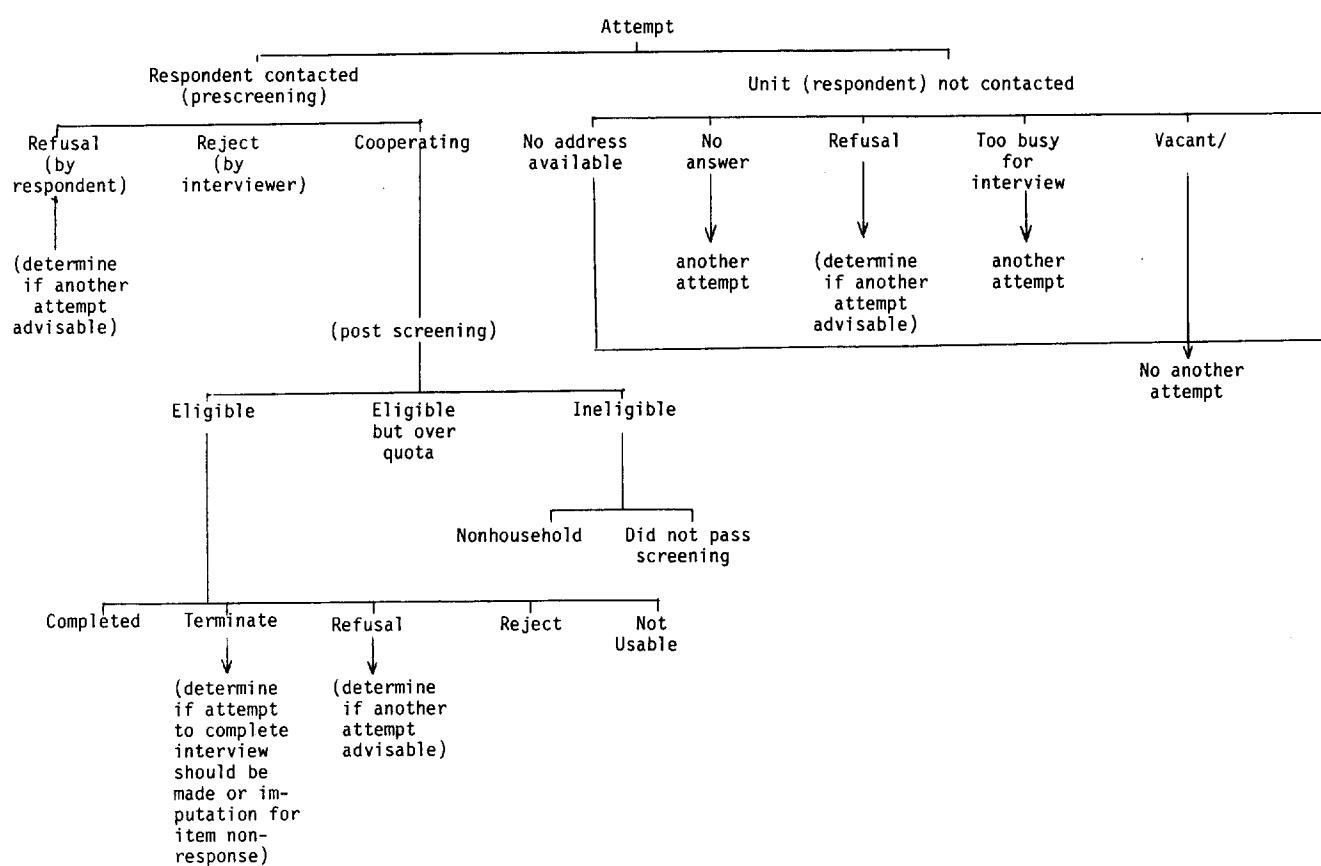
XI. REPORTING NONRESPONSE

In order to monitor nonresponse over time and for particular cases it is necessary to have a systematic and continuing record of response/nonresponse.

As survey data are collected there are many situations that determine a response or the lack of it.

Each situation has a reason behind its response/nonresponse outcome. Some sampling units may be refusals, others may be vacant and still others may not have been contacted at all. Thus, to examine all these cases carefully it is helpful to have for each survey a response outcome as it is suggested in table 13 for personal and telephone surveys.

TABLE 13
RESPONSE OUTCOME FOR PERSONAL AND TELEPHONE SURVEYS



Based on Table 13 Accountability Tables can be constructed whose purpose is to exhibit any reduction from the number of units in the sample to the units for which responses are made. The units are classified

into categories for exclusion, for ineligibility, for various types of response/nonresponse, etc. For example, for a particular survey on Accountability Table may look as follows:

TABLE 14
ACCOUNTABILITY TABLE

Name of Survey	:	Income Survey
Survey Population	:	
Target Population	:	
Type of Data Collection	:	Personal Interview
Selected Sample Size	:	
Number of Contacted Units	:	
Number of Eligible Units	:	
Number of Vacant Units	:	
Number of No one at Home	:	
Number of Refusals	:	
Refusals by Reasons	:	
Number of Temporary Absent	:	
Other type of nonresponse	:	

Accountability Tables may be classified by stratum, interviewing area, interviewer in order to facilitate the management of the survey, including making decisions on whether additional support is required for certain areas, interviewers on reassignment of units to other interviewers. Further accountability tables may serve two main purposes.

First, they include the ingredients from which nonresponse rates can be calculated. The second objective is to enable the management of the survey to be improved. For example, if accountability tables are kept up to date, then the management of the survey will be able to decide whether and where special efforts are required.

REFERENCES

- 1.- An investigation of the properties of raking ratio estimators with simple random sampling. *Survey Methodology Journal* (1977) pp. 62-83 by ARORA H.R. and BRACKSTONE G.J.
- 2.- Edit and Imputation in the Canadian Labour Force Survey. *Proceedings of ASA* (1978). pp. 425-430 by ASHRAF A. and MACREDIE I.
- 3.- Comparison of the biases of the "Hot Deck" Procedure with "equal weights" Imputation. Panel on Incomplete Data. U.S. National Research Council (1979) pp. 422-447 by BAILAR B.A. and BAILAR I.
- 4.- Nonresponse in complex multiphase surveys. *Proceedings of ASA* (1980) pp. 639-642 by DREW J.H. and FULLER W.
- 5.- Two efficient random imputation procedures. *Proceedings of ASA* (1981) by KALTON G. and KISH L.
- 6.- Compensating for Missing Survey Data. Research Report by KALTON G. Survey Research Centre, University of Michigan (1983).
- 7.- Models for Nonresponse in sample surveys (1982). *JASA* pp. 237-250 by LITTLE R.J.A.
- 8.- Imputation Methodology (1980) Panel on Incomplete Data. U.S. National Academy of Science. By PLATEK R. and GRAY G.B. Part V, Chapter 17, 18, 19, pp. 255-333.
- 9.- Some Aspects of Nonresponse Adjustments. *Survey Methodology* 1985 VOL. II NO. 1 pp. 1-14, by PLATEK R. and GRAY G.B.
- 10.- Some factors affecting nonresponse by PLATEK R. *Bulletin of the International Statistical Institute* 47(3), 347-366, also *Survey Methodology Journal* December 1977, VOL 3 NO. 2 pp. 191-214.
- 11.- On the definition of Response Rates by PLATEK R. and GRAY G.B. *Survey Methodology*, June 1986 VOL. 12 No. 1 pp. 17-27.
- 12.- The Inter action of refusal and non-at-home bias (1977) U.S. *Journal of Marketing Research* by WILCOX J.B.
- 13.- The use of post-stratification for adjustments of nonresponse bias. (1977) *Bulletin of ISI*.
- 14.- The use of ratio estimates to compensate for nonresponse bias in certain economic surveys. (1978) *Proceedings of ASA*.
- 15.- Weighted Hot-Deck imputation of medical expenditures based on a record check subsample (1981). *Proceedings of ASA*.

METODOLOGIA Y TRATAMIENTO DE LA NO-RESPUESTA

TRADUCCION: M.^a TERESA GARCIA DEL VALLE IRALA
CATEDRATICA DE ESTADISTICA
UNIVERSIDAD DEL PAIS VASCO
EUSKAL HERRIKO UNIBERTSITATEA

I. INTRODUCCION A LA NO-RESPUESTA

a) La información estadística se obtiene en gran parte de censos y encuestas. Cualquiera que sea su origen la información sufrirá las carencias debidas a la no-respuesta. Se reconoce en general que la no respuesta es una importante medida de la calidad de los datos, si bien esta importancia se tiene en cuenta en distintos grados. La no respuesta puede introducir sesgo en la estimación e incrementar la varianza muestral ya que reduce el tamaño de la muestra. Como ya sabemos la varianza muestral es inversamente proporcional al tamaño de la muestra. Así, una estimación basada en una m.a.s. con una tasa de respuesta del 80% tendrá una varianza muestral que es un 12,5% mayor que la varianza de la estimación correspondiente a una tasa del 90%. La relación entre el tamaño de la no-respuesta y el sesgo, aunque quizás más importante, es menos obvia ya que depende tanto de la magnitud de la no-respuesta como de la diferencia de características entre los que responden y los que no responden. Supongamos, sin embargo, que el sesgo debido a la no-respuesta es proporcional a la tasa de no-respuesta. Entonces, el porcentaje de sesgo es el mismo para una tasa dada de no respuesta, con independencia del tamaño de la muestra. Pero la varianza muestral se verá afectada

por los cambios en el tamaño de la muestra.

Consiguentemente es posible especular sobre la importancia relativa de los dos componentes, tanto en lo referente a la interpretación de los resultados de una encuesta cuanto sobre la manera de asignar el presupuesto en la fase de diseño. Para ilustrar esto tomemos el coeficiente de variación, que es una medida importante de la calidad de los datos. Cuando el tamaño de la muestra se incrementa, el coeficiente de variación de una estimación disminuye. Sin embargo el porcentaje de sesgo se mantiene mientras no haya un cambio en el tamaño de la no-respuesta. Así, en relación con el tamaño del coeficiente de variación, el sesgo asume un papel más importante e incluso dominante en lo referente a la estimación. Esto implica que en muestras grandes, generalmente diseñadas para obtener estimaciones a nivel nacional, los diseñadores han de valorar más el tamaño del sesgo que el tamaño de la varianza. Sin embargo en muestras pequeñas el sesgo debido a la no respuesta puede que no sea tan importante en relación con el tamaño de la varianza. En este caso, la varianza tendrá el papel dominante debido al incremento en su tamaño, causado por la no respuesta. Se debe tener en cuenta que todo

esto afecta también a la estimación por intervalo. En muestras pequeñas, los intervalos de confianza pueden contener al verdadero valor debido a una varianza muestral grande pero esto no suele suceder con muestras grandes donde la varianza muestral es relativamente pequeña.

Este sencillo análisis puede tener implicaciones prácticas importantes. Por ejemplo, si el propósito de una encuesta es obtener estimaciones preferentemente a nivel nacional, se debe prestar más atención a reducir sesgos. En otras palabras, deben emplearse más recursos en la atención a las posibles fuentes de sesgos: por ejemplo, al entrenamiento de los entrevistadores y al control del proceso de realización de los cuestionarios. Si las estimaciones se refieren a pequeñas áreas sería aconsejable hacer lo opuesto. Desafortunadamente los resultados de una encuesta se usan en la práctica en ambos niveles, por lo que las anteriores observaciones, aunque útiles, han de usarse prudentemente. Todavía y en un sentido más práctico en las diferentes etapas de una encuesta, la no respuesta o algunos de sus componentes pueden ser muy útiles a la hora de enjuiciar problemas operativos y valorar la validez de los datos.

b) Como base para el estudio de la no-respuesta vamos a reseñar varias etapas fundamentales que deben considerarse antes de proceder a la realización de una encuesta o censo. Primero, se debe identificar y definir cuidadosamente la población objetivo. Un ejemplo sencillo de ésta, lo constituye la población de una ciudad en un área metropolitana y ejemplos de marcos que la representen son las guías de la ciudad (con señas) y las guías telefónicas. Segundo, hay que definir un marco muestral o censal que represente a la población objetivo. El marco puede ser una lista, un área o una mezcla de ambos. Un ejemplo de un marco de áreas que represente a la población en una ciudad puede ser un mapa de la ciudad con sus calles y bloques. Las guías telefónicas no cubren a todos los habitantes y puede ser necesario utilizar un marco de áreas para hallar una muestra de la población residual excluida de la guía telefónica.

Seleccionada la muestra, hay que obtener los datos en un determinado intervalo de tiempo (por ejemplo una semana, un mes, ...) de forma de hallar las estadísticas que se correspondan con ese período de referencia específico. Cuanto más tiempo se emplee en la encuesta mayor será el desfase entre el momento en que se obtienen los datos de la encuesta y el período de referencia. El resultado puede ser un incremento en los errores

de respuesta, y en consecuencia un incremento tanto del sesgo, como de la varianza de la respuesta. Desde luego puede haber errores en la respuesta por otras causas además de la indicada y así puede haber errores debido a: i) interpretaciones no correctas de las preguntas, ii) confusiones del encuestador, iii) los encuestados han guardado pocos datos (por ejemplo pocos datos de impuestos), iv) los encuestados no pueden o no desean expresarse claramente durante la encuesta, v) respuestas incorrectas dadas a propósito por los encuestados (por ejemplo renta, actitud hacia cuestiones delicadas, ...). Estos problemas de errores en la respuesta aparecen sea cual fuere el método de obtención de los datos, es decir, i) entrevista personal, ii) entrevista telefónica, iii) encuesta por correo o cualquier combinación o variación de estos métodos. Los sentimientos personales del entrevistador se pueden transmitir a los encuestados de forma que se produzcan errores correlados. Las entrevistas telefónicas pueden aliviar pero no eliminar por completo esta tendencia a los errores correlados. Las encuestas postales son útiles en aquellos casos en los que no existan dificultades para la contestación del cuestionario; las hay de dos tipos: a) aquéllas en las que se pasan a recoger los cuestionarios, b) y aquéllas en las que los propios encuestados los devuelven por correo. Depende del tipo de intervención del entrevistador o de la ausencia de toda interacción el que pueda o no haber errores correlados. El entrevistador puede tener la oportunidad de ayudar a los que responden en caso de dificultad, evitando algunos errores de respuesta.

Además de los errores de respuesta pueden faltar datos debido a la no respuesta. La no-respuesta se da bajo cualquier forma de recogida de datos. Sin embargo la entrevista personal suele dar tasas más bajas de no respuesta que la telefónica. Las encuestas por correo suelen presentar la tasa más alta de no-respuesta. La no-respuesta a una pregunta particular suele deberse más a la complejidad de la encuesta o al diseño del cuestionario que al método de recogida de datos. Este tipo de no respuesta implica una necesidad de imputar valores y trae como consecuencia errores de imputación cuya extensión frecuentemente es difícil determinar.

Hay diversas maneras de tratar la no-respuesta, así como diferentes procedimientos de imputación de los datos que faltan, utilizándose éstos según las condiciones de la encuesta o los objetivos que se pretendan.

Un diagrama de flujos (tabla 1) resume las diferentes

etapas a través de las cuales se obtienen los datos de una encuesta, empezando con la población objetivo y terminando con la tabulación y la estimación. Entre estas dos etapas, hay muchos problemas que pueden dar lugar a omisión de unidades, falta de cobertura, datos que faltan (no respuesta) y errores en las preguntas. Desde luego no son las únicas etapas o los únicos problemas en la obtención de datos de encuestas. Aunque la mayoría de los problemas relacionados con la calidad de los datos se corresponden con las etapas y problemas presentados en el diagrama.

Como se señala en la Tabla 1, los problemas actuales o potenciales de listados o marcos por áreas incompletos se producen antes de la selección de la muestra. Esto produce errores de cobertura en marcos muestrales cuando no representan adecuadamente la población objetivo en el momento de la encuesta. Lo que ocurre por las siguientes razones: i) Unidades de la población objeto que no están en el marco. (Falta de cobertura). ii) Unidades que no están en la población objeto y aparecen en el marco. (Exceso de cobertura). iii) Hay unidades que se repiten en el marco muestral (duplicación que produce exceso de cobertura). Habrá también falta de cobertura en la muestra que se basa en un marco muestral en el que no haya posibilidad de seleccionar unidades de la población objetivo al estar omitidas del marco muestral. Esto se indica conectando líneas desde errores en el listado del marco (caja 5) y errores en el marco por áreas (caja 6) a la selección muestral (caja 7). En general, las unidades muestrales perdidas debidas a la falta de cobertura no se conocen durante la operación de recogida de datos. Generalmente, sólo después de haberse completado la recogida de datos (caja 8) y de haberlos procesado, lo cual incluye las etapas desde la depuración previa del cuestionario hasta la estimación, se revela la falta o el exceso de cobertura en una o más de las células resultantes de la post-estratificación; por ello sólo hay una línea conectando la imputación por falta de cobertura (caja 17) a la estimación (caja 21).

Otra rama en el diagrama de flujos procede de la operación de coleccionar datos (caja 8), cuando puede ocurrir la respuesta/no-respuesta de una unidad (cajas 10 y 11 respectivamente), o la unidad puede ser no existente o estar incorrectamente incluida (caja 9), contribuyendo a un exceso de cobertura (caja 13). La respuesta de una unidad va seguida de la depuración previa del cuestionario (caja 12), que se puede revelar (i) completo y consistente (ii) algunos blancos o inconsistencias (no respuesta a algunas preguntas) o (iii)

un cuestionario totalmente inútil (cajas 14, 15, 16). Mientras que las no respuestas a algunos items o inconsistencias suelen ser detectables en esta etapa, es imposible en general detectar errores en las respuestas si no hay alguna razón lógica para ello. Así que se procedería directamente a la estimación a pesar de que hubiese errores en la respuesta (caja 18). Se pueden producir errores de imputación debidos a la imputación realizada ante la no-respuesta a una pregunta o de toda una unidad (caja 19) y también éstos pueden no ser detectados. Consecuentemente se procede a la estimación (caja 21).

A la no respuesta de una unidad o a un ítem (cajas 11, 15 y 16) le sigue la imputación de los datos que faltan (caja 19). Se debe hacer constar que dependiendo del procedimiento de imputación y de la disponibilidad de fuentes de datos externas, se puede realizar la imputación inmediatamente después de haber detectado no-respuesta (por ejemplo, fichero caliente o por otros métodos de duplicación). Los otros métodos de imputación deben esperar a la etapa de estimación (por ejemplo, ajuste de las ponderaciones para alargar la muestra, sustitución por los datos del mes pasado), de ahí que haya líneas de puntos conectando la imputación por no respuesta y la estimación.

Cualquier procedimiento de imputación (caja 19) puede dar lugar a un error de imputación, que se puede presentar a niveles micro o sea en datos que pertenecen a unidades muestrales, y posiblemente también niveles macro si los errores de imputación después de ponderar los datos no se cancelan; de ahí que haya una línea conectando ambas direcciones, estimación y errores de imputación.

La estimación (caja 21) y tabulación (caja 22) culminan con un análisis de los errores cuadráticos medios (no se incluyen en el diagrama de flujos), que incluyen al sesgo proveniente tanto de errores de respuesta como de imputación a la varianza muestral (al menos que los datos se hayan obtenido mediante censo), y a la varianza no muestral proveniente de variaciones en los errores de respuesta e imputación.

Los datos de encuestas suelen requerir ajustes por diversas diferencias causadas tanto por valores que faltan como por errores de respuesta. En la Tabla 2 se muestra un modelo de imputación/error en la respuesta que ilustra las etapas en las que una unidad particular contendría valores específicos. Se puede ver que de una característica encuestada y se den diferentes tipos

errores dependiendo de lo que ocurre en la unidad. Por ejemplo, si la unidad responde (caja 2) pero falla al responder con respecto a la característica y (no

respuesta a una pregunta, caja 5) se puede dar un error de imputación $NR_{iy}^{\epsilon 1}$ en el valor imputado Z_{iy}^I (caja 7).

t_i = Según la unidad sea o no seleccionada

δ_i = Según la unidad responde o no

δ_{iy} = Según se consiga respuesta a un ítem o no

y_i = Valor verdadero

Y_i = Valor observado

R_i^{ϵ} = Error de respuesta

Z_{iy} = Valor imputado

NR_i^{ϵ} = Error de imputación

R_i^B = Sesgo de respuesta

NR_i^{BI} = Sesgo de imputación

$R_i^{\sigma^2}$ = Varianza de respuesta

$NR_i^{\sigma^2}$ = Varianza de imputación

TABLA 1
DIAGRAMA DE FLUJO PARA LOS DATOS DE UNA ENCUESTA

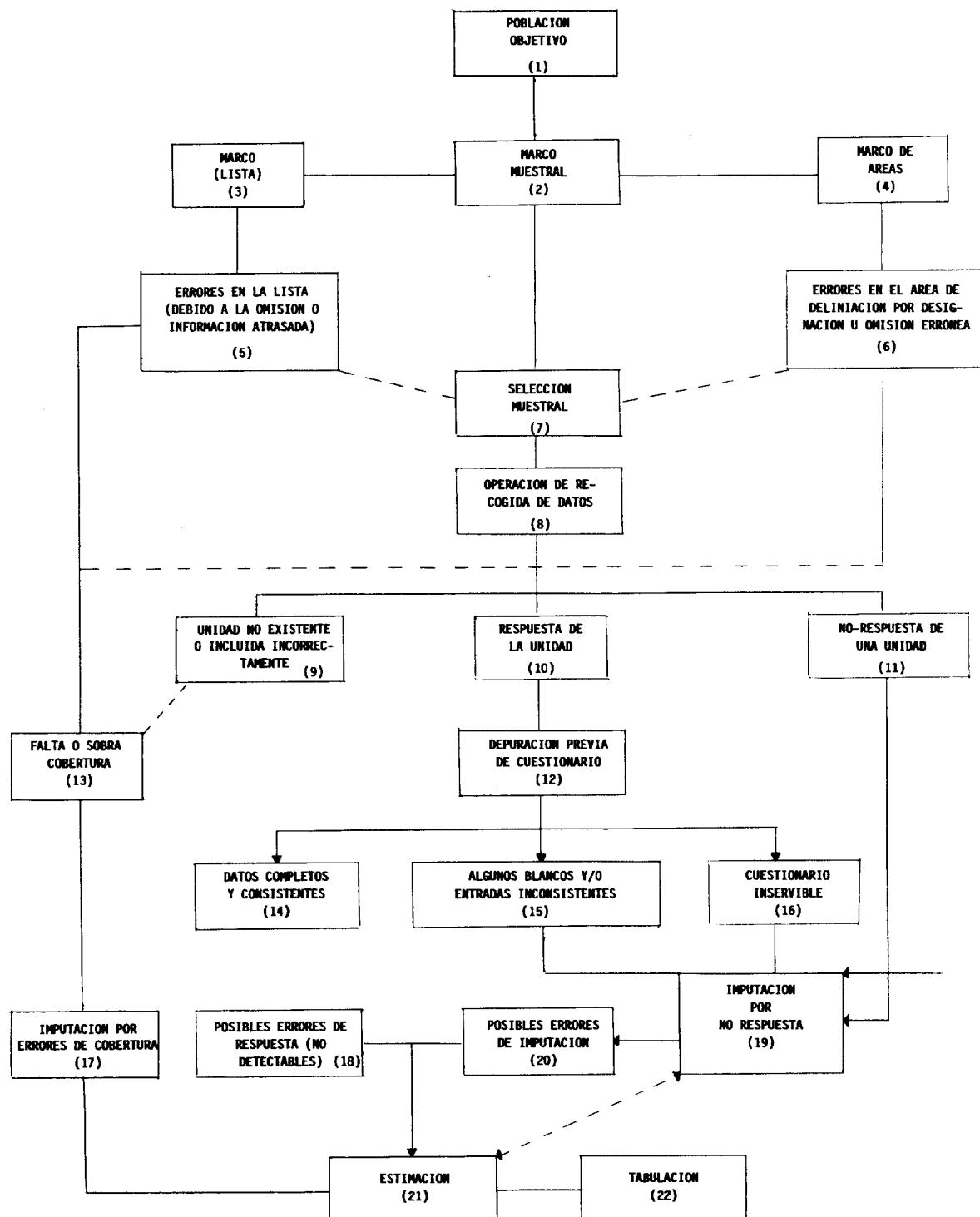
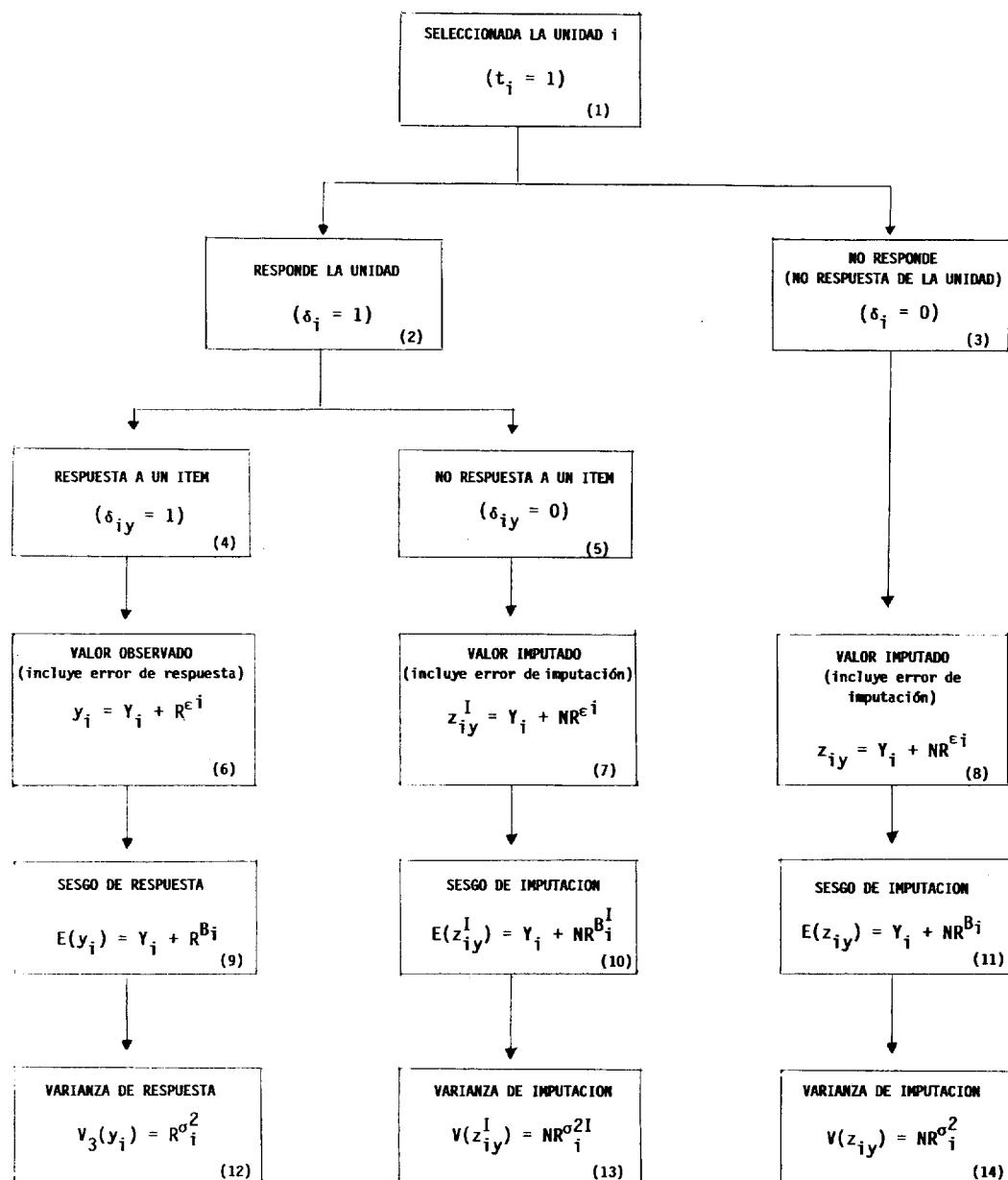


TABLA 2
MODELO DE ERROR DE IMPUTACION/RESPUESTA
(Nivel de unidad)



$$u_{iy} = \delta_i [\delta_{iy} y_i (1 - \delta_{iy}) z_{iy}^I] + (1 - \delta_i) z_{iy}$$

= valores observados o imputados para datos que faltan por no respuesta de una unidad a un ítem

II. LA DEFINICION DE LAS TASAS DE RESPUESTA/NO RESPUESTA

Recientemente y en muchos países se ha observado un incremento de la no respuesta. Consecuentemente, hay una mayor necesidad que antes de controlar y vigilar las tasas de no respuesta, hacer comparaciones entre encuestas, países, organizaciones de encuestas y también asegurar cierto grado de comparabilidad. Ha habido intentos de estandarizar la definición de tasa de respuesta y su complemento, la tasa de no respuesta; véase por ejemplo, Kviz (1977), Cannell (1978). Wiseman y MacDonald (1980) describen problemas provenientes de definiciones inconsistentes de tasas de respuesta relacionadas con encuestas telefónicas.

Hay también problemas de terminología inconsistente respecto a la respuesta/no respuesta en las encuestas. Se han utilizado términos tales como tasa de cumplimentación, tasa de contacto y tasa de falta o exceso de cobertura en trabajos y artículos que tratan la recogida de datos. Mientras que estos términos son claros en un trabajo individual, suelen resultar confusos y sujetos a interpretaciones conflictivas cuando se comparan varios trabajos.

Al considerar los problemas de respuesta/no respuesta, hay que hacer una distinción entre tasas de no respuesta de una unidad o a una pregunta. La no respuesta de unidades se da en el nivel en el que se recogen datos durante el primer contacto. Ejemplos de este nivel pueden ser una casa, individuo, almacén, establecimiento. Sin embargo, en caso de muestreo en varias etapas, puede haber no respuesta de todas las unidades dentro de un conglomerado o incluso de unidades muestrales primarias (psu) así que la no respuesta debida a una unidad se puede aplicar al conglomerado seleccionado o psu así como a una casa o individuo.

La no respuesta a una pregunta se da ya dentro del cuestionario, en este caso hay preguntas contestadas y preguntas no contestadas.

Sin embargo si una unidad falla en responder, automáticamente no responde a ninguna pregunta. La no respuesta de unidades y la no respuesta a preguntas deben ser tratadas de forma diferentes.

Las tasas de respuesta pueden referirse a toda la muestra o a una parte de la muestra tales como áreas que dependen del diseño o también se pueden referir a áreas administrativas como las asignadas a un entrevistador o a un grupo de entrevistadores supervisados por un inspector de campo.

a) Componentes de respuesta/no respuesta

Para poder definir las diversas tasas de respuesta y discutir sus usos y aplicaciones, es necesario clasificar la población objeto (de la muestra o el censo), en varios componentes según el tipo de respuesta/no respuesta. La tabla 3 consigue este propósito, indicando la mayoría de los componentes importantes de la encuesta que se usan en las tasas. Una vez que se define la población objetivo (caja 1) de una encuesta, se determina el marco muestral de N unidades (caja 2).

Para realizar una encuesta es necesario decidir tanto el procedimiento de recogida de datos como el diseño muestral (caja 3).

$t_i = 1$ ó 0 según que la unidad i sea seleccionada o no

Σ = sumatorio sobre N
 i

Para muestras de tamaño n , $\Sigma t_i e_i$ son elegibles (caja 5) y $\Sigma t_i(1 - e_i)$ son inelegibles (Caja 5).

$e_i = 1$ ó 0 según sea la unidad i elegible o no.

A veces no se determina el criterio de elegibilidad si la unidad no puede ser contactada, mientras que otras veces este criterio es obvio por razones físicas, tales como casas vacías/no vacías en una encuesta sobre hogares.

Las unidades no elegibles $\Sigma t_i(1 - e_i)$ (caja 5) se pueden dividir en dos grupos: i) unidades no entrevistadas y que no debían serlo $\Sigma t_i(1 - e_i)(1 - \delta_i)$ (caja 6). ii) unidades incorrectamente entrevistadas $\Sigma t_i(1 - e_i)\delta_i$ (caja 7). Se espera que no haya unidades en la caja 7 o que sean muy pocas. Sin embargo, si las hay, deben ser suprimidas de la muestra $\delta_i = 1$ ó 0 según la unidad i responda o no.

Las $\Sigma t_i e_i$ unidades (caja 8) se dividen en $\Sigma t_i e_i \delta_i$ unidades que responden (caja 11) estas últimas proveen de información que no se puede usar en la encuesta y apenas se sabe nada de estas unidades excepto quizás su localización geográfica.

Las unidades que responden $\Sigma t_i e_i \delta_i$ se pueden dividir primero entre $\Sigma t_i e_i \delta_i$ y $\Sigma \delta_i$ unidades, que han contestado a todas las preguntas, pero con posibles errores

de respuesta (caja 9) y $\sum t_i e_i \delta_i [1 - \pi(\delta_{iy})]$ unidades con no respuesta a algunas preguntas de al menos alguna característica pero no en todas (caja 10). Aquí $\delta_{iy} = 1$ ó 0 según que la unidad entrevistada responda o no a la pregunta o característica y.

En (caja 9), $\delta_{iy} = 1$ para la unidad i y para todos los ítems, mientras que en (caja 10), $\delta_{iy} = 0$ para algunas pero no para todas las preguntas. Para una pregunta particular y algunas de las $\sum t_i e_i \delta_{iy}$ personas que han respondido (caja 12) provienen del grupo que ha respondido a todas las preguntas (caja 9), mientras que el resto proviene de esas unidades que responden pero no a todas las preguntas (aunque si a la pregunta y). Los $t_i e_i \delta_i (1 - \delta_{iy})$ que no responden a la pregunta y provienen de aquellas unidades que responden pero que tienen preguntas (una o varias) sin respuesta (caja 10) entre las preguntas sin respuesta está la y.

Los que responden al ítem y (caja 12) se pueden descomponer en tres componentes (i) aquellas unidades con ítem y libre de errores de respuesta (ii) aquéllas a los que se les ha detectado error en la respuesta para el ítem y (iii) aquéllos con error en la respuesta y pero no detectado (cajas 15, 16 A y 16 B respectivamente).

b) Definición de diversas tasas

La muestra de $n = \sum t_i$ unidades descompuesta en la tabla 1 de la sección (2) en unidades elegibles, unidades que responden y que no responden, rechazos, preguntas con respuesta y sin respuesta, etc., conducen a diferentes tipos de tasas que pasamos a definir seguidamente. Para cada tasa, el numerador es un subconjunto particular del denominador. Cuando es posible, la tasa se define en términos de las unidades tal y como vienen en la tabla 3.

(i) Tasa de elegibilidad

La tasa de elegibilidad viene dada por:

$$\bar{e} = \frac{\sum t_i e_i}{\sum t_i} = \text{Caja 8} / \text{Caja 4}$$

Wiseman y McDonald (1980) usaron el término *tasa de indicencia* pero aplicaron el término únicamente a personas seleccionadas de muestras telefónicas que contestaron en la fase de criba para determinar su elegibilidad para una encuesta.

La tasa de elegibilidad demuestra la calidad del diseño de una encuesta para seleccionar unidades elegi-

bles a partir de un marco, cuando la elegibilidad no se puede determinar sin un contacto u observación rápida. En la fase de criba, la tasa provee de información para determinar cuántas unidades elegibles resultarán en la etapa de recogida de datos de una encuesta. Así, la tasa se puede emplear en la fase de diseño si hay datos sobre elegibilidad disponibles de estudios anteriores. Dependiendo de la naturaleza y del procedimiento de la encuesta, la elegibilidad de las unidades puede no ser determinable entre unidades que no han sido contactadas o incluso entre unidades que rehusaron. Hay dos alternativas para definir la tasa de elegibilidad y las tasas de respuesta (que las definiremos más tarde) entre unidades elegibles. Se puede suponer, en el caso de estimaciones conservadoras de la calidad de los datos y de la calidad del procedimiento de recogida de datos de una encuesta, que todos los no-contactos y rechazos serían elegibles incluso aunque realistamente la proporción de unidades elegibles entre los que no responden es a menudo más baja que entre los que responden y no responden; para los cuales se puede obtener una cota inferior para la tasa de respuesta y una cota superior para la tasa de elegibilidad.

Alternativamente, se puede suponer la misma proporción de unidades elegibles entre unidades cuya elegibilidad no se puede determinar como entre aquéllos cuya elegibilidad se conoce. Bajo este supuesto tendremos una ligera sobreestimación de la tasa de elegibilidad y de otras tasas.

(ii) Tasas de respuesta y de cumplimentación

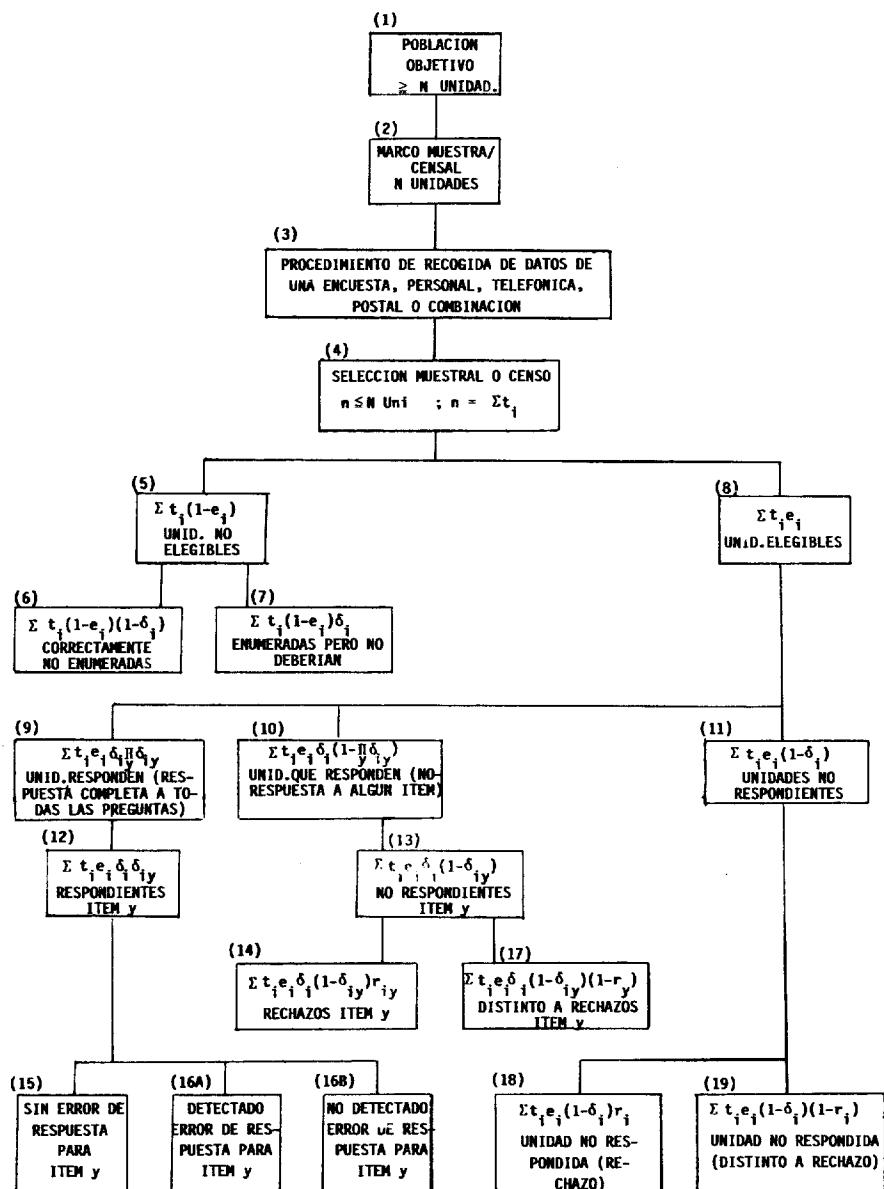
a) De acuerdo con una de las dos definiciones alternativas que propone el comité federal estadounidense sobre metodología estadística (1978), la tasa de respuesta es el porcentaje de muestra elegible de la cual se obtiene información. La tasa de respuesta se define:

$$\begin{aligned} R_{(1)} &= \frac{\sum t_i e_i \delta_i}{\sum t_i e_i} \\ &= (\text{Caja 9}) + (\text{Caja 10}) / (\text{Caja 8}) \end{aligned}$$

La anterior es la tasa de respuesta comúnmente empleada en la práctica ya que proporciona el porcentaje de muestra para el cual se obtienen ciertos datos útiles, una vez suprimidas las unidades no elegibles. En el denominador se incluyen todos los tipos de no-respuesta de las unidades elegibles.

Se usa la inversa de la tasa anterior en una célula ajustada como una ponderación ajustada en compensación

TABLA 3
COMPONENTES DE RESPUESTA/NO RESPUESTA



$$e_i = 1,0 \quad (\text{unidad elegible/no elegible})$$

$$\delta_{iy} = 1,0 \quad (\text{respuesta/no respuesta item } y)$$

$$t_i = 1,0 \quad (\text{seleccionada/no seleccionada})$$

$$r_i = 1,0 \quad (\text{de acuerdo con si la unidad la rechaza o no})$$

$$\delta_i = 1,0 \quad (\text{respuesta/no respuesta unidad})$$

Para $r_i = 0$, (principalmente "nadie en casa" o ausente temporalmente)

de datos que faltan debido a unidades que no responden; por ejemplo, tales tasas se usan frecuentemente en la LFS canadiense como ponderaciones de ajuste (véase Platek y Gray, 1985).

La tasa anterior o su complemento la tasa de no-respuesta, se usan frecuentemente para valoraciones administrativas y operativas de la organización de una encuesta. Estas tasas se usan también para valorar la habilidad de un encuestador para contactar con los encuestados y conseguir información útil, así por ejemplo, se estudian las tasas de respuesta/no respuesta que ha tenido cada entrevistador. La tasa de no respuesta incluye tanto los rechazos que pueden controlarse con diplomacia y habilidad como los no contactos los cuales no quedan bajo el control del entrevistador. Por estas razones siempre que se puede se subdividen las tasas de no respuesta.

Kviz (1977) definió la tasa de cumplimentación como una tasa similar a la anterior, que incluye la muestra completa en el denominador. Esta tasa provee de una estimación más conservadora de la calidad ya que incluye en el denominador incluso a unidades elegibles vacías. Otra definición:

$$R_{(2)} = \frac{\sum_i t_i \delta_i}{\sum_i t_i [\delta_i + (1 - \delta_i)r_i]}$$

es el porcentaje de veces en los que un encuestador obtiene entrevistas en las señas muestrales, cuando logra contactar. La unidad i rechaza o no según $r = 1$ ó 0. Esta es la definición de tasa de cumplimentación dada por O'Neill, Groves y Cannel (1979). Si se puede determinar la elegibilidad de todas las unidades que se contactan entonces se puede dar quizás una definición mejor en relación con las unidades elegibles.

$$\begin{aligned} R_{(3)} &= \frac{\sum_i t_i \delta_i e_i}{\sum_i t_i e_i [\delta_i + (1 - \delta_i)r_i]} \\ &= (\text{Caja 9}) + (\text{Caja 10}) / (\text{Caja 9}) + (\text{Caja 10}) + (\text{Caja 18}) \end{aligned}$$

donde e_i , el criterio de elegibilidad se define después de la Tabla 1.

Las tasas anteriores pueden ser útiles en encuestas personales y telefónicas en las que la no respuesta incluye tanto el no contacto como el rechazo. Estas tasas no son prácticas en encuestas postales a menos que haya un seguimiento telefónico o personal de las no-respuestas; en la mayoría de las encuestas postales

puras se dan tanto respuestas como no-respuestas sin que se conozcan sus causas. Cuando las tasas anteriores son significativas, miden la habilidad del método de recogida de datos para obtener la colaboración de los encuestados responsables de las unidades seleccionadas, siempre que hayan podido ser contactados. Los que no han podido ser contactados suelen estar fuera del control de los encuestadores y dependiendo del procedimiento seguido en la encuesta suelen ser normalmente eliminados de las tasas.

Klecka y Tuchfarber (1979) definieron la tasa de respuesta como tasa de cumplimentación ya que supusieron, siendo quizás este supuesto un tanto irreal, que todos los rechazos eran elegibles. La tasa de cumplimentación sería entonces una estimación conservadora del comportamiento del método de recogida de datos respecto a lograr la cooperación de las unidades elegibles.

Alternativamente se puede suponer que la proporción de elegibilidad entre los que rechazaron contestar es la misma que entre los que completaron su cuestionario, y otras unidades cuyo criterio de elegibilidad se conoce.

(iii) Tasas de contacto

Hauck (1974) definió la tasa de contacto como el porcentaje de unidades muestrales que son contactadas y viene dado por:

$$R_{(4)} = \frac{\text{entrevistas terminadas} + \text{rechazos (contactados)}}{\text{entrev. termin.} + \text{rechazo (contact. y no contact.)}}$$

supuso que los no contactos eran elegibles, proporcionando así una estimación conservadora del éxito en contactar unidades muestrales. En rechazos se incluyen las entrevistas incompletas que son rechazos para algunas preguntas (Caja 10 de la Tabla 3).

La expresión algebraica para la tasa de contacto viene dada por:

$$R_{(4)} = \frac{\sum_i t_i \delta_i e_i + \sum_i t_i (1 - \delta_i) r_i \hat{e}_i}{\sum_i t_i \delta_i \hat{e}_i + t_i (1 - \delta_i) r_i \hat{e}_i + \sum_i t_i (1 - \delta_i) (1 - r_i) \hat{e}_i}$$

$$= \frac{(Caja 9) + (Caja 10) + (Caja 18)}{(Caja 9) + (Caja 10) + (Caja 18) + (Caja 19)}$$

donde

$\hat{e}_i = e_i = 1$ ó 0 si se conoce el criterio de elegibilidad, y para los no contactados.

$\hat{e}_i = 1$, de acuerdo con la definición de Hauck.

$\bar{e} = \bar{e}$, tasa promedio de elegibilidad entre aquellas unidades cuyo criterio de elegibilidad se conoce.

La tasa de contacto mide la habilidad de la organización o de los encuestadores para contactar, tengan o no éxito en conseguir la cooperación de los que responden.

(iv) Tasa de rechazo (tasa de no-rechazo)

Hauck (1974) por un lado, Wiseman y MacDonald (1980) por otro dan dos definiciones de las tasas de rechazo:

$$\begin{aligned} F_1 &= \frac{\text{número de rechazos}}{\text{número de entrevistas completas y rechazos}} \\ &= \frac{\sum_i t_i \hat{e}_i (1 - \delta_i)}{\sum_i t_i \delta_i + \sum_i t_i \hat{e}_i (1 - \delta_i) r_i} \\ &= \frac{(Caja 18)}{(Caja 9) + (Caja 10) + (Caja 18)} = 1 - R(3) \end{aligned}$$

y

$$\begin{aligned} F_2 &= \frac{\text{número de rechazos}}{\text{número de todas las unidades seleccionadas}} \\ &= \frac{\sum_i t_i (1 - \delta_i) r_i}{\sum_i t_i} \\ &= (Caja 18) / (Caja 4) \end{aligned}$$

Teniendo en cuenta el criterio de elegibilidad, la tasa de rechazo viene dada por:

$$F_3 = \frac{\sum_i t_i \hat{e}_i (1 - \delta_i)}{\sum_i t_i \hat{e}_i}$$

$$= (Caja 18) / (Caja 8)$$

La tasa de rechazo mide la falta de capacidad de la organización de la encuesta o de los entrevistadores en conseguir la cooperación de las unidades para proveer datos de encuesta útiles. Y esto en relación a las unidades contactadas, en relación a la muestra total o en relación a la muestra elegible. Se puede desear determinar una tasa de rechazo pura sin que aparezcan las unidades no contactadas que a menudo están fuera del control de los entrevistadores y de esta forma estudiar la eficiencia del cuestionario o el efecto que un determinado tema produce en el grado de cooperación de los entrevistados. Alternativamente, se puede preferir examinar la tasa de rechazo como una de las componentes de la no-respuesta en su conjunto.

(V) Tasa de respuesta/no-respuesta para cada pregunta

Un diseño complejo del cuestionario puede dar lugar a no-respuestas a preguntas específicas por otras razones distintas al rechazo, como se señaló en la Caja 17. Un pregunta personal o controvertida o el querer acabar la entrevista puede dar lugar a un rechazo en proveer datos para una pregunta específica como se señala en la Caja 14.

Así, se puede medir la tasa de no-respuesta general al ítem **y** en relación con todas las unidades a las que se responde y viene dada por:

$$R_y = \frac{(Caja 13)}{(Caja 9) + (Caja 10)}$$

o si el ítem **y** tiene importancia únicamente entre ciertas unidades encuestadas (cuestionarios) pero no para todos, se puede medir la no-respuesta del ítem en relación a aquellas unidades que responden **y** para las cuales es relevante la cuestión **y** (elegibles).

Consecuentemente, se pueden definir para cada pregunta las tasas de respuesta, no respuesta y elegibilidad de manera análoga o como se hizo con las unidades, reemplazando en las tasas el número de unidades (elegibles/no elegibles) (responden/no responden) por el número de unidades que responden (elegibles o relevantes para una pregunta **y**, irrelevantes, respondan o no respondan al ítem **y**, ...) respectivamente.

La mayoría de las tasas definidas para las unidades y diferentes a las tasas de contacto tienen su versión para preguntas sin más que hacer las sustituciones apropiadas en las expresiones. Suele ser más difícil hacer constar las razones para la no respuesta a un ítem que las razones para la no-respuesta de unidades, ya que a menudo la no-respuesta a un ítem se detecta sólo a través de una rutina de depuración e imputación.

(vi) Tasas ponderadas y tasas características

En el caso de muestras con diferentes ponderaciones muestrales π_i^{-1} para las unidades, como por ejemplo, en el muestreo proporcional al tamaño (pps), todas las tasas anteriores se pueden definir como tasas ponderadas aplicando la ponderación muestral π_i^{-1} junto a la variable indicador t_i en cada una de las expresiones. En el caso de muestras autoponderadas en un área o clase para la que se calculan las tasas, los pesos muestrales son redundantes.

Sin embargo, en el muestreo (pps) en la etapa final suele suceder que las unidades grandes responden más que las pequeñas y las tasas ponderadas de respuesta, con pesos más pequeños que los muestrales aplicados a las unidades grandes en vez de a las pequeñas, suelen ser menores que las no ponderadas, que se basan en el conteo de unidades como se muestra en la Tabla 3.

Las tasas de respuesta ponderadas estiman la proporción de población que habría respondido a la encuesta bajo condiciones de encuesta muy similares, mientras que las tasas de respuesta no ponderadas proveen de una medida del comportamiento de un conjunto de datos de una muestra o submuestra que pertenecen a un área o clase específicas.

Al estimar la tasa de no-respuesta para toda la población en lugar de para la muestra como lo hacen las tasas no ponderadas, las tasas ponderadas pueden proporcionar información confusa sobre la calidad de los datos ya que se puede distorsionar la distribución de la característica en la muestra. La ventaja de las tasas ponderadas es que se añaden unidades a los niveles de población en lugar de a los muestrales, de modo que, se obtiene una estimación de la tasa que se obtendría a niveles censales bajo condiciones similares de recogida de datos.

(vii) Tasas del tamaño de respuesta/no-respuesta

Las tasas de no respuesta varían de encuesta a encuesta. Algunas encuestas tienen tasas de no-respuesta

por encima del 50% y otras sólo del 4% al 6%. Es importante darse cuenta que es el objetivo de la encuesta el que determina si las tasas de respuesta son muy altas o muy bajas. Si el objetivo de la encuesta es estimar un ítem que sólo se da en un 10% en la población, entonces una tasa de no-respuesta del 5% le afectaría considerablemente. Por otro lado, se puede pensar en una situación en la que una no-respuesta alta no es necesariamente inútil. Como todos los estadísticos, las tasas de respuesta tienen varianza muestral y no muestral y bajo un modelo de probabilidad de respuesta (Platek, Gray 1983) se puede hallar una expresión para la varianza de esta tasa.

Una aproximación a la varianza de la tasa de respuesta, según cualquier definición viene dada por:

$$V(\hat{R}) = \frac{\hat{R}(1 - \hat{R})}{n} F$$

donde $R = E_1 E_2 (\hat{R})$; y E define la esperanza.

E_1 = valor esperado en la distribución muestral

E_2 = valor esperado en las diferentes formas de no respuesta de las unidades seleccionadas.

n = tamaño muestral

F = efecto del diseño

F puede valer entre 1 y 2 dependiendo del diseño muestral y de la forma de realizar el cuestionario. Así, F podría estar cercana a 1 si el diseño es muestreo aleatorio simple y el cuestionario se envía por correo.

En las encuestas telefónicas podría estar cercano a 1,5 debido al efecto conglomerado al tener un entrevistador una asignación determinada. En las entrevistas personales puede valer 2 si se combinan el método de entrevista con un diseño por conglomerados.

III. NO-RESPUESTA EN LA ENCUESTA SOBRE FUERZA DE TRABAJO

Aunque la encuesta sobre la fuerza de trabajo en Canadá es una encuesta consolidada y controlada, tiene un problema constante en la no-respuesta. Se guarda un registro detallado de la no-respuesta total, y ésta se puede partitionar en varios componentes cada uno de ellos tiene causas diferentes y requiere un tratamiento distinto. Se pueden reconocer los siguientes: i) Familias ausentes temporalmente; ii) Nadie en casa; iii) Rechazo; iv) Ningún entrevistador disponible; v) Malas condiciones climatológicas; vi) Otras causas.

i) Comportamiento de la no-respuesta en el tiempo

En la Tabla 4, se muestran algunas tendencias en el comportamiento de las tasas de no-respuesta en la LFS.

La tasa de no-respuesta global se incrementa agudamente en el mes de Julio y algo menos en Agosto. Esto se debe a ausencias temporales por vacaciones. El tamaño de la no-respuesta viene determinado sobre todo por ausencias temporales y nadie en casa. Las tasas de rechazo han sido bastante regulares durante varios años con alguna tendencia reciente a la baja y aparecen bastante altas en los meses de verano. Una tendencia significativa es que la encuesta revisada sobre fuerza de trabajo en 1975 tuvo una tasa de no-respuesta relativamente alta. Esto se debió probablemente a que se renovó casi por completo el equipo de entrevistadores y a que los supervisores se encontraron sobrecargados: nuevos procedimientos, nuevas muestras y duros entrenamientos. Cuando la encuesta se asentó y los supervisores pudieron dedicar más tiempo a la tasa de respuesta, la no-respuesta se redujo a niveles más aceptables.

TABLA 4
TASAS DE NO RESPUESTA EN LA LFS DE CANADA POR COMPONENTES

	AÑO	MES	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
	Total													
	1973		7.3	7.2	6.8	7.9	7.0	8.4	15.1	10.9	6.5	5.7	5.2	6.6
A	1974		6.0	6.0	6.4	8.3	7.0	6.8	10.4	8.8	5.6	5.5	4.3	4.6
-	<u>1975</u>		4.3	4.7	4.6	4.7	4.7	5.8	7.6	6.3	4.3	4.5	4.3	5.3
	1975		10.2	9.3	8.7	7.5	7.9	8.9	13.2	11.0	7.5	7.5	7.0	7.6
B	1976		7.9	7.7	7.3	8.8	9.2	8.0	11.7	9.2	5.8	5.9	6.0	5.3
	Ausente													
	Temporal													
	1973		1.8	2.2	1.9	2.4	1.8	3.8	9.1	5.6	1.6	1.3	1.2	1.7
A	1974		1.7	1.8	1.9	2.0	1.5	2.0	6.1	4.7	2.0	1.7	1.0	1.4
-	<u>1975</u>		1.4	1.6	1.6	1.2	1.2	2.2	4.2	3.0	1.1	0.9	0.7	1.2
	1975		2.0	2.2	2.1	1.5	1.6	2.6	7.1	5.0	1.7	1.4	1.2	1.4
B	1976		1.9	2.1	2.3	2.3	1.8	2.4	6.2	4.1	1.7	1.4	1.3	1.3
	Nadie en casa													
	1973		2.5	2.1	2.0	2.6	2.5	2.7	3.2	2.3	2.1	1.9	1.6	2.0
A	1974		1.5	1.7	1.8	2.8	1.9	1.8	1.7	1.7	1.4	1.7	1.4	1.2
-	<u>1975</u>		1.0	0.9	1.0	1.2	1.1	1.3	1.2	1.2	1.1	1.4	1.2	1.6
	1975		2.7	2.2	1.8	1.5	1.9	1.8	2.2	2.1	1.8	1.7	1.7	1.8
B	1976		2.2	1.9	1.5	2.8	2.8	2.4	2.4	2.2	1.7	2.0	2.2	1.8
	Rechazo													
	1973		1.7	1.9	1.9	2.0	2.0	1.9	1.9	2.3	2.1	2.0	1.9	1.7
A	1974		1.6	1.6	1.7	2.1	2.4	2.3	2.1	1.9	1.6	1.4	1.3	1.2
-	<u>1975</u>		1.2	1.2	1.2	1.4	1.6	1.4	1.4	1.3	1.3	1.2	1.4	1.3
	1975		2.1	2.2	2.1	2.0	1.9	1.8	1.7	1.6	1.6	1.5	1.5	1.4
B	1976		1.5	1.6	1.6	1.9	2.5	2.4	2.0	1.9	1.6	1.6	1.7	1.6
	Otras													
	1973		1.3	1.0	1.0	0.9	0.7	0.5	0.9	0.7	0.7	0.5	0.5	1.2
A	1974		1.2	1.0	0.9	1.4	1.2	0.7	0.5	0.5	0.6	0.7	0.6	0.8
-	<u>1975</u>		0.7	1.0	0.8	0.9	0.8	0.9	0.8	0.8	0.8	1.0	1.0	1.2
	1975		3.4	2.7	2.7	2.5	2.5	2.5	2.2	2.3	2.4	2.9	2.6	3.0
B	1976		2.3	2.1	1.9	1.8	2.1	1.6	1.1	1.0	0.8	0.9	0.8	0.6

A Son tasas de la antigua encuesta sobre fuerza de trabajo

B Son tasas de la nueva encuesta sobre fuerzas de trabajo

TABLA 5

TASA DE NO RESPUESTA EN LA LFS DE CANADA POR COMPONENTES

AÑO	MES	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
Total													
1977		6.1	5.9	5.8	4.5	5.8	5.7	8.4	5.6	4.6	4.3	4.3	4.2
1978		5.0	5.2	5.7	5.3	5.7	5.1	7.9	5.9	4.7	4.6	5.0	4.5
1979		5.4	5.8	5.8	5.1	5.1	5.2	7.5	5.9	4.9	4.5	4.2	4.8
1980		5.3	5.4	5.2	5.3	5.9	6.1	7.6	5.9	4.7	4.3	4.2	4.4
Ausente													
Temporal													
1977		1.9	1.9	2.2	1.4	1.9	2.3	5.1	2.5	1.5	1.3	1.1	1.1
1978		1.7	2.1	2.5	1.7	1.7	1.9	4.6	2.9	1.6	1.4	1.2	1.2
1979		1.8	2.1	2.1	1.5	1.4	1.7	4.1	2.6	1.6	1.2	1.1	1.5
1980		1.6	1.7	1.7	1.7	1.6	2.2	4.0	2.7	1.5	1.2	1.0	1.3
Nadie en casa													
1977		1.8	1.9	1.6	1.4	2.0	1.7	1.6	1.4	1.4	1.4	1.3	1.2
1978		1.4	1.3	1.4	1.5	1.8	1.4	1.4	1.3	1.4	1.5	1.8	1.5
1979		1.6	1.6	1.5	1.5	1.7	1.6	1.6	1.6	1.7	1.8	1.6	1.6
1980		1.8	1.5	1.5	1.7	1.9	1.7	1.6	1.4	1.4	1.5	1.5	1.4
Rechazo													
1977		1.6	1.5	1.4	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
1978		1.7	1.3	1.4	1.6	1.7	1.5	1.5	1.4	1.4	1.4	1.4	1.4
1979		1.4	1.5	1.5	1.5	1.6	1.5	1.5	1.3	1.3	1.2	1.2	1.3
1980		1.3	1.5	1.4	1.5	1.9	1.7	1.6	1.5	1.3	1.3	1.3	1.4
Otros													
1977		0.8	0.6	0.6	0.4	0.6	0.4	0.4	0.4	0.4	0.3	0.6	0.5
1978		0.6	0.5	0.4	0.5	0.5	0.3	0.4	0.3	0.3	0.3	0.6	0.4
1979		0.6	0.6	0.7	0.6	0.4	0.4	0.3	0.4	0.3	0.3	0.3	0.4
1980		0.6	0.7	0.6	0.5	0.5	0.5	0.4	0.3	0.5	0.3	0.4	0.3

ii) No respuesta y tiempo de permanencia en una encuesta

La Tabla 6 provee un ejemplo de tasas de no respuesta

en la LFS según el número de veces que una familia ha estado en la encuesta. A esto le sigue una breve discusión sobre varios componentes de las tasas totales de no respuesta.

TABLA 6
TASAS DE NO RESPUESTA (%) SEGUN LA PERMANENCIA DE HOGARES
EN LA LFS (1979-1980)

Número de meses en la encuesta	Tasas de no respuesta (%)			
	Total no respuesta	Rechazo	Nadie en casa	Ausencia temporal
1	8.04	1.43	2.96	2.94
2	5.09	1.21	1.44	1.99
3	4.71	1.32	1.10	1.90
4	4.65	1.46	1.09	1.79
5	4.62	1.51	0.99	1.77
6	4.45	1.52	0.78	1.73

Sobre la base de los resultados que aparecen en la Tabla 6 se pueden hacer los siguientes comentarios:

- (a) La tasa total de no-respuesta fue más alta durante el primer mes, claramente porque los entrevistadores tuvieron mayores dificultades en encontrar a la gente en casa no sabiendo cuándo era el mejor momento para localizarlos; esto confirma la alta tasa de no-respuesta debida a nadie en casa. La tasa decrece abruptamente en el segundo mes y continúa decreciendo durante el 3º y 4º mes.
- (b) La tasa de rechazo decrece en el segundo mes, incrementa gradualmente en el 3º, 4º y 5º, y se nivela para el 6º. El decrecimiento inicial representa probablemente el efecto de una intervención con supervisión; el incremento siguiente se debe probablemente a la carga acumulada de varias encuestas.
- (c) La tasa de nadie en casa decrece rápidamente del primer mes al 2º en casi un 50%. Decrece de nuevo del segundo al tercer mes y más despacio durante el cuarto y quinto mes. En el sexto mes se da un decrecimiento mayor. El comportamiento de la tasa nadie en casa durante los seis meses de permanencia de una familia en la encuesta se debe probablemente al hecho que cuanto más tiempo esté una familia en la encuesta mejor sabe el entrevistador en qué momento son localizables.
- (d) La tasa de ausente temporalmente decrece durante los seis meses, particularmente del primero al segundo mes. La tasa temporalmente ausente no debería depender tanto del tiempo de permanencia de la familia y por ello parece que los entrevistadores han confundido la no-respuesta debida a nadie en casa con la de ausente temporalmente.

TABLA 7
TASAS DE NO RESPUESTA (%) POR EL TIPO DE AREAS (PROMEDIOS MENSUALES: 1985)

Tipo de área	Proporción de la muestra	No respuesta total	Rechazo	Nadie en casa	Ausente temporal
NSRU ¹	48.03 %	4.9 %	1.2 %	1.4 %	1.8 %
- urbana ²	17.58 %	5.2 %	1.1 %	1.4 %	2.2 %
- rural ²	30.45 %	4.8 %	1.3 %	1.4 %	1.6 %
SRU ³	51.08 %	5.7 %	1.8 %	1.8 %	1.8 %
- construida ⁴	36.34 %	5.6 %	1.7 %	1.7 %	1.8 %
- límite ⁴	10.47 %	4.8 %	1.6 %	1.2 %	1.6 %
- apartamentos ⁵	4.27 %	9.9 %	2.5 %	4.0 %	2.6 %

1 NSRU¹ son áreas que no están dentro de los SRU y contienen centros rurales y pequeños núcleos urbanos.

2 Toda unidad primaria de muestreo en una NSRU se divide en una parte rural y otra urbana.

3 SRU: Unidades representativas; son ciudades cuya población excede de 15.000 personas o cuyas características particulares hace que sea considerada como unidad autorrepresentada.

4 Las SRU se estratifican en subunidades y las subunidades se clasifican en zonas construidas o zonas de frontera dependiendo de sus posibilidades de crecimiento. Hablando con generalidad los hogares que pertenecen a la frontera en la SRU son los que están en las afueras o en las áreas suburbanas.

5 En 17 de grandes ciudades en Canadá hay estructuras separadas de apartamentos en edificios que contienen por lo menos cinco pisos y 30 o más unidades.

Dentro de la NSRU la tasa total de no respuesta fue más alta en la parte urbana, debido a tasas de ausentes temporales entre los hogares urbanos. Las tasas de nadie en casa en las partes rural y urbana fueron las mismas, pero las tasas de rechazo fueron un 20% más altas en las áreas rurales que en las urbanas dentro de las NSRU.

Dentro de la SRU, las áreas construidas tuvieron una tasa de no respuesta total más alta que las áreas frontera debido a los componentes nadie en casa y ausente temporalmente.

Parece que la gente que vive en el centro de las ciudades son más difíciles de contactar que la gente que vive en las áreas frontera; las diferencias sin embargo no son grandes.

Los apartamentos (SRU) son los que tienen la tasa de no respuesta más alta que cualquier otra área como se muestra en la Tabla 7. De hecho la tasa de no respuesta dentro de los apartamentos (SRU) casi doblaba la tasa de no respuesta de las otras dos clasificaciones. Las tasas de rechazo, nadie en casa y ausente temporalmente fueron también más altas dentro de los apartamentos.

La tasa nadie en casa fue casi tres veces mayor en la encuesta de apartamentos que en la de no-apartamentos. Esta gran diferencia puede deberse a un estilo de vida diferente de los habitantes de los apartamentos. Las familias que viven en apartamentos son en general muy pequeñas, siendo lo más normal que sea una sola persona; tienden a moverse más y son difíciles de encontrar en casa, mientras que las familias que no viven en apartamen-

tos suelen tener niños y son de un tamaño mayor. Otro problema suele ser que los entrevistadores a menudo tienen dificultades para entrar en los edificios.

Mientras que la tasa ausente temporalmente fue generalmente dos veces mayor en los apartamentos que en los no-apartamentos, la diferencia fue más pequeña durante los meses de Julio y Agosto que durante los otros 10 meses. Esto se debe posiblemente al hecho que es más fácil que personas que viven solas o matrimonios sin hijos disfruten de sus vacaciones en otras estaciones distintas al verano, que el que lo hagan las familias que tienen niños en edad escolar.

La tasa de rechazo fue casi siempre mayor en la muestra de apartamentos que en el resto, aunque esta diferencia no fue tan grande en las tasas nadie en casa y ausente temporalmente.

Los resultados recientes indican sin embargo que la diferencia en las tasas de rechazo entre apartamentos y no-apartamentos se van agrandando de tal forma que la tasa de rechazo en la muestra de apartamentos casi dobla la tasa correspondiente a la muestra de no-apartamentos.

La razón por la cual examinamos las tasas de no-respuesta de acuerdo con diferentes clasificaciones, como

la de SRU y NSRU, es que esta aproximación nos ayuda a establecer relaciones entre los diversos tipos de áreas en términos de los comportamientos de las tasas de no-respuesta. Por ejemplo, la tasa de no-respuesta es siempre mayor en la SRU que en la NSRU y se considera extraño el resultado de una encuesta cuando esto no sucede. Lo mismo se mantiene para las tasas nadie en casa y de rechazo. Otro ejemplo es la relación dos a uno entre la tasa de no-respuesta en la muestra de apartamentos y la correspondiente en la muestra de no-apartamentos. Si las tasas de no-respuesta se incrementan más allá de los niveles esperados, entonces el conocimiento de estas relaciones nos ayuda a analizar la situación y a buscar remedios.

iii) Incidencia de la no-respuesta en las asignaciones del entrevistador

Como era de esperar, las tasas de no-respuesta varían con los entrevistadores, muchos entrevistadores de hecho consiguen el 100% respuesta y muy pocos son los que no consiguen por lo menos un 75%. Es interesante disponer de la distribución de los entrevistadores según sus niveles de tasas de no-respuesta como lo muestra la Tabla 8.

TABLA 8
TASAS DE DISTRIBUCION DE ENTREVISTAS SEGUN NO RESPUESTA
(PROMEDIO MENSUAL: 1980)

Tasa no respuesta total (%)	Número de entrevistados	Porcentaje de entrevistados total
0.0	159	15.0
0.1 to 5.0	434	40.8
5.1 to 10.0	333	31.3
10.0 to 15.0	98	9.2
15.1 to 20.0	29	2.7
más 20.0	10	1.2
TOTAL	1,063	100.0

Los datos de la Tabla 8 se basan en todos los entrevistadores que encuestaron por lo menos 20 hogares y representan promedios durante los doce meses de 1978. Esta Tabla indica que el 56% de los entrevistadores consiguieron tasas de no respuesta de 5,0 ó mejores.

El 13% de los entrevistadores con tasas de no respuesta mayores que el 10.0 representaban el 31% o sea casi

1/3 del total de la no-respuesta. También resulta interesante ver que el 59% de los entrevistadores no tuvieron ninguna no-respuesta del tipo *nadie en casa* y el 54% no encontró *rechazos*. Además, el 71% y el 70% de los entrevistadores consiguieron tasas de *nadie en casa* y *rechazo* del 2% o mejores. Estos porcentajes reflejan el éxito de los programas de entrenamiento, control y vigilancias así como de los procedimientos y técnicas de entrevistar diseñados para maximizar los niveles de respuesta en la LFS.

IV. NO RESPUESTA SEGUN EL TIPO DE ENCUESTA

Los factores que más influyen en las tasas de no respuesta son los temas delicados y controvertidos (como renta, criminalidad, información sobre educación y nutrición; la última además significa una carga adicional para el encuestado al requerirle pruebas clínicas). Nuestra experiencia ha sido que el esfuerzo hecho por incrementar las tasas de respuesta en la LFS ha beneficiado a otras encuestas. Sin embargo, un efecto combinado de cuestiones controvertidas, complejidad y longitud del cuestionario puede resultar demasiada carga para el encuestado, aumentando la no respuesta incluso cuando la encuesta puede ser un suplemento de la LFS.

De hecho esto es lo que ha sucedido con la encuesta sobre Gastos en Consumo (SCF). Sin embargo, se observa que cuando se utiliza para la SCF una muestra independiente, la carga es menor y el encuestado

suele cooperar dando lugar a una tasa más pequeña de no-respuesta que cuando es un suplemento de la LFS.

Un segundo dato importante es el método de recogida de datos. Generalmente las entrevistas personales tienen tasas de no-respuesta más bajas que cualquier otro método refiriéndose siempre a encuestas de temas comparables y con un mismo nivel de controversia. Así por ejemplo, las encuestas sobre viajes tuvieron una tasa en no respuesta de un 7,3% al realizarse por entrevista personal, de un 10% al hacerlas por teléfono y de un 16% al utilizar el correo. Las encuestas sobre hábitos de fumar tuvieron un nivel de no-respuesta del 8% al hacerse mediante entrevista personal y de un 10% al hacerse por teléfono. Se debe señalar que ambas encuestas se realizaron como suplemento a la LFS.

TABLA 9
TASA DE NO RESPUESTA POR EL TIPO DE ENCUESTA

Nombre de la encuesta	Año	Método de recogida	Tasa de no respuesta (1)	Suplemento a la LFS
Encuesta de Gasto en consumo (marco)	1972	Envío-recogida	28.3	Sí
" "	1973	"	18.1	No
" "	1974	"	25.7	Sí
" "	1975	"	20.2	No
" "	1976	"	27.1	Sí
" "	1977	Personal	20.3	No
" "	1978	Envío-recogida	34.4	Sí
" "	1979	"	28.1	No
Actividades recreativas seleccionadas	1978	"	13.0	Sí
Encuesta sobre nutrición	1972	Personal	50.0 ⁽²⁾	No
Equipamiento del hogar	1978	"	9.1	Sí
" "	1979	"	11.5	Sí
" "	1980	"	15.9	Sí
Encuesta sobre viajes	1979	"	7.3	Sí
Hábitos de fumar	1979	"	8.0	Sí
Estudio sobre viviendas	1979	"	20.0	No
Encuesta sobre viajes	1979	Teléfono	10.0	Sí
Hábitos de fumar	1979	"	10.0	Sí
Criminalidad	1979	"	15.0	No
Graduación 1976	1979	"	14.5	No
Encuesta sobre viajes (1977)	1977	Envío-dev. correo	16.0	Sí
Test metodológico	1977	"	18.0	Sí

1 Tasas de no resp. = $\frac{n^2}{N} \times 100$ donde N es el total muestras y n_2 es el número de unidades que responden

2 Incluye 25% que rehusó ir a la clínica

V. EL PAPEL DE LA NO RESPUESTA EN UNA ENCUESTA

Aunque, la medida de respuesta/no respuesta según viene definido por diversas tasas es bastante sencilla, sin embargo, su importancia, interpretación y control requieren discusión.

Miremos el papel de la no-respuesta desde otro punto de vista, desde el punto de vista de las operaciones de diseño, recogida de datos, análisis y dirección de la encuesta.

i) Diseño

En la etapa de planificación es importante tomar una decisión sobre el nivel de tolerancia de la no-respuesta en la muestra. Se suele admitir en encuestas a nivel nacional, y para cuando los que responden y los que no responden tienen características similares, un nivel de no-respuesta del (20-34%). En este caso el sesgo será pequeño aunque incrementará la varianza. Lo mismo se argumentaría para estimaciones de tendencias y pre-encuestas. Sin embargo, si las estimaciones deben ser precisas y se requieren a niveles más pequeños que el nacional, entonces el tamaño de la no respuesta se debe mantener entre el 5 y el 7% ya que se pueden producir bolsas de no respuesta en áreas locales.

Otro papel importante que juega la respuesta/no-respuesta en esta etapa es el coste de la encuesta. Es importante localizar el coste de los diversos factores de tal manera que se consiga una no-respuesta suficientemente baja que sirva a los fines de la encuesta. Generalmente es mejor aceptar muestras más pequeñas que las planificadas en un principio y destinar más dinero a la recogida de datos y al seguimiento. Esto será particularmente apropiado si se espera grandes diferencias entre los que responden. Los diseñadores de encuestas con experiencia pueden estimar todo lo dicho con bastante precisión. Además pueden identificar de una manera más formal un importante número de factores que también afectarán a la no respuesta.

Es evidente que los siguientes factores afectan a la no respuesta: marco muestral, método de entrevista, selección, entrenamiento y control del equipo, longitud y tipo de lenguaje del cuestionario, tipo de preguntas, tipo de área, viabilidad y número de devoluciones y finalmente publicidad. Idealmente el diseñador de una encuesta desearía poder realizar uno o varios estudios pilotos para contrastar la importancia de por lo menos algunos de estos factores. A falta de experimentación, una discusión a fondo basada en la experiencia pasada y en la intuición

es esencial para llegar al diseño final.

En la planificación y desarrollo de una encuesta se deben tener en cuenta varios factores para llegar al diseño final. Estos factores se clasifican en tres grupos:

Grupo I: Estos factores pertenecen a la etapa de diseño muestral; factores que tienen un efecto indirecto en las tasas de respuesta.

- a) marco muestral
- b) tamaño muestral
- c) estratificación
- d) reparto de la muestra por estratos
- e) procedimiento muestral dentro de los estratos
- f) reparto de la muestra por etapas, especialmente conglomerados de la muestra.

Grupo II: Factores que pertenecen al diseño del cuestionario y a los procedimientos de recogida de datos y de relaciones públicas; cuestiones que afectan directamente a las tasas de no-respuesta.

- a) listado de unidades para la selección muestral
- b) tema y tipo de encuesta
- c) procedimiento mediante el cual se realiza el cuestionario
- d) longitud y complejidad del cuestionario
- e) cuestiones controvertidas para encuestadores y encuestados.
- f) selección, desarrollo y control del equipo de campo incluyendo a los entrevistadores
- g) tipo de área para la encuesta
- h) manejo y coste de las encuestaciones repetidas y procedimientos de seguimiento
- i) publicidad/medios.

Grupo III: Estos factores pertenecen a la etapa de procesamiento y análisis después de que se ha completado la encuesta y no hay oportunidad ya de tratar en el campo la no-respuesta.

- a) depuración previa e imputación
- b) procedimientos de estimación, incluyendo ajustes de las ponderaciones en las células.
- c) varianza y estimación de la varianza
- d) publicación y análisis de datos.

Todos estos factores están interrelacionados, y por lo menos, aquéllos del Grupo I y II tienen su importancia en la incidencia de la no respuesta, que a la vez afecta al error cuadrático medio de las estimaciones.

Sobre la base de un modelo de error para los datos de una encuesta que incluye la respuesta muestral, y los componentes del error de imputación, se puede obtener la siguiente expresión para el error cuadrático medio (MSE) de una estimación de una encuesta basada en procedimientos de imputación.

$$MSE = \gamma + B^2$$

$$\gamma = SV + SRV + CRV + VRR + CVRR$$

$$B = SB + RB + IB$$

donde SV: varianza muestral

SRV: varianza de respuesta simple

CRV: varianza de respuesta correlada

VRR: varianza que proviene del status de respuesta/-no-respuesta de las unidades seleccionadas

CVRR: covarianza proveniente de la covarianza entre los status respuesta/no respuesta de pares de unidades seleccionadas

B: sesgo

SB: sesgo muestral

RB: sesgo de la respuesta

IB: sesgo de la no respuesta o imputación.

La varianza muestra SV se ve afectada por todos los factores del Grupo I y Grupo III (a) y (b) y casi siempre se incrementa con el tamaño de la no-respuesta después de la aplicación de los procedimientos de depuración previa e imputación.

El sesgo muestral SB se ve afectado por dos factores del Grupo I incluyendo (a) marco muestral incompleto y (e) procedimientos muestrales, dentro de los estratos a no ser que se realice un muestreo no probabilístico o por cuotas. También se ve afectado por un par de factores del Grupo II: (a) listado incompleto de unidades, (f) caso de equipo de campo poco entrenado y entrevistadores que pueden no entrevistar a las unidades tal y como ha sido diseñado en la muestra. El sesgo muestral, sin embargo, no se verá afectado por un incremento en las tasas de no respuesta.

Es importante señalar que cada encuesta tiene sus

propios requerimientos en relación con el diseño, cuestionario y procedimiento de realización de éste. En lo que concierne al diseño, un factor importante que afecta a la no respuesta son los factores (e) y (f) del Grupo I el procedimiento de muestreo de los estratos y el reparto de la muestra entre las diferentes etapas (si las hay). Por ejemplo, una muestra no asignada por conglomerados puede producir una tasa de no respuesta más alta que un diseño por conglomerados. Esto puede deberse a que al no ser el muestreo por conglomerados y si las entrevistas son personales se ha de viajar más, el coste de ello hace que se restrinja la encuestación repetida.

El diseño del cuestionario, así como los diversos procedimientos para su contestación y otros factores del Grupo II, tienen su efecto en los componentes de la varianza no muestral SRV + CRV + VRR, en el sesgo de respuesta RB, en la tasa de no respuesta y consecuentemente en el sesgo de no respuesta que se puede incrementar con la tasa de no respuesta, después de la aplicación de los procedimientos de depuración previa e imputación. Además todos los componentes de la varianza no muestral tienden a incrementarse cuando los tamaños muestrales que responden tienden a decrecer.

Algunos ejemplos de casos específicos de los factores del Grupo II que afectan directamente a la tasa de no respuesta serán mencionados. Estos incluyen procedimientos por los que se realizaron los cuestionarios, longitud y complejidad de los cuestionarios y tipo de áreas para la encuesta.

Por ejemplo, se demostró en la sección 2.5 y Tabla 3 cómo los procedimientos mediante los cuales se realizan los cuestionarios Grupos II (c), afectan a las tasas de no-respuesta.

Grupo II (d) longitud y complejidad del cuestionario: la longitud del cuestionario no sólo afecta a la tasa de no respuesta sino también al procedimiento de encuestación a emplear. Cuestionarios largos y complejos pueden necesitar de la entrevista personal, mientras que los cortos y sencillos se pueden realizar por correo. Como el cuestionario es especialmente importante en el diseño de una encuesta se discute más a fondo en la sección 2.10.

Grupo II (g) tipo de área de la encuesta: suele ser más fácil encontrar a alguno en casa en las pequeñas comunidades urbanas o rurales que en las grandes ciudades, sobre todo en los apartamentos donde las unidades familiares constan de uno o dos individuos que suelen estar casi siempre ausentes.

Estos ejemplos muestran factores a tener en cuenta en la etapa de planificación y desarrollo dentro del diseño de una encuesta. Si se presta atención a los factores del Grupo II, se pueden evitar serios problemas de no respuesta.

En la etapa de diseño se puede argumentar que los problemas asociados a la no-respuesta no pueden resolverse sencillamente comenzando con una muestra mayor de la necesaria y así permitir una no respuesta potencial, ya que con la presencia de no respuesta inicial, la muestra no sería ya una muestra probabilística. Además, las unidades encuestadas con probabilidad de respuesta son generalmente diferentes en formas y grados de las unidades que realmente responden. Es verdad que si suponemos que la probabilidad de respuesta es la misma para todas las unidades encuestadas con probabilidad de respuesta, el sesgo se eliminaría ajustando (inflando) los pesos muestrales de las unidades muestrales con probabilidad de respuesta. Sin embargo, la probabilidad de respuesta puede depender de la característica de interés y solución antes enunciada puede dar lugar a sesgo en la no-respuesta, dependerá de las relaciones entre la característica de interés y la probabilidad de respuesta.

ii) Cuestionario

El cuestionario es fundamental en el desarrollo de una encuesta. Su longitud, claridad y estructura afectan considerablemente a la magnitud de la no-respuesta. La definición más sencilla de cuestionario es la de un grupo de preguntas diseñadas para conseguir información de una persona sobre un tema. Dentro del campo de técnicas de entrevistas, el cuestionario puede ser desde una lista de temas poco definidos hasta un conjunto de cuestiones muy estructuradas sin otras opciones de respuesta distintas que las listadas.

En la mayoría de las encuestas, el cuestionario es un método importante de estandarizar y controlar el proceso de recogida de datos. Si no hubiese unas preguntas e instrucciones claras, los entrevistadores cambiarían inevitablemente el significado y el énfasis de las preguntas y posiblemente de las respuestas. El cuestionario juega un papel importante en un proceso complejo (la entrevista) en el que la información se transfiere de los que la tienen a los que la necesitan. El cuestionario es el medio a través del cual la información que los usuarios necesitan se expresa en términos operativos y se presenta de manera que el encuestado pueda proporcionarla. Para que la transferencia de información sea efectiva el cuestionario debe satisfacer tanto los requisitos de los u-

uarios como de los encuestados. El contenido total del cuestionario, el estilo de encuestar, el formato, la longitud, todo está muy unido al método de recogida de datos y al tema de la encuesta. Cada método de recogida de datos, entrevista personal, telefónica o por correo crea sus propias condiciones y forma un marco importante al cual un buen cuestionario se debe adaptar a través de un estilo apropiado en las preguntas, contenido, formato, longitud. Por ejemplo, en las entrevistas personales suele ser posible que el entrevistador mediante observación directa tome ciertos datos como el sexo del entrevistado, el tipo de casa. También, el cara a cara es un factor muy motivador, para el encuestado. Generalmente, cuando el cuestionario es largo y complejo la única elección es la entrevista personal. En las entrevistas telefónicas se pierde la interacción social entre entrevistador y entrevistado, lo que puede afectar a la cooperación que preste este último. Para tener éxito el cuestionario debe basarse fundamentalmente en la comunicación verbal, siendo de menor importancia el tema sobre el que verse. Sin embargo, en algunas encuestas controvertidas (encuestas sobre criminalidad), el que exista distancia entre entrevistador y entrevistado puede facilitar el que se responda a la encuesta. En las encuestas por correo, el cuestionario adopta el papel de entrevistador. Debe introducir la encuesta, motivar al encuestado para que coopere y guiarle hasta que la termine.

El cuestionario es una fuente importante de errores de no-respuesta que pueden ocurrir básicamente por dos razones: a) *no contacto* (nadie en casa, ausente temporalmente, mal tiempo ...) o b) *rechazo*. Lo último puede ser debido a la no-respuesta de una unidad (rechazo a participar en la encuesta) o no respuesta a una pregunta (rechazo a contestar una o más preguntas). El cuestionario puede hacer poco por eliminar el *no contacto* pero juega un papel importante en prevenir el rechazo. Supuestos poco reales sobre el conocimiento o la memoria del encuestado el uso de un lenguaje muy difícil y técnico o demandas excesivas son causas de no-respuesta y tienen su raíz en el cuestionario.

El cuestionario es también una fuente importante de errores de respuesta, que pueden ocurrir debido a la complejidad de las cuestiones y al diseño del cuestionario. En tales casos, los encuestados responden pero dan respuestas incorrectas debidas a malentendidos o a no conocer el material relevante.

Como hemos visto el cuestionario es causa inevitable de errores ajenos al muestreo pero se debe hacer todo lo posible para evitarlos. El grado en el que el cuestiona-

rio consigue disminuirlos depende de su diseño. Cada nueva encuesta presenta nuevos problemas y escollos, y debemos anticiparnos y tenerlos en cuenta a la hora de diseñar el cuestionario.

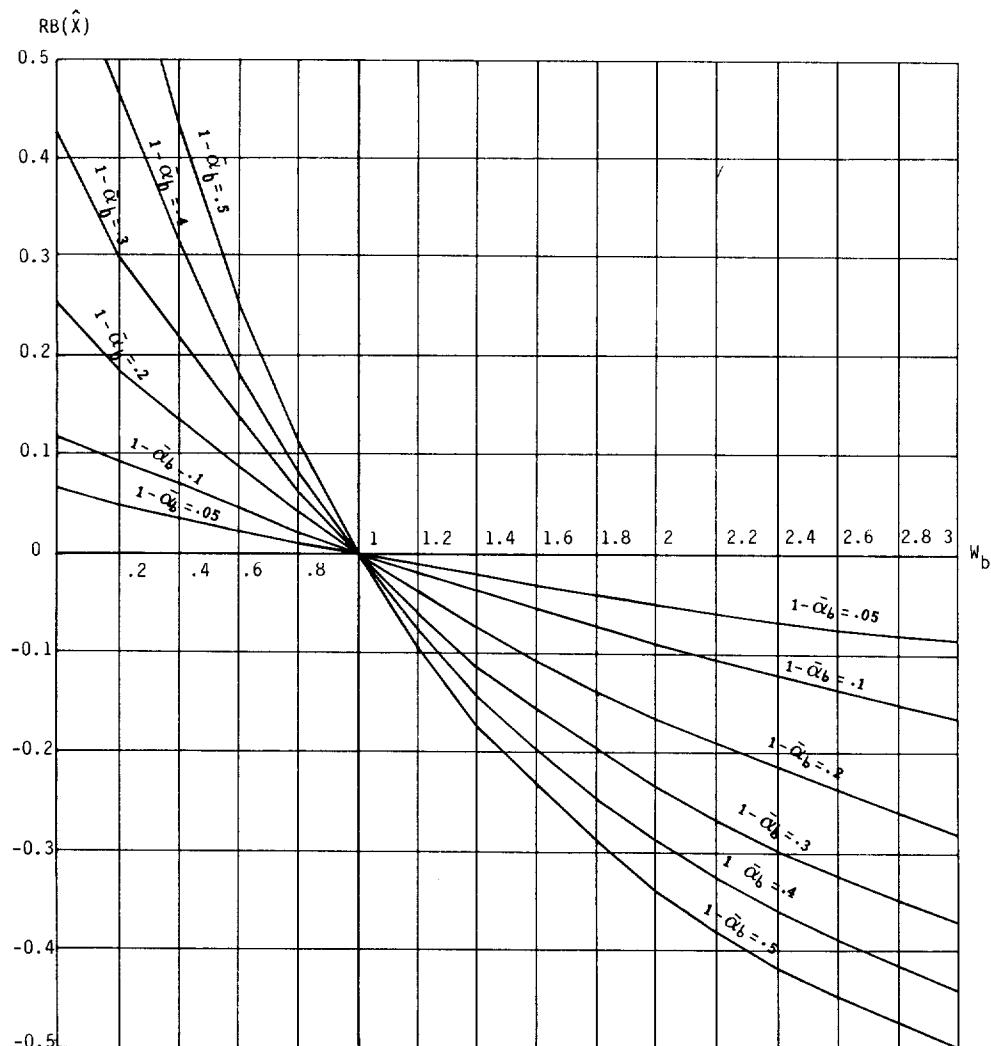
iii) Sesgo relativo debido a la no-respuesta

Habiendo examinado los diversos factores que pueden afectar al tamaño de la no-respuesta en una encuesta en fase de diseño, podemos deducir que un estadístico especialista en diseño de encuestas puede hacerse una idea de lo que es factible esperar en lo referente a la no-respuesta.

Consecuentemente, un factor importante al planificar una encuesta es decidir el nivel tolerable de no-respuesta y un estadístico experimentado puede estimar adecuadamente el nivel de respuesta esperable en una encuesta determinada, dependiendo de las condiciones en que se realiza la encuesta. Para ayudar a los estadísticos de encuestas a determinar el nivel de tolerancia adecuado, R. Platek, M.P. Single y Tremblay, en un trabajo publicado en un libro sobre muestreo y medida editado por Academic Press desarrollan una fórmula y un gráfico que permiten estimar el límite superior del sesgo relativo de una determinada característica de la encuesta y para un determinado método de imputación.

SESGO RELATIVO DE IMPUTACION DEL ESTIMADOR DE TASA DE NO RESPUESTA

$$(1 - \bar{\delta}_b) & W_b$$



W_b = Unidades que responden/Unidades que no responden

Consideramos la estimación del total poblacional X de una determinada característica x . La población se puede clasificar en los que responden y los que no responden de forma que:

R : número de unidades que responden

X_R : número de unidades que responden y con la característica x

X_N : número de unidades que no responden y con la característica x

N : número de unidades que no responden

$X = X_R + X_N$: población con la característica x

$T = R + N$: número de unidades

Supongamos que imputamos Z_N a X_N entonces $\hat{X} = X_R + Z_N$ será la estimación de X . En general, podemos expresar el sesgo de X así:

$$\begin{aligned}\text{Sesgo } (X) &= (X_R + Z_N) - X = X_R + Z_N - X_R - X_N = \\ &= Z_N - X_N = N(\bar{Z}_N - \bar{X}_N)\end{aligned}$$

a) Si no se realiza ningún ajuste para paliar la no respuesta, el sesgo se debe íntegramente a \bar{X}_N .

b) Si se aplica la corrección, entonces:

$$B(\hat{X}) = N(\bar{X}_R - \bar{X}_N)$$

En la práctica a veces el estadístico no puede estimar el sesgo en la no-respuesta. Sin embargo, sería deseable establecer una cota superior (positiva o negativa) para el sesgo.

Se puede argumentar, con cierta justificación, que si se imputa \bar{Z}_N a \bar{X}_N se espera que Z_N esté más cercano a \bar{X}_N que \bar{X}_R . De otra forma no se necesitaría calcular la imputación para \bar{X}_N . Esto se debe a que el estadístico piensa que se da la siguiente desigualdad.

$$|\bar{Z}_N - \bar{X}_N| \leq |\bar{X}_R - \bar{X}_N|$$

Bajo estas hipótesis el límite superior del sesgo sería $(\bar{X}_R - \bar{X}_N)$. Si este no fuese el caso, aplicaríamos el método más sencillo y barato de compensar la no-respuesta (estimando \bar{X}_N mediante \bar{X}_R).

En el caso en que el límite superior se pueda obtener por $(\bar{X}_R - \bar{X}_N)$, el sesgo relativo se puede expresar mediante

$$RB(\hat{X}) = \frac{\hat{X} - X}{X} = \frac{(1 - \alpha_b)(1 - W_b)}{\alpha_b + W_b(1 - \alpha_b)} = \frac{(1 - \alpha_b)(1 - W_b)}{1 - (1 - \alpha_b)(1 - W_b)}$$

$$\text{donde } W_b = \frac{\bar{X}_N}{X_R} =$$

$1 - \alpha_b$ = tasa de no-respuesta

b = unidad de balance (área)

Se ve en el gráfico cómo los cambios relativos afectan a la tasa de respuesta. Por ejemplo: si $1 - \alpha_b = 0,2$ y $W_b = 1,5$ entonces $RB(\hat{X}) = -0,09$ (0,9%). Si por métodos inapropiados se reduce $(1 - \alpha_b)$ a 0,10 pero W_b se incrementa hasta 2 entonces $RB(\hat{X}) = -0,09$ y el coste añadido para reducir los resultados de la no-respuesta no produce reducción en el sesgo. De modo que, si el valor de α_b se obtiene de encuestas anteriores o de otro tipo de estudios, entonces es posible obtener la magnitud del sesgo relativo con la ayuda del gráfico. En otras palabras se puede determinar la tasa de respuesta que se ha de conseguir en una encuesta de forma que el sesgo quede dentro de límites tolerables.

iv) Recogida de datos

Los distintos usos posibles de las diferentes tasas de no-respuesta/respuesta justifican la importancia de su estudio en la etapa de recogida de datos. Una primera clasificación de estas tasas sería *contacto, no contacto*. Nadie en casa o ausente temporalmente que son ejemplos de no contacto y tienen una orientación operativa. El otro tipo es la verdadera no-respuesta, cuando se ha tomado contacto con el que debe responder pero no se obtiene una respuesta aceptable o ni siquiera contesta. Hay soluciones operativas para el problema de no contacto. Por ejemplo, en las encuestas personales o telefónicas son importantes el momento y la forma de llamar al que tiene que responder. El tamaño de las asignaciones y el

tiempo dedicado a la recogida de datos deben ser adecuados. Un examen más profundo del *no contacto* puede revelar que algunas unidades, por ejemplo, casas, deben ser inelegibles. La inelegibilidad varía con el tipo de información requerida.

Por ejemplo, una casa vacía es inelegible para una encuesta de familias. Pero es elegible para una encuesta de casas. Otras casas pueden ser en su mayoría apartamentos ocupados por jóvenes solteros y empleados. Incluso otras unidades pueden tener otras características. De ahí la importancia de conocer el origen de la no respuesta ya que ayuda a determinar la estrategia apropiada ante un tipo particular de no-respuesta.

El problema de los rechazos es diferente. Las tasas de rechazo en principio no son claras. Así un entrevistador puede preferir registrar un rechazo como *nadie en casa* o la persona a entrevistar puede sencillamente no contestar a la puerta como forma de rechazar y se la registrará como *nadie en casa*.

En una encuesta por correo no se sabe con certeza si el que tiene que responder ha recibido el cuestionario o si habiéndolo recibido se olvidó de devolverlo por correo. En el propio proceso de entrevistas, el entrevistador se puede encontrar con unidades que no deberían estar y otras con el cuestionario completo o parcialmente completo.

Encuestas que tratan un tema delicado pueden influir no sólo en los rechazos de esa encuesta concreta, sino que al afectar a los encuestados, influyen también en encuestas posteriores. Para evitar las bolsas de no-respuesta ya conocidas, es deseable definir de tal manera la población a encuestar que éstas no se produzcan. Pero se debe tener en cuenta la dificultad de inferir a partir de la población encuestada las características de la población objetivo.

Un papel importante que juega la no-respuesta en la etapa de recogida de datos se refiere al tamaño de la no respuesta en ciertas situaciones o áreas. Por ejemplo, si las tasas de no respuesta de un entrevistador o de unas áreas se producen antes de que finalice la recogida de datos, se pueden identificar los entrevistadores y las áreas que necesitan apoyo para conseguir tasas de respuesta satisfactorias. También si se realizan tabulaciones preliminares pueden éstas indicar si la muestra que ha respondido es demasiado pequeña o está distribuida demasiado desigualmente como para obtener estimaciones satisfactorias de las características importantes de la

población. Se puede realizar entonces un esfuerzo mayor en la recogida de datos de forma que se corrijan estos defectos.

También puede suceder que el entrevistador en un descuido codifique la respuesta incorrectamente produciendo datos no válidos que deben descartarse. También englobaremos dentro del problema de no-respuesta la motivación del encuestado de forma que dé una respuesta válida.

En relación con la motivación, supongamos que la unidad encuestada es neutral respecto a la encuesta y consideremos qué es lo que puede influirle en que responda o no. Factores tales como: dificultad en comprender la pregunta, tiempo, intimidad, discrepancias, dificultad en recordar la información, cuestiones personales o embarazosas son ejemplos de motivos que pueden llevarle a no responder. Por otro lado factores que le llevan a responder pueden ser: interés por la encuesta, deseos de ayudar, sentido del deber, comprensión de la importancia de los resultados de la encuesta, etc.

El problema es cómo acentuar la motivación positiva y reducir la negativa hasta que la balanza se incline a favor de la respuesta.

El diseñador de encuestas debe tener presente todo lo que afecta al elemento clave, el encuestado, en especial a su motivación para responder y a su nivel de comprensión.

v) Motivación del encuestado

En la vida cotidiana generalmente cuando uno pregunta obtiene una respuesta, ¿qué es entonces lo que lleva a los encuestados a no responder? Hay tres razones: violación del derecho a la intimidad, carga de respuestas y hostilidad general al Gobierno. Al tratar con estos problemas, es importante verlos desde el lado de los encuestados, no desde el lado de las ideas preconcebidas del diseñador de encuestas. Los encuestados suelen percibir los cuestionarios como pesados cuando no los entienden o no ven en qué les concierne. En otras condiciones sin embargo, se pueden percibir las mismas cuestiones como interesantes y el encuestado se siente motivado para participar en la encuesta. En general, existen dos vías por las que se puede motivar al encuestado: relaciones públicas y relaciones con el encuestado.

Por relaciones públicas entendemos actividades dirigidas al público general y que pueden tomar formas diversas. Un objetivo a largo plazo es crear un clima entre

el público en general que tiende a motivarlo para cooperar en las encuestas. La imagen que el encuestado percibe de la organización de una encuesta puede ser un factor importante en su motivación para responder. La extensión y la forma de disseminar los datos se pueden usar para impresionar al público sobre la importancia del papel de la agencia estadística. La agencia estadística debe estar alerta ante las posibilidades de un comentario favorable en los medios de comunicación y también sobre la necesidad de tratar las críticas apropiadamente y a tiempo. Unas buenas relaciones públicas además de mantener una imagen favorable se pueden usar para publicar actividades específicas y requerir cooperación para llevarlas a cabo. Una publicidad mal hecha puede tener un efecto negativo. Por regla general las campañas publicitarias son más adecuadas para los censos de población que afectan a todos y en los que existen por tanto la necesidad de motivar a toda la población. Sin embargo, cuando la tasa tamaño muestral/tamaño de la población es pequeña, es más efectivo, desde el punto de vista del coste (en el marco de una única encuesta) evitar la publicidad directa y concentrarse en las relaciones con el encuestado.

vi) Relaciones con el encuestado

Una definición inicial del concepto relaciones con el encuestado, puede ser aquélla que comprende cualquier acción dirigida hacia el encuestado individual y que puede afectar a su actitud y motivación con respecto a la encuesta. Como antes señalábamos los problemas que más preocupan son la violación del derecho a la intimidad, la carga de respuestas y la hostilidad o desconfianza. En relación con la hostilidad y la desconfianza ya se han mencionado las relaciones públicas y la imagen de la agencia. Es muy importante la identificación tanto del entrevistador como de la agencia. Diversas personas indeseables usan el pretexto de estar haciendo una encuesta para establecer contacto o conseguir entrar en una casa. Es importante presentar identificación oficial y que los materiales de la encuesta den la impresión de que son realmente oficiales.

Cartas de presentación, ejemplos de los usos de los datos, folletos en donde se describen los objetivos de la encuesta, y el organismo que la realiza, son medios excelentes para evitar la hostilidad y la desconfianza.

El derecho a la intimidad se suele relacionar con el contenido del cuestionario aunque las reacciones de los encuestados son muy diferentes. Existen diversos procedimientos para minimizar este efecto en los encuestados y se debe diseñar un procedimiento específico para cada si-

tuación. En algunos casos lo mejor es permitir al encuestado responder de manera totalmente anónima.

Esto se puede lograr autoentrevistándose el encuestado con un cuestionario que no contenga identificación alguna. Bastante a menudo, sin embargo, resulta importante tener un código de área o una designación muestral para ponderar y estimar. En estos casos se debe tener cuidado con que el encuestado no perciba los medios utilizados para identificar sus contestaciones. Desde luego, está claro que aunque las encuestas anónimas pueden ser beneficiosas desde el punto de vista de que el encuestado siente que se le respeta su derecho a la intimidad, estos beneficios pueden ser contrarrestados por el hecho de que este tipo de encuestas no permiten el seguimiento de los que no responden.

Además de asegurar el derecho a la intimidad, diversas organizaciones han experimentado diferentes formas de compensar, a los encuestados por su tiempo y esfuerzo.

vii) Análisis

Cuando miramos los resultados de una encuesta y encontramos tasas de no-respuesta bajas, entonces los datos de la encuesta merecen nuestra confianza. Por otro lado, las tasas de no-respuesta altas suelen tomarse como indicadores de una pobre calidad de los datos. La interpretación de las medidas de no-respuesta son incluso más difíciles cuando tratamos diseños complejos ya que la concentración de la no-respuesta puede ser mayor en un área que en otra. La mayoría de los estadísticos han usado las tasas de respuesta como indicadores de la calidad de los datos. Por ello en toda encuesta se suelen tomar y evaluar los datos de no-respuesta. Sin embargo la no-respuesta sola puede o no afectar a la calidad de los datos. Hablando estrictamente el error medio cuadrático que es igual al sesgo² + varianza provee de una base uniforme para evaluar los resultados de una encuesta. Pero la no-respuesta afecta a su magnitud como se puede ver en la etapa de diseño. Su efecto depende también del tipo de estadística que se analiza. Intuitivamente parece que cuando deseamos estimar la media de una característica el tamaño del sesgo debido a la no-respuesta lo determinan tanto las diferencias entre los que responden y los que no responden como el tamaño de la no-respuesta. Pero, si tenemos razones para creer que los que responden y no responden no difieren mucho entonces el sesgo debido a la no-respuesta será pequeño e independiente del tamaño de la no-respuesta. En tal situación puede no importar un seguimiento de los que no responden. Por otro lado, si la diferencia es grande, entonces el sesgo será grande, a menos que el tamaño de la

no-respuesta sea muy pequeño. En este caso, es aconsejable un seguimiento o cualquier forma de reducir el tamaño de la no-respuesta. En la práctica nunca existe una situación clara. Un analista suele estar interesado en estimar varias medidas de las características. Por tanto un procedimiento seguro es tener una no-respuesta baja.

Podemos expresar este razonamiento intuitivo en una fórmula para la media.

Media

Supongamos:

R: número de unidades que responden

M: número de unidades que no responden

N: tamaño de la población

Por tanto $N = R + M$

r, m, n es la nomenclatura para los valores muestrales.

Si no se compensa la no-respuesta, la media de la gente que responde \bar{y}_r se utiliza para estimar \bar{Y} .

\bar{Y} es el número de unidades que tienen unas características determinadas.

$$E(\bar{y}_r) = \bar{Y}$$

$$\begin{aligned} \text{y el sesgo } B(\bar{y}_r) &= \bar{Y}_r - \bar{Y} = \bar{Y}_r - (\bar{R}\bar{Y}_r + \bar{M}\bar{Y}_n) \\ &= \bar{M}(\bar{Y}_m) + \bar{Y}_r(1 - \bar{R}) \end{aligned}$$

$$\text{pero } 1 - \bar{R} = \bar{M}$$

$$\bar{M}(\bar{Y}_r - \bar{Y}_m)$$

que demuestra lo que antes se discutió intuitivamente. Muestra también las condiciones bajo las cuales \bar{y}_r es un estimador insesgado de \bar{Y} .

Total

Si quisieramos estimar la población total entonces parece natural que el sesgo en presencia de no-respuesta

será pequeño sólo si es pequeño el número de unidades que no responden. En este caso debemos intentar hacer la no-respuesta lo más pequeña posible. Las siguientes fórmulas lo muestran:

$$E(\hat{Y}) = y_r$$

por tanto

$$B(\hat{Y}) = y_r - \bar{Y} = Y_r - (Y_r + Y_m) = -M\bar{Y}_m$$

La expresión anterior es cero si $M = 0$ ó $\bar{Y}_m = 0$

Varianzas y correlaciones

Similamente podemos examinar el efecto de la no-respuesta en otras estadísticas como las varianzas y las correlaciones. Sin embargo en estos casos la intuición está muy limitada. Se debe por tanto recurrir a expresiones matemáticas sobre el sesgo de la no-respuesta. Al tratar con más de una variable, los mecanismos del error son más complejos y pueden operar en direcciones opuestas. Las estimaciones de varianza y covarianza deberán tener en cuenta los errores adicionales de muestreo y no muestreo causados por la no-respuesta.

Cuando se utilizan como ajustes de las ponderaciones las inversas de las tasas de respuesta y así se infla una muestra deficiente, las estimaciones de la varianza y covarianza pueden tener pequeños sesgos en relación con sus verdaderos valores poblacionales siempre que se apliquen métodos de replicación apropiados. El sesgo depende tanto de la forma en que se realizan las repeticiones en el análisis de varianza/covarianza como de la manera en que se define las células de ajuste al ajustar las ponderaciones. Convendría partir de un modelo de error muestral/no muestral parecido al de Platek y Gray (1983) o al de J. Lessler (1983) y examinar el sesgo en la estimación varianza/covarianza.

Cuando se utiliza el fichero caliente (hot deck) o datos de encuestas anteriores en sustitución de los valores que faltan, generalmente se trata a estos datos sustitutos como reales. En estas circunstancias las estimaciones de la varianza subestimarán casi siempre, especialmente en el caso de sustitución mediante fichero caliente, que no es más que una duplicación de los resultados de la encuesta a partir de unidades similares a las que no han respondido. Cuando la sustitución se hace mediante encuestas históricas, si estos datos están muy correlados con los actuales, se puede justificar el tratarlos como respuestas observadas de cara a los análisis de varianza/covarianza.

Dirección de la encuesta

Definimos la dirección de la encuesta como una serie de controles que aseguran que la encuesta produce estadísticos de calidad aceptable. Algunas de las medidas tienen carácter preventivo. Otras intentan medir la efectividad de tales programas para mantener la calidad. Un análisis cuidadoso de las tasas de no respuesta puede ayudar a identificar los puntos débiles de los diversos programas. Por ejemplo, si falla en el entrenamiento de los entrevistadores o estos carecen de un conocimiento adecuado de las diferentes formas de llevar a cabo una encuesta o no entienden bien los conceptos, se incurrirá en un sesgo mayor. También la falta de adherencia en los procedimientos de encuestación incrementará la varianza de respuesta correlada. Dos programas que pueden ayudar en el análisis de las tasas de no-respuesta y de sus efectos en la calidad de los datos son la observación y la re-entrevista. La observación suele darse en el momento en que se está realizando la entrevista. Provee de una oportunidad de entrenamiento (re-entrenamiento o de motivación para el entrevistador) y de identificar los problemas de campo. Algunos entrevistadores cuyas tasas de no-respuesta son altas deben ser sometidos a observaciones más frecuentes. Así las medidas de la no-respuesta determinarán la intensidad y frecuencia de los programas de control.

Una reentrevista es una entrevista a unidades a las que ya se ha entrevistado durante la semana de la encuesta.

Las diferencias que se observan entre los dos conjuntos de respuesta se atribuyen a diversas fuentes entre las que se encuentran la propia entrevista y reentrevista. Es la actuación del entrevistador la que nos importa en relación con las tasas de no-respuesta en ciertas áreas y en ciertas categorías de la población. Se debe examinar cualquier cambio en las tasas de respuesta intentando establecer una razón para ello. De nuevo aquí, las medidas de las tasas de no-respuesta determinan la intensidad y frecuencia de las reentrevistas.

Un uso más directo de la no-respuesta sería mirar y examinar las asignaciones de cada entrevistador, su entrenamiento, cobros, así como el diseño muestral. Por ejemplo, se esperaría un nivel bastante uniforme de no-respuesta debido a nadie en casa para la misma encuesta realizada en condiciones similares. Pero si hay un fuerte cambio al repetir una encuesta entonces se debe identificar las causas. Una podría ser que se hubiesen modificado los procedimientos de realizar las entrevistas. Otra posibilidad es que los entrevistadores no hayan preparado su trabajo adecuadamente. Desde el punto de vista del diseño pudiera ser que las poblaciones objetivo y muestra hubiesen cambiado desde la etapa de diseño. Se deben examinar todas estas posibilidades. La dirección de la encuesta debe tener presente que los errores de respuesta y no-respuesta pueden estar correlados. Por ejemplo, si un encuestado potencial se siente presionado para proporcionar información que preferiría no dar, puede dar datos incorrectos. A esto nos referimos como el fenómeno de un encuestado reticente.

VI. TRATAMIENTO DE LA NO-RESPUESTA

Se ha mencionado que en la etapa de diseño cierto conocimiento del origen de la no-respuesta y de sus efectos en los datos conducirían sin duda a minimizar el tamaño de la no-respuesta. Sin embargo, la lucha más importante con la no respuesta se da en la recogida de datos así como en las etapas de su procesamiento y análisis.

i) Encuestaciones repetidas (callbacks)

La no-respuesta se puede reducir mediante esfuerzos persistentes de los entrevistadores y motivando a los que no responden para que respondan. Estos esfuerzos persistentes son en su mayor parte encuestaciones repetidas que son parte esencial de cualquier encuesta. Es claro intuitivamente que las familias que están en casa la primera vez que se va suelen diferir de las que no están. Así, si se tiene en cuenta únicamente a los que estaban en casa en el primer intento se obtienen unos resultados muy diferentes de los que se obtendrían si se repitiesen las entrevistas que no se pudieron realizar la primera vez. Por tanto, en el diseño del procedimiento de recogida de datos debe tenerse en cuenta la encuestación repetida a fin de reducir la no-respuesta y sus sesgos.

Las encuestaciones repetidas toman formas diversas, dependiendo del tipo de encuestas y de las formas de recoger los datos. Las encuestaciones repetidas son la forma más limpia de reducir la no-respuesta. No se necesitan supuestos. Los otros métodos se basan en la aceptación por el diseñador del riesgo de sesgo causado por un método particular frente al coste de hacer encuestaciones reiteradas si éstas fueran posibles.

En las encuestaciones por correo la encuestación repetida suele ser una carta recordando al individuo de la unidad seleccionada la importancia de la encuesta; en la carta se suele incluir una nueva copia del cuestionario por si acaso el encuestado perdió la primera copia. Se debe también indicar que no se tenga en cuenta este aviso si ya ha enviado la primera copia.

En las encuestas por correo no suele haber más de tres repeticiones y éstas se realizan antes de que finalice el plazo de recogida de datos y comience la fase de procesamiento de la información. A veces, las encuestaciones repetidas se realizan por teléfono o mediante entrevista personal. Estos métodos, especialmente la encuesta personal, pueden ser contraproducentes ya que generalmente se procede a una encuesta por correo para minimi-

zar el coste y evitar los conglomerados en la muestra que tienden a aumentar la varianza muestral.

En el caso de encuestas telefónicas las encuestaciones repetidas suelen realizarse también por teléfono ya que generalmente esto resulta más barato que la entrevista personal o el envío por correo. Se suelen permitir más repeticiones telefónicas que de cualquier otro tipo. Suelen darse unidades que rechazan la entrevista telefónica y unidades que nunca contestan a las llamadas. En tales casos el (los) intento(s) final(es) debe ser la entrevista personal. Esta como ya se ha señalado antes puede producir una muestra más agrupada que la realizada únicamente por teléfono.

Finalmente, en el caso de entrevistas personales, las encuestaciones repetidas se hacen o volviendo a llamar a las puertas o por teléfono. Si se consigue contactar por teléfono, la llamada se puede utilizar o para obtener los datos o para acordar una hora en la que realizar la entrevista. El número de encuestas repetidas en este caso suele estar limitado por la duración de la recogida de datos que suele ser mucho más corta en encuestas personales que en las realizadas por correo y a veces también que en las telefónicas.

Cualquiera que sea el procedimiento por el que se realiza la encuesta, se debe controlar que las encuestaciones repetidas se lleven a cabo en las unidades seleccionadas a fin de proveer de inputs para un sistema de imputación, estudiar su efecto en la calidad de los datos y determinar estrategias en los diseños muestrales para procedimientos de futuras encuestaciones repetidas. En la etapa de diseño se puede estimar el número de encuestas repetidas sobre la base de la experiencia pasada. Es importante tener la varianza de la información obtenida en las encuestas repetidas y el coste asociado con éstas. Diversos factores tales como las características demográficas y socioeconómicas de los entrevistados, estrategias del entrevistador tales como horario en las repeticiones y coste de obtener la información afectan al éxito de las encuestaciones repetidas.

Un número de estudios llevados a cabo por diversos individuos y organizaciones confirman que:

i) Todas las encuestas muestran que el número de respuestas incrementa sustancialmente cuando crece el número de repeticiones.

ii) En la mayoría de las encuestas, la estimación de

la media después del primer intento de realizar la encuesta difiere sustancialmente de la estimación tras varias reiteraciones.

Se realizaron diversos experimentos en relación con el método de encuestación en las entrevistas repetidas. Por ejemplo, se puede entrevistar a cualquier persona, aleatoriamente o a un individuo específico. Ambos demostraron diferencias en las estimaciones en las distintas repeticiones. Para poder establecer el número óptimo de encuestas repetidas es importante saber el coste de las diversas operaciones a ellas asociadas. Para más detalles véase Durbin Stuart (1954), Drew, Fuller (Actas de ASS, 1980), Cochran, Deming, etc.

ii) Encuestación delegada (Proxy)

La segunda técnica más importante es elegir datos de una unidad alternativa. Por ejemplo las instrucciones de la encuesta pueden indicar que se puede encuestar, en el caso de que no sea posible hacerlo a la persona indicada, a cualquier otro miembro de la familia que tenga más de 20 años. Obviamente, las respuestas tienen mayor probabilidad de ser incorrectas en estos casos. Sin embargo, no es peligroso para cierto tipo de encuestas. Pero si la información requerida es confidencial o personal entonces la encuestación delegada puede ser confusa. Se debe realizar un pre-estudio de la encuestación delegada antes de aplicarla a cada caso particular.

Sin embargo, debido a las restricciones de tiempo y coste es virtualmente imposible obtener respuesta de cada individuo sin recurrir a la encuestación delegada. Por esta razón se acepta la encuestación delegada en la encuesta sobre mano de obra en Canadá. Generalmente sólo se encuesta a una persona de la familia y ésta responde sobre el comportamiento de todos los miembros. A veces, se necesitan entrevistas separadas de los miembros de una casa por ejemplo si hay huéspedes. En promedio se ha encontrado que las encuestas delegadas suponen el 50% de los que responden. Además, al aceptar las encuestas delegadas es posible obtener respuestas completas para todos los miembros de una casa prácticamente en el 100% de los hogares que responden. De hecho, hay menos de un 0,2% de las casas en las que se entrevista (tanto por delegación como sin ella) en las que se obtienen respuestas sobre algunas pero no sobre todos los miembros de la casa. La aceptación de encuestados delegados es por tanto un medio muy efectivo de reducir la no-respuesta.

iii) Rechazos-seguimiento. Siempre que sea posible debe hacerse un seguimiento de los hogares que rechazan

la encuesta. En la mayoría de los casos esto significará que un entrevistador experimentado o el representante de la oficina regional harán de efectuar una visita personal.

En áreas donde no sea posible este tipo de seguimiento, se enviará en estos casos una carta con información adicional sobre la encuesta y sobre el uso que se hará de los datos. Se debe enfatizar la importancia de la encuesta y el valor de la colaboración del encuestado. La experiencia en Canadá muestra que del 20 al 30% de los hogares que rechazaron pudieron ser encuestados al mes siguiente. En el caso de hogares a los que no se les puede persuadir para que respondan se aconseja que no sean visitados de nuevo a no ser que haya un cambio en la composición de la casa.

iv) Teléfono

El procedimiento de entrevistas por teléfono que se usa en la LFS envuelve una combinación de visitas personales y llamadas telefónicas y se lleva a cabo en grandes áreas urbanas. Los entrevistadores han de realizar la primera entrevista personalmente dejando las entrevistas telefónicas para los meses siguientes, siempre que el encuestado esté de acuerdo en contestar por teléfono. La experiencia muestra que aproximadamente el 75% de los hogares pueden ser entrevistados por teléfono, cuando se permite este procedimiento.

Aunque la razón principal para realizar la encuesta por teléfono en la LFS es reducir costos, la entrevista por teléfono tiene ciertas ventajas desde el punto de vista de la no-respuesta. a) Permite que se realice la entrevista en su momento sean cuales sean las condiciones climatológicas. b) Está especialmente indicada para personas solas, familias pequeñas e inquilinos de apartamentos a los que suele ser difícil localizar en casa y a los que se suele encontrar a la noche (es más fácil para el entrevistador telefonear por la noche que ir personalmente). c) Permite realizar la entrevista de la manera que más convenga al que ha de responder (si en un momento no es adecuado se puede quedar para otro). Y d) tiene la posibilidad de reducir la no-respuesta al dar más oportunidades y tiempo para la reiteración de la encuesta.

v) Quota

En el muestreo por cuotas, se estratifica la población según unas variables que se espera que estén muy relacionadas con la respuesta. Por ejemplo, se selecciona una muestra de hogares y se dan instrucciones al entre-

vistador, para que obtenga entrevistas de personas que se encuentran en ciertas categorías de edad, sexo y renta.

El peligro de este procedimiento está en que la selección efectiva la hace el entrevistador entre aquéllos que se encuentran en esas categorías. La cuestión es cuánto riesgo(sesgo) los diseñadores de encuestas están dispuestos a aceptar, frente al coste de encuestas repetidas. Los muestreos por cuota no tienen no-respuesta ya que el entrevistador no registra a aquéllos que intenta entrevistar sin conseguirlo. Además el mecanismo de selección en el muestreo por cuotas está mal definido y es por tanto imposible estimar el error muestral. Los sesgos se deben a diferencias en la disponibilidad del que responde. Su ventaja es que es rápido y barato.

vi) Sustitución en el campo

Un método de reducir el tamaño de la no-respuesta en el momento de recogida de datos es proceder a una sustitución usando unidades que no habían sido seleccionadas para la muestra. Hay dos tipos básicos de sustituciones a usar:

- a) Selección de las unidades aleatoriamente
- b) Selección de sustitutos designados especificados.

a) Selección al azar

Se selecciona probabilísticamente unidades adicionales para reemplazar las no-respuestas. Generalmente los sustitutos de una unidad que no responde se seleccionan a partir de una población restringida, subgrupo (por ejemplo, del mismo bloque, o del mismo estrato) o del grupo de estratos donde hubo no-respuesta). En cada caso se supone que las características del sustituto elegido de un subgrupo se parecerán más a la unidad que no responde que un sustituto escogido de entre toda la población. En muchos procesos aleatorios, los sustitutos potenciales se seleccionan antes de proceder a la recogida de datos de esta forma se emplea menos tiempo en esta última. Hay generalmente dos elecciones operativas i) Cada entrevista tiene un sustituto potencial, ii) Se les comunica a los entrevistadores cuando éstos lo requieren. Cada procedimiento tiene sus ventajas y desventajas.

En la práctica, es mejor proceder a la selección aleatoria de las unidades sustitutas en las oficinas centrales en lugar de dejárselo al entrevistador; de esta forma se evita que haya una selección sesgada. En algunos pro-

cedimientos los sustitutos potenciales se seleccionan previamente a la recogida de datos, de esta forma se evitan retrasos o problemas como sesgos en la selección si se sacan los sustitutos durante o después de la recogida de datos. Se debe enfatizar, sin embargo, que este procedimiento no elimina el sesgo debido a la no-respuesta inicial. Sólo sustituye de manera insesgada una unidad por otra que falló en responder, pudiendo tener o no las mismas características.

Selección de un sustituto específicamente designado

Identifica una o más unidades que se supone que tienen características similares a los que no responden. A veces es una unidad en la proximidad pero no tiene por qué serlo. En el procedimiento que usa sustitutos designados específicamente, se selecciona a menudo al vecino más cercano, debido a esa tendencia de las unidades vecinas, a tener características similares en comparación con unidades seleccionadas al azar. Si el vecino más cercano falla en responder, por ejemplo rechaza la entrevista o no está nunca disponible, se selecciona la siguiente unidad y se repite el procedimiento hasta que una unidad responde. Además del criterio de la distancia para escoger el sustituto, se debe intentar buscar una unidad de características similares a la seleccionada originalmente. Se debe tener cuidado en prevenir la re-selección de una unidad ya seleccionada y que ha respondido. La clave para decidir si merecen la pena los procedimientos de sustitución, es contrastar si el uso de sustitutos provee de mejores valores para los que no responden que aquéllos que se obtienen por otros métodos alternativos de imputación. Al examinar los dos métodos de sustitución, es evidente que la primera regla de sustitución remueve el sesgo muestral pero no remueve el sesgo de la no respuesta. La segunda tiende a remover el sesgo debido a la no-respuesta pero falla en remover el sesgo muestral: Sin duda los procedimientos de sustitución tienen sus ventajas y sus desventajas. La primera ventaja es que resulta una forma conveniente de balancear la muestra en relación con el tamaño muestral. La otra es la reducción de la varianza muestral debido a un incremento en el tamaño muestral efectivo.

Al examinar más profundamente los métodos de sustitución se deben considerar tanto desde el punto de vista operativo como del error medio cuadrático. Por ejemplo, es probable que la proximidad provea de un sustituto mejor que un procedimiento ponderado en el que se sustituya por un promedio. Sin embargo, esto puede no ser cierto si se dispone de más información sobre los que no respon-

den en un grupo particular. Cuando tratamos con una muestra pequeña cuyo diseño es fuertemente estratificado la sustitución suele ser más apropiada que otros métodos. Cuanto más se sabe acerca del que no responde, mejor se puede realizar la sustitución. Pero en grandes encuestas en las que se tiene poca o nula información sobre los que no responden la sustitución suele ser peor que otros procedimientos de ponderación. Una de las mayores desventajas de la sustitución es que los entrevistadores se esfuerzan menos en obtener las primeras entrevistas.

vii) El uso de incentivos

El método está bien establecido y se usa a menudo. Un

incentivo puede ser menos costoso que una llamada adicional. La planificación de incentivo se basa en el coste de las llamadas adicionales que habrían de efectuarse en un esfuerzo por completar la entrevista. Pero el uso de incentivos financieros tiene diversos inconvenientes. El primero es que se puede convertir en una práctica aceptada y el coste de las encuestas se incrementa. El segundo puede ser que para conseguir el incentivo se dé información errónea. Se puede argumentar, sin embargo, que el proveer información es una carga y los encuestados deben recibir alguna compensación.

VII. PROBLEMAS CONCEPTUALES Y TIPOS DE NO RESPUESTA

Al pasar la información de la etapa de recogida de datos a la tabulación se identifican diferentes tipos de respuestas que se presentan en la Tabla 10.

En la Tabla 10, se puede ver que dos de los tres tipos de cuestionarios que provienen de la etapa de depuración previa (edit) requieren retoques antes de proceder a la estimación. Estos son los cuestionarios que no se pueden usar (7), los que contienen algunos blancos o entradas inconsistentes (6) y los que no tienen respuesta (3). Un cuestionario inservible puede provenir de una unidad que no ha respondido o de una unidad que ha respondido pero no a todas las preguntas. En cualquier caso, es necesario la imputación (8). Unidades que no responden son ponderadas de alguna manera excepto en el censo. Los cuestionarios deficientes (no respecto a preguntas) se encuentran en dos categorías (6) entradas no consistentes o blancos.

Las entradas inconsistentes son imposibilidades lógicas o posibilidades altamente improbables. Parece obvio que si las entradas son imposibilidades lógicas al detectarse deberán ajustarse aunque no incidan demasiado en los datos. El ajuste elimina dificultades a los analistas de la materia que publican los datos.

En el caso de entradas plausibles pero improbables, se enfrenta uno a la dificultad de elegir entre dejar datos que parecen provenir de una distribución rara o remover los valores extremos de la distribución que parece representar la situación real. Idealmente se debe optar por una u otra cosa sobre la base de la experiencia con el mecanismo del error y con la naturaleza de la distribución sobre la que se basa. En cualquiera de los casos hay que poder identificar los problemas por ejemplo, sería conveniente disponer de reglas de depuración previas para cuando aparecen sucesos imposibles o altamente improbables y un método de cómo tratarlos (imputación).

Hay una diferencia fundamental entre depuración previa e imputación. Consideremos el conjunto de todas las combinaciones posibles de códigos en un cuestionario, la depuración previa se puede definir como la división del conjunto en dos subconjuntos disjuntos. Las combinaciones que se consideran aceptables y las que se consideran inaceptables, las últimas contienen blancos no válidos y entradas inconsistentes. Así, la depuración previa es básicamente un diagnóstico y operativamente se puede definir mediante un conjunto de reglas. Por otro lado, la imputación pertenece por naturaleza al tratamiento de datos.

La imputación es, tal como la hemos definido, la asignación de datos a campos vacíos (incluyendo la no-respuesta total) o el reemplazamiento de datos incluidos de ciertos campos siguiendo ciertas reglas. No hay un método insesgado conocido de imputación pero algunos métodos son más adecuados que otros.

Es posible, en lugar de imputar la no-respuesta en el momento en que se preparan las tabulaciones de la encuesta, presentar estas informaciones sobre el tamaño de la no-respuesta. En este caso los usuarios podrían elegir entre diversos métodos de imputación a partir de los datos tabulados. A primera vista, esta aproximación tendría ciertas ventajas al dar a los usuarios la oportunidad de seleccionar su propio método de imputación. Hay, sin embargo, serias desventajas. Los usuarios utilizando diversos métodos de imputación darían estimaciones contradictorias, creando problemas de consistencia e integración de los datos. Además, visualmente la agencia que recoge los datos puede decidir mejor el método de imputación dada su proximidad a las fuentes de datos. Por estas razones, generalmente son las agencias las que realizan la imputación en lugar de los usuarios. Toda la filosofía de la imputación se basa en que si el procedimiento es apropiado, ya sea para la no-respuesta o para los blancos que resultan de un fallo detectado en la depuración previa, se obtenga una relación más lógica entre datos clasificaciones cruzadas y se rebaje el error medio cuadrático de los estimadores.

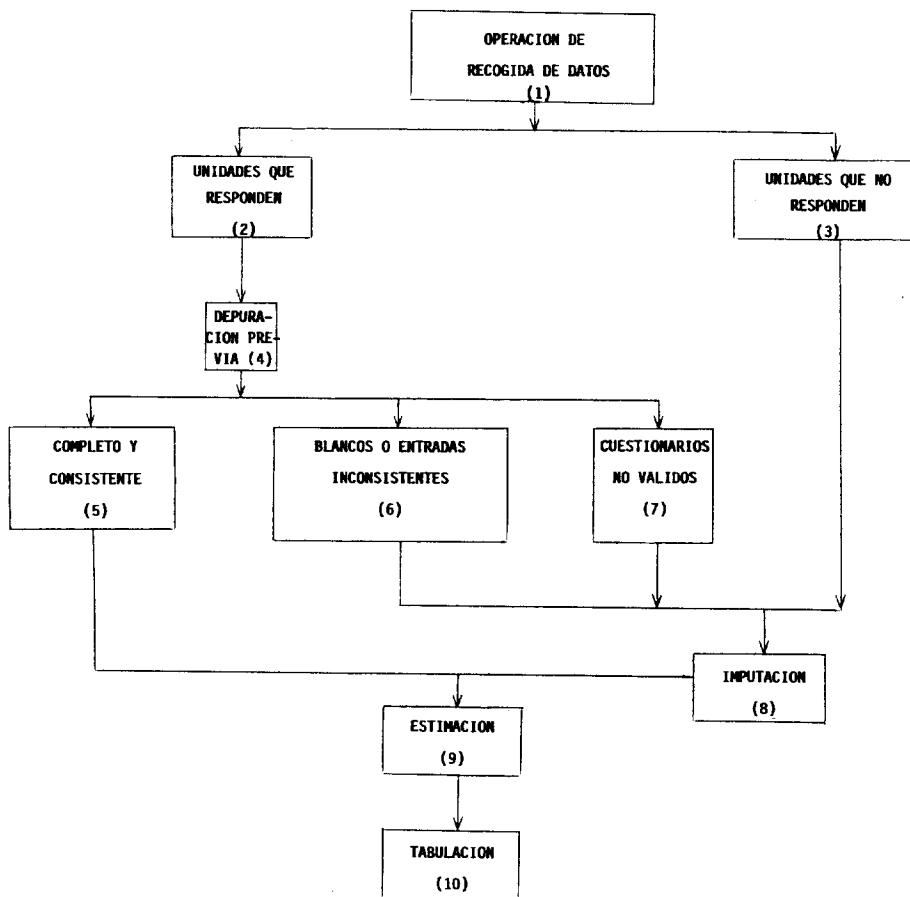
La situación más sencilla se da cuando hay sólo un valor que se puede imputar, en un campo de forma que después de la imputación el valor sea consistente. A este caso se le denomina imputación determinista. Por ejemplo, si la esposa aparece codificada como masculino sólo hay un valor posible a imputar al sexo que sea consistente con el resto de la información. A veces hay más de un valor que lo hace consistente. Si es éste el caso, se elegirá aquel valor particular que es más predominante en relación a la frecuencia total o más plausible. Un ejemplo de este tipo se encuentra en la encuesta sobre mano de obra. Así, si una persona entre 15 y 16 años no ha llenado la característica sobre su actividad laboral en los meses que van de otoño a primavera se le asigna como asistiendo a la escuela, aunque es posible que no asista a la escuela. Mientras la proporción de tales casos sea pequeña, el efecto de esta imputación será un incremento pequeño en el sesgo pero habrá reducción en la varianza.

En otras situaciones cuando se puede razonablemente

imputar un intervalo de valores, necesitamos otros criterios. Uno sería el minimizar el error medio cuadrático de las estimaciones resultantes. La cuestión, sin embargo, es a qué estimaciones nos referimos. Ante el incremento continuo de la demanda de datos tabulados de diferentes e imprevistas maneras, no se sabe qué error medio cuadrático hay que minimizar. Además no se conocen los diferentes agregados a los que unos datos pueden contribuir y sus diferentes formas de tabulación. Consecuentemente se suele preferir usar algún otro criterio que produzca una entrada más apropiada para un campo dentro de

un conjunto particular de datos y en relación con otra información del conjunto. En otras palabras, ¿cómo se puede predecir el mejor valor de un campo sobre la base de conocer los otros campos del conjunto? Un buen ejemplo de este tipo de imputación es el uso de los datos del mes previo en la encuesta sobre mano de obra: para una determinada persona, difícilmente se encontrará un valor imputado mejor, particularmente en aquellos casos en los que características demográficas cambian lentamente. Si no disponemos de información pasada se tiene que recurrir a otros métodos de imputación.

TABLA 10
DIAGRAMA DE FLUJO PARA CADA UNIDAD MUESTREADA



Este es un diagrama muy simplificado del proceso y se expone únicamente para que sirva a la argumentación

VIII. METODOS DE IMPUTACION

Durante varios años se han estado usando diversos procedimientos de imputación para los datos que faltan debidos a la no-respuesta, tanto en encuestas de hogares como en censos. El uso de un determinado procedimiento ha sido justificado en base a la experiencia, a la intuición y oportunidad. Se suponía uniforme la probabilidad de que las unidades respondiesen y se ignoraba frecuentemente el sesgo causado por la no-respuesta. La imputación es el proceso de asignar valores a datos que faltan produciendo así un conjunto de datos completo. Con la imputación al igual que con el ajuste de las ponderaciones se busca reducir el sesgo debido a la no-respuesta. La imputación opera sobre los datos, de forma que los resultados obtenidos por los diferentes análisis sean mutuamente consistentes; lo que no sucede cuando el análisis se realiza sobre un conjunto de datos incompleto. Si bien, y cualquiera que sea el método de imputación utilizado, no podemos asegurar una mejora en el sesgo respecto del sistema de datos incompletos. Al fin y al cabo la imputación es un procedimiento de generar datos. No aumenta la cantidad de información disponible y de hecho el uso de una submuestra aleatoria entre los que responden sirve para reducir en lugar de para incrementar el tamaño de la muestra efectiva. Ya que el muestreo con reemplazamiento tienen una varianza mayor. Sin embargo, podemos esperar que el conocimiento de diversas relaciones y su aplicación a los datos que faltan reducirá el sesgo debido a la no respuesta.

Se suelen aplicar diversos métodos de imputación y seguidamente vamos a describir unos pocos:

i) Imputación deductiva

Actualmente este método se aplica en situaciones en las que las respuestas que faltan se puede deducir del resto de la información proveniente del conjunto de datos. Por ejemplo, si falta el sexo del que responde y las personas tienen nombre femenino y está casada con un varón, se puede deducir que es hembra. Este es un método popular y se usa a menudo.

ii) Fichero caliente (Hot deck)

Es un método usual de ajustar conjuntos de datos para valores no observados y admite diversas variantes. Generalmente el fichero caliente es un procedimiento de duplicación. Cuando falta un valor, se duplica un valor ya existente en la muestra para reemplazarlo. El principal propósito del fichero caliente es reducir el sesgo debido a la no respuesta. Para reducir este sesgo, el procedi-

miento de fichero caliente incorpora un método de clasificación. Todas las unidades muestrales se clasifican en grupos disjuntos de forma que sean lo más homogéneas posible dentro de los grupos. A cada valor que falte, se le asigna un valor del mismo grupo. De modo que el supuesto implícito que se está utilizando es que dentro de cada clasificación la no respuesta sigue la misma distribución que los que responden. Tal supuesto impone una fuerte restricción para las variables de clasificación. Estas variables han de estar correladas con los valores que faltan y con los valores de los que contestan. Si esto no se mantiene el fichero-caliente reduce sólo en parte el sesgo debido a la no-respuesta i) produce un conjunto de datos limpios, esto es, un conjunto de datos completo y claro; ii) reduce el sesgo mientras preservemos las distribuciones conjuntas y marginales. Por ejemplo si sustituyéramos un valor que falte por la media, la distribución de los valores muestrales resultaría afectada. Y si escogiéramos aleatoriamente un valor entre los datos se reduciría la distorsión de la distribución pero no el sesgo.

Como método de imputación los procedimientos de fichero caliente tienen ciertos rasgos atractivos entre los que se encuentran los siguientes: a) los procedimientos conducen a una post-estratificación sencilla; b) no se presentan problemas especiales de encajar conjuntos de datos; c) no se necesitan supuestos fuertes para estimar los valores individuales de las respuestas que faltan.

Sin embargo, los procedimientos de fichero caliente tienen ciertas desventajas. Carecen de un mecanismo de probabilidad, lo que imposibilita calcular su confianza sin algún procedimiento de modelización. Existe la posibilidad de usar varias veces a una misma unidad que ya ha respondido.

Al evaluar los métodos de fichero caliente sería conveniente saber cómo afectan al sesgo y a la confianza de las estimaciones principales el tamaño de los grupos de clasificación (clases de ponderación), la frecuencia de los datos que faltan y la elección de los items de encaje. Fellegi y Holt, Bailar y Bailar (1978) y Cox B. y Falsom R.E. (1978) han realizado trabajos teóricos en relación con los ficheros calientes.

iii) Fichero caliente modificado

La esencia de este método es la clasificación y encaje de donantes potenciales y receptores utilizando

un considerable número de variables. El encaje se hace sobre bases jerárquicas en el siguiente sentido: si no se encuentra un donante para encajar con un receptor en todas las variables de control, se eliminan algunas variables consideradas como menos importantes para de esta forma, conseguir el encaje a un nivel inferior. Los niveles jerárquicos se definen con este propósito siendo el nivel más bajo aquel en el que se encuentra un donante.

iv) Regresión

Mediante este procedimiento la pregunta y a la cual hay que imputar valores, se le regresa en las variables de control (X_1, X_2, \dots, X_p). Tomando los datos de unidades en las que sí hay respuesta para y . Las variables de control pueden ser tanto cuantitativas como cualitativas. Las variables cualitativas se incorporan al modelo de regresión por medio de variables ficticias. A través de un desarrollo cuidadoso de un modelo explícito, la imputación por regresión potencialmente produce valores más cercanos al verdadero que otros métodos de imputación. La construcción de un buen modelo de regresión es costosa consume tiempo y no es realista considerar su aplicación a todas las preguntas con datos faltantes de una encuesta. Parece mejor reservar esta técnica para las variables principales.

v) Ponderación

Uno de los procedimientos más usuales de tratar la no respuesta en la etapa de procesamiento y estimación es el área de balance que depende del diseño, en las que se inflan las ponderaciones mediante la inversa de la tasa de respuesta. En un área de balance b , una estimación del total de una característica viene dado por $X_b = \sum_{i=1}^{n_b} x_i / \pi_i$ donde x_i es la respuesta, π_i es la probabilidad de inclusión, n_b es el tamaño muestral en las áreas de balance. Si sólo responden m_b unidades entonces la ponderación π_i^{-1} es inflada por la inversa de la tasa de respuestas m_b/n_b o sea por n_b/m_b .

Las áreas de balance deben determinarse preferentemente en la etapa de planificación en lugar de en la etapa de procesamiento y pueden ser estratos individuales, grupos de estratos, una provincia, una unidad muestral primaria o un conglomerado. La elección de un área de balance es bastante crucial ya que las tasas de no-respuesta y sesgo pueden diferir de área en área.

Un problema metodológico importante lo constituye la determinación de un tamaño óptimo, o por lo menos apro-

piado, de un área de balance. Por apropiado entendemos un tamaño adecuado que nos asegure tanto una tasa de respuesta suficiente, como las ventajas debidas a las medidas de homogeneidad que ayuden a reducir el sesgo de la no-respuesta. Se puede mostrar que la inflación de peso asumida en todos los conjuntos de datos en un área de balance, para compensar a los que no responden es equivalente a sustituir cada elemento que no responde de ese área por el valor medio ponderado de los que responden. Si una característica tiene una medida de homogeneidad alta (que se incremente al decrecer el tamaño del área), entonces la ponderación (o sustitución por el valor medio) en áreas pequeñas tenderá a dar valores medios más similares al valor real de la característica de los que no responden en aquellas áreas que en áreas mayores. Por tanto el sesgo debido a la no-respuesta tenderá a ser menor en el caso de áreas pequeñas de balance que en áreas grandes. ¿Y qué pasa con la varianza? Según las áreas de balance se van haciendo más pequeñas, la ponderación por la inflación se va haciendo más variable ya que es más inestable la variación en las tasas de respuesta entre muchas áreas pequeñas de balance en contraposición a pocas y grandes áreas de balance y esta inestabilidad en las ponderaciones hará incrementar la varianza. Claramente hay una contraposición al determinar el tamaño de las áreas de balance entre áreas pequeñas con la ventaja de la homogeneidad y áreas grandes para asegurar la estabilidad en los ajustes de las ponderaciones. Los dos valores extremos serían por un lado toda la muestra y por el otro las de tamaño uno. Sin embargo, en este último caso se nos plantea el problema de que hacer si la unidad no responde. En lugar de sustituir por el valor medio tendríamos que recurrir al análisis de regresión o a modelos de superpoblación, (como una aproximación totalmente diferente a la sustitución o a emplear valores históricos). La elección del tamaño de las áreas de balance depende no sólo de la medida de homogeneidad sino también del diseño muestral, del tamaño muestral y de la tasa de respuesta. Encuestas con tasas de respuesta bajas requerirán áreas de balance mayores que aquéllas con altas tasas de respuesta. Se podría empezar por áreas pequeñas de balance y adoptar un procedimiento para englobarlas en otras si la tasa de respuesta no alcanza un cierto nivel (que no debería situarse muy por debajo de la tasa de respuesta global), de forma que se minimiza la inestabilidad de la ponderación. El juntar o englobar áreas de balance añade una nueva dificultad a la estimación de la varianza ya que habría que tener en cuenta las probabilidades de agrupar 1, 2, 3, 4, ... áreas de balance así como de elección de 1, 2, 3 ó 4 áreas de balance. Este procedimiento se lleva a cabo en la encuesta sobre fuerza de trabajo; la necesidad de agrupar es demasiado frecuente

como para tener que garantizar un tratamiento especial en estos casos para estimar la varianza. Consecuentemente si se van a realizar cálculos de varianza u otros análisis distintos que simples promedios o totales, se deben incorporar al diseño muestral aquellas áreas de balance que se espera sean estables. Si se considera la estimación de la varianza, para evitar la necesidad de agrupar áreas de balance la probabilidad de tasas de respuesta suficientemente altas deberá ser elevada. Esto desanimará de utilizar áreas de balance pequeñas para la no respuesta.

vi) Duplicación

En lugar de ponderar por la inversa en la tasa de respuesta en las áreas de balance, se podría duplicar un número de unidades suficientes m_b' entre los que no responden y de esta forma alcanzar el nivel original de n_b unidades. Sin embargo, se puede demostrar que esta forma de actuar añade un nuevo componente a la varianza que obtenemos si utilizamos ponderaciones.

La principal ventaja de la duplicación sobre la ponderación radica en que nos asegura la obtención de ponderaciones integrales, en lugar de fraccionadas, aplicadas a cada conjunto de datos. En ciertos tipos de datos por ejemplo, el número de personas con ciertas características, las ponderaciones integrales tienden a evitar errores de redondeo en los datos de subclasificaciones. Cuando se estiman medias, proporciones o ciertos tipos de totales cuantitativos como el producto nacional bruto, el uso en ponderaciones integrales en lugar de fraccionales no ofrece ventaja.

A parte de los comentarios hechos en el párrafo anterior, los problemas metodológicos relacionados con las ponderaciones en las áreas de balance se aplican también a la duplicación en las áreas de balance.

vii) Sustitución

Otro método importante de imputación es la sustitución de la no-respuesta por datos históricos (los del mes anterior) o por datos provenientes de una fuente externa (ficheros de la administración, otras encuestas, datos del censo). Una vez que se ha procedido a la sustitución de la mayor parte posible que no responde por datos históricos o externos, la ponderación o duplicación se puede ver afectada dentro de las áreas de balance. En el caso de ponderaciones, se inflaría la ponderación π_i^{-1} por el factor $n_b(m_b+m_b')$, donde los m_b respondentes se obtuvieron como antes y de los (n_b-m_b) no respondientes se imputaron

datos históricos a m_b' . Con este método de imputación la varianza muestral se reduce respecto de la que se obtiene en un esquema de ponderaciones ya que el tamaño muestral se incrementa de m_b a algún valor entre m_b y (m_b+m_b') . Esta muestra incrementada, o sea la que incluye valores históricos, no será tan buena como sería la auténtica formada por (m_b+m_b') datos ya que los datos históricos o de fuentes externas no suelen coincidir totalmente con la información actual al menos que las características de las unidades en las que se ha procedido a la sustitución no hayan cambiado.

Alternativamente se puede desechar duplicar los que responden, por ejemplo tomar una muestra de respondientes de igual tamaño que el número de no-respondentes y aplicar una ponderación de 2 en lugar de inflar la ponderación para todos los respondientes. Sin embargo, si se prefiere no duplicar al grupo que no responde para el que se ha efectuado sustitución con información histórica, entonces se deberá submuestrear de entre los m_b que responden $n_b-(m_b+m_b')$, que denotamos m_b^* de forma de conseguir una muestra de n_b unidades en el área de balance (b) a partir de (m_b+m_b') unidades. El total estimado para un área de balance b sería:

$$\hat{x}_b = \sum_{i=1}^{m_b} w_i x_i / \pi_i + \sum_{j=m_b+1}^{m_b+m_b'} x_j' / \pi_j$$

donde x_j' es el valor que se imputa a la unidad j y $w_i = 1$ ó 2 (2 cuando se toma una submuestra aleatoria de $n_b-m_b-m_b'$ unidades entre los m_b que responden). El valor esperado de \hat{x}_b en todas las formas posibles de duplicación es:

$$\hat{x}_b^* = (n_b - m_b')/m_b \cdot \sum_{i=1}^{m_b} x_i / \pi_i + \sum_{j=m_b+1}^{m_b+m_b'} x_j' / \pi_j$$

Consecuentemente, $V(\hat{x}_b) = V(\hat{x}_b^*) + (\text{componente adicional de varianza que resulte de submuestrear entre los respondientes})$. \hat{x}_b^* no es lo mismo que la estimación

$$n_b/(m_b+m_b') \left[\sum_{i=1}^b x_i / \pi_i + \sum_{j=m_b+1}^{m_b+m_b'} x_j' / \pi_j \right]$$

y la varianza de \hat{X}_b^* es diferente de la estimación cuando se aplica la ponderación inflada $n_b/m_b + m_b'$.

El procedimiento de estimación trata con todo lo anterior incluyendo ponderaciones o duplicaciones en áreas de balance que dependen del diseño. Si se dispone de datos históricos o externos para algunos de los que no responden, éstos se podrían utilizar en la imputación de las áreas de balance antes que otros métodos como ponderación o duplicación. En lugar de áreas de balance, se pueden utilizar clases de ponderaciones con propósitos de imputación y éstas se discuten en el siguiente párrafo.

Las clases de ponderaciones se distinguen de las áreas de balance en que las primeras comprenden características de las unidades últimas (tipos de vivienda, grupos especiales de renta, ...) como opuestas a las áreas geográficas aunque se podrían concebir grupos de áreas según determinadas características que no estén seleccionadas con el diseño muestral. Generalmente se definen las clases de ponderaciones, así como las áreas de balance, antes de la etapa de recogida de datos y luego se hacen ajustes agrupando áreas si las tasas de respuesta son muy bajas o la muestra demasiado pequeña para emplear cualquier otro tipo de ajuste de las ponderaciones. En algunos procedimientos de imputación, sin embargo, se definen las clases de ponderación después de la etapa de recogida de datos y se emplean técnicas analíticas como el análisis factorial para determinar el conjunto más eficiente de ponderaciones. Despues de determinar las clases de ponderaciones, los procedimientos de estimación son similares a los que se usan en las áreas de balance. Los sesgos y varianzas (al menos en términos de probabilidades de inclusión individuales y conjuntos de las unidades últimas y de otros parámetros no relacionados con el diseño) son idénticos. Sin embargo, a menos que la varianza tenga en cuenta el diseño muestral, las varianzas de las estimaciones que pertenecen a áreas de balance y a clases de ponderaciones serán bastante diferentes. Para utilizar clases de ponderaciones en el procedimiento de imputación se requiere saber ciertas cosas de los que no responden (tales como clase de renta, tamaño del hogar). En la práctica este procedimiento no se puede usar si no se dispone de esta información.

La fórmula de estimación utilizada en los métodos de imputación aquí discutidos es:

$X = \sum_b X_b$ estima el total de una característica donde b es el área de balance o la clase de ponderación. La estimación para un área dada de balance a una clase de ponderación viene dada por:

$$X_b = \sum_{i=1}^{m_b} w_i x_i / \Pi_i + \sum_{j=m_b+1}^{m_b+m_b'} w_j x_j / \Pi_j$$

donde Π_i o Π_j es la probabilidad de inclusión y m_b es el número de unidades que responden entre las n_b unidades del área de balance b .

x_i = valor en la respuesta de la unidad $i = X_i$ (valor verdadero + $R\epsilon_i$ (error en la respuesta)

x_j = valor histórico para la unidad j (si es que se dispone de este valor) asignado a la unidad j que no respondió. De entre las $(n_b - m_b)$ unidades que responden m_b tienen datos históricos en el área de balance b .

$x'_j = X_j$ (valor verdadero) + $R\epsilon_j$ (error en la respuesta o en el valor histórico en relación a X_j).

w_i y w_j son ponderaciones, son diferentes según el método.

viii) Tasa RAD (Ranking Ratio)

Este procedimiento puede ser útil como forma de ajuste cuando las clases a ponderar (clases de ponderación) se definen en términos de tablas de contingencia de ciertas características. La base subyacente de la técnica es hacer que las distribuciones marginales ponderadas se comporten igual que en la población. Hay dos razones fuertes para utilizar la tasa RAD: 1) Cuando se forman muchas clases mediante tabulaciones cruzadas de las características, el ajuste de las ponderaciones de cada clase por separado puede resultar muy variable e inestable debido al tamaño pequeño de las clases. 2) El ajustar las ponderaciones de cada clase requiere un conocimiento de la distribución conjunta de las características de las clases en la población, mientras que la tasa RAD sólo requiere el conocimiento de las distribuciones marginales de la población.

Las W 's constituyen la distribución poblacional de ambas características, mientras que las q 's constituyen la distribución muestral.

Edad de la Población				Edad de la muestra			
1	2K	Total	1	2	..K	Total
1	w_{11}	w_{12}	w_{1K}	$w_{1.}$	1	q_{11}	q_{12}
2	w_{21}	w_{22}	w_{2K}	$w_{2.}$	2	q_{21}	q_{22}
	$w_{.1}$	$w_{.2}$	$w_{.K}$	1		$q_{.1}$	$q_{.2..}$
						$q_{.K}$	1

La tasa RAD emplea un procedimiento iterativo para hacer que las ponderaciones de las distribuciones marginales muestrales ($q_{.1} \dots$)

($q_{.2} \dots$) se igualen a las distribuciones marginales de la población ($w_{.1} \dots$)

($w_{.2} \dots$)

El sesgo en la tasa RAD depende de la proximidad de las distribuciones de las células entre dos poblaciones y de la diferencia en las medias entre los que responden y no responden en una célula.

TABLA 11
PROCEDIMIENTOS PARA TRATAR LA NO RESPUESTA E IMPLICACIONES

<u>Etapa de la encuesta</u>	<u>Tratamiento de los datos que faltan</u>	<u>Implicaciones</u>
Planificación	<ul style="list-style-type: none"> - Tasas de no respuesta prevista a partir de la experiencia obtenida en encuestas similares respecto al método de recogida de datos. - Se incrementa la muestra para permitir la no-respuesta (sobre muestreo). - Estudios piloto para contrastar el cuestionario y los trabajos de campo. 	<ul style="list-style-type: none"> - Se determinan diversas estrategias en la etapa de campo y procesamiento (encuesta repetida, imputación, etc.). - La varianza muestral se comportará como se esperaba; sin embargo sigue habiendo sesgo debido a la no-respuesta. - Se intenta reducir el error en las respuestas así como las no-respuestas a una pregunta.
Entrenamiento de los encuestadores	<ul style="list-style-type: none"> - Pruebas de aptitud, guión de la encuesta, asignación agrupada que les permite encuestaciones repetidas. 	<ul style="list-style-type: none"> - Los entrevistadores experimentados podrán minimizar la no-respuesta y por tanto el sesgo sobre todo en las encuestas continuas. En encuestas ad hoc poco puede uno hacer excepto encuestaciones repetidas que aumentan los costos por unidad seleccionada.
Recogida de datos	<ul style="list-style-type: none"> - Intentos de encuestaciones repetidas hasta que se obtienen las máximas respuestas posibles; intentos de evitar los rechazos por medio de diplomacia y medios de comunicación. - Sustitución por otras unidades de campo no seleccionadas originalmente. - Sustitución por otras unidades de un grupo de unidades reservadas (incremento de la muestra para compensar la no-respuesta). - Submuestreo entre los que no responden. - Se ignora la no-respuesta en el campo. 	<ul style="list-style-type: none"> - Más encuestas repetidas incrementan el costo por unidad, especialmente si no se puede determinar el momento propicio a través de los vecinos; al reducir el nivel de no-respuesta se reduce la varianza muestral y el sesgo debido a la no-respuesta. - Reducen la varianza muestral pero pueden existir sesgos muestrales y de no respuesta. - Reduce la varianza muestral y el sesgo no muestral pero puede seguir habiendo sesgo debido a la no-respuesta. - Se incrementa el coste debido al esfuerzo extra que se requiere y se reduce el sesgo de no-respuesta aunque no se elimina por completo. - Se reduce la varianza muestral, posibles subestimaciones de los totales a menos que se ajusten las ponderaciones en la fase de procesamiento. - Reducción del error de no-respuesta, nosu eliminación mediante un método apropiado de imputación. - Reducción en la varianza muestral a través de un estimador de la razón o usando fuentes de datos independientes para la imputación. - Incremento ligero en el coste del proceso debido a un procesamiento posiblemente complejo para tratar los datos que faltan. - Dificultades en el análisis de los datos de una encuesta compleja más allá de lo que proviene de la muestra SRS, debido a datos que faltan y a las imputaciones realizadas por ellos. - Dificultad de avisar a los usuarios si hay datos faltantes cuando éstos son analizados.
Etapa de procesamiento y estimación	<ul style="list-style-type: none"> - Imputación de la no-respuesta. - Formación de las células de ajuste. - Ajuste de ponderaciones o sustitución explícita de los datos que faltan. 	

IX. APROXIMACION A LA PROBABILIDAD DE RESPUESTA

Aunque se han detectado variaciones en las tasas de respuesta entre unidades según sus características no se ha examinado suficientemente el efecto sobre el sesgo y la varianza de las unidades individuales, tanto de los que responden como de los que no responden.

Para facilitar un examen más detallado de este efecto Platek y Grey (1978) han desarrollado una metodología que trata el sesgo y la varianza que se corresponde con cada uno de los métodos de imputación. El desarrollo de la expresión para el sesgo y la varianza de las estimaciones se basa en el concepto fundamental de que una unidad, si es seleccionada, responde o no con una cierta probabilidad de respuesta incorporada a cada unidad. Es una extensión de la aproximación desarrollada por Platek, Singh y Tremblay en 1977 con respecto a los censos.

i) El concepto de probabilidad de respuesta

Un examen concienzudo del comportamiento de las tasas de respuestas en diversas encuestas bajo diferentes condiciones nos lleva a un supuesto natural de que en la población, cada unidad tiene una probabilidad de responder (si es seleccionada), dadas las condiciones de la encuesta. En otras palabras en la mayoría de las situaciones, la población no está formada de una dicotomía de gente que responde/gente que no responde sino de un conjunto potencial de encuestados, algunos de los cuales responderán y otros no, dependiendo de las condiciones de la encuesta. Este concepto se encuentra implícito en la aplicación de la técnica de Politz-Simon S. (1949).

Hemos definido la probabilidad de respuesta como la probabilidad de que una unidad seleccionada provea de una respuesta completa o parcial al cuestionario de una encuesta. La probabilidad de respuesta puede estar relacionada con el momento del día en que se realiza la entrevista. Puede también estar correlacionada con características como edad, sexo, nivel de renta del que responde, tamaño del hogar (número de personas que viven en él), procedimiento de entrevista, tema de la encuesta e interés por él. Las probabilidades de respuesta son generalmente más altas en las entrevistas personales que en las encuestas telefónicas o por correo pero no siempre.

En la definición de la probabilidad de respuesta de una unidad del párrafo anterior quedó implícito que una unidad que responde puede o no contestar una pregunta determinada. Por tanto, podemos definir una probabilidad de respuesta a un ítem de una determinada pregunta como la probabilidad de que una unidad seleccionada y que res-

ponde responda a un determinado ítem. Esta probabilidad depende de la complejidad de la pregunta y por lo menos en los casos de encuestas personales y telefónicas de la habilidad del encuestador para explicar o clarificar las preguntas al que responde.

Para desarrollar más el concepto de probabilidad de respuesta, examinamos la posible correspondencia entre las tasas de respuesta que dependen de la característica y las diferencias en probabilidades de respuesta para una característica particular. Consideraremos una situación de una enumeración completa (censo) de una población de N unidades de las cuales Y tienen la característica y y $N-Y$ no la tienen. Hay que estimar el número total de unidades con la característica y . Supongamos sin embargo, que en el momento de la enumeración, tenemos $R < N$ que responden y $M = (N-R)$ que no responden. Supongamos además que de los R que responden están divididos en Y_R con la característica y y $(R-Y_R)$ que no la tienen. Supongamos también que los M que no responden estén divididos en Y_M con la característica y y $(M-Y_M)$ que no tienen la característica. Se divide por tanto la población en cuatro categorías tal y como aparecen en la Tabla 12.

TABLA 12
CENSO DE POBLACION PARA ESTIMAR Y
(Situación sencilla)

	Responden	No responden	Total
Tienen la característica	Y_R	Y_M	Y
No tienen la característica	$(R - Y_R)$	$(M - Y_M)$	$(N - Y)$
TOTAL	R	M	$N = (R + M)$

El parámetro que es desconocido, excepto cuando hay respuesta completa, se puede expresar de dos formas:

$$\text{i) } N(Y/N) \text{ ó ii) } Y_R(Y/Y_R)$$

De la tabla anterior, se puede ver que la tasa total de respuesta= R/N es también la probabilidad media de respuesta de N unidades. Similamente, la tasa de respuesta entre personas con la característica igual a Y_R/Y , que iguala a la probabilidad media de respuestas entre todas las unidades con característica y .

Entre las personas que no tienen la característica, una expresión similar para la tasa de no-respuesta, o de la probabilidad media de no-respuesta iguala a $(R-Y_R)/(N-Y)$.

Estas tasas pueden o no igualar a la tasa de respuesta total N/N . La extensión en la que la tasa se desvía de R/N determina el sesgo de la no-respuesta de la estimación.

El valor total Y es desconocido excepto cuando hay respuesta total pero se puede estimar mediante cualquiera de las siguientes expresiones equivalentes:

$$(a) \tilde{Y} = N(Y_R/R)$$

$$(b) \tilde{Y} = Y_R(N/R)$$

en el caso a) (Y_R/R) estima la proporción (Y/N) de personas con la característica y , suponiendo que tienen las mismas características los que responden y los que no responden, o sea, se supone que $(Y_R/R) = (Y/N) = (Y_M/M)$.

Alternativamente en b) (N/R) la inversa de la tasa de respuesta total estima la tasa de respuesta inversa (Y/Y_R) para personas que tienen la característica bajo el supuesto de que las tasas de respuesta (o las probabilidades promedio en respuesta) son idénticas entre las personas que tienen la característica y las que no la tienen. La estimación del parámetro Y se podría escribir como $\tilde{Y} = (Y_R/R)$ bajo a), que es obviamente a lo mismo que se llega bajo b), propiamente $(N/R)Y_R$.

ii) Diversas aproximaciones a la probabilidad de respuesta

El uso de una determinada aproximación dentro del concepto de probabilidad de respuesta tiene consecuencias teóricas y prácticas que trataremos brevemente más adelante.

a) Dicotomía: Supongamos que la población se divide en estratos de encuestas con probabilidad 1 para los que responden y 0 para los que no responden. Bajo este supuesto habría sesgo si las características de interés estuviesen correlacionadas con las probabilidades de respuesta 1 ó 0. El sesgo de no respuesta que ocurre bajo este supuesto se reduce a la diferencia entre las dos medias que toma la característica de interés dentro del conjunto de los que responden y en el conjunto de los que no responden. Esta diferencia se puede detectar mediante un submuestreo entre los que no responden, sin embargo, el hacer esto implica que hay una probabilidad distinta de cero entre los que no responden, una contradicción para el

supuesto dicotómico.

b) Probabilidad constante: Supone que cada unidad de la población objetivo tiene la misma probabilidad de respuesta. Así, una submuestra entre los que responden es idéntica a la muestra de los que responden.

c) Probabilidades variables: Supongamos que la población objetivo está compuesta de respondientes, cada uno teniendo ciertas probabilidades de respuesta según las condiciones de la encuesta. Bajo este supuesto, el sesgo en la estimación dependerá de la magnitud de la correlación entre la característica de interés y la probabilidad de respuesta de las unidades.

iii) Implicaciones teóricas de la probabilidad de respuesta

a) Sesgo de respuesta: El sesgo de no respuesta, al que tan a menudo se hace referencia en la literatura sobre el tema, suele referirse a las estimaciones de estadísticos como medias, totales, ..., después de compensar los datos que faltan debido a la no-respuesta. Un método común de imputación es el ajuste de ponderaciones por la inversa en la tasa de respuesta en una célula de ajuste. Se puede mostrar que el sesgo de no respuesta de una estimación es proporcional a la correlación entre magnitudes de probabilidades de respuestas y valores de los estadísticos que se están estimando. Bajo una aproximación dicotómica, las probabilidades de respuesta son 0, 1 y la correlación existe cuando las características tienden a diferir entre los que responden y los que no responden. En la literatura en la que se discute el sesgo de la no respuesta, se supone generalmente la aproximación dicotómica, en la que se supone que los que responden y los que no responden forman estratos separados y las diferencias en sus características se analizan teóricamente usando otras fuentes. Platek, Singh y Tremblay (1977) y Platek y Grey (1983) en la aproximación mediante probabilidad variable a la respuesta derivan una aproximación más general al sesgo de la no-respuesta, de la que un caso particular es el dicotómico. En el caso de probabilidades iguales de respuesta, en el que se usan ponderaciones para inflar la submuestra que responde hasta llegar a la muestra total, no hay sesgo debido a la no-respuesta.

b) Error medio cuadrático: La aproximación a la probabilidad de respuesta posibilita desarrollar una teoría general de los componentes de muestreo y no muestreo de la varianza en el contexto de la no-respuesta y de su imputación. La teoría se puede derivar a partir del mues-

treo, y modelos de error tanto de respuesta como de imputación difieren según los diferentes procedimientos de imputación de los datos que falten. El desarrollo es más general bajo la probabilidad variable de respuesta que bajo probabilidad igual o bajo probabilidad dicotómica de respuesta. Estos dos casos se pueden derivar como casos especiales del de probabilidad variable de respuesta pero ésta última aproximación no se puede derivar a partir de las otras dos.

c) Estudios de asignación y formación de las células de ajuste: Las células de ajuste son importantes en el área de imputación. En lugar de emplear, por ejemplo, ajustes de ponderaciones para pasar de la muestra de los que responden a la total, los ajustes de ponderaciones se pueden realizar en células separadas definidas por submuestras separadas en las células de post estratos de unidades primarias de muestreo. A las células de post estratos se les suele llamar clases de ponderaciones, mientras que en las que dependen del diseño se les llama áreas de balance. Los procedimientos de imputación son idénticos para los dos tipos. Sin embargo las expresiones del error cuadrático medio, particularmente de la varianza muestral de las estimaciones basada en los dos tipos de células, son bastante diferentes.

iv) Implicaciones prácticas de la aproximación a la probabilidad de respuesta

La aproximación por la probabilidad variable de respuesta es más realista que las otras aproximaciones, la dicotomía o la probabilidad constante de respuesta. Sólo una pequeña parte de la población de la muestra seleccionada no tendría posibilidad de responder (duros) o total probabilidad de responder (gente siempre disponible y muy cooperadores con el encuestador). Muchos rechazos duros se convertirían en posibles respondientes a través del uso de incentivos a la respuesta (Gewer, 1977), tacto del entrevistador y relaciones públicas. Otro supuesto poco realista son las probabilidades iguales de respuesta de todas las unidades, ya que en la práctica, es difícil definir células de ajuste que contengan unidades con probabilidades de respuesta iguales.

Sin embargo, la aplicación de un ajuste de ponderación fijo en una célula por la inversa de la tasa de respuesta implica el supuesto de probabilidades constantes de respuesta para todas las unidades en una célula. La consecuencia es un sesgo de no-respuesta, dependiendo de las correlaciones entre las probabilidades de respuesta actuales y el valor de la característica en cada célula.

Las implicaciones prácticas de la aproximación a la

probabilidad de respuesta son inherentes a los procedimientos de imputación de fichero caliente, particularmente para no-respuesta a un ítem pero también para no-respuesta de toda una unidad. En un procedimiento de fichero caliente, las unidades se agrupan según tipos en células, aplicando algún criterio de similaridad. Cuando falta un dato por no respuesta ya sea de unidad o de ítem, se usa un donante de la célula para imputar el valor que falte. La selección del donante puede ser sistemática o aleatoria, a menudo sobre la base de una función de distancia sobre pares de unidades. Se suelen escoger conglomerados, medidos por algún tipo de distancia como proximidad geográfica para la elección de donantes. Sin embargo, a través de la formación de una célula de donantes, la tasa de respuesta, que es generalmente la tasa de respuesta a un ítem, es una estimación de la probabilidad promedio de respuesta a un ítem, en la célula donde están los datos del donante potencial.

Usando la estimación de Horvitz-Thompson, Platek y Gray han desarrollado una fórmula general de imputación para \tilde{Y} .

$$\tilde{Y}(G) = \sum_{i=1}^N t_i w_i w_{ai} \{ \delta_{iy} [\delta_{iy} y_i + (1-\delta_{iy}) z_{iy}^I] + (1-\delta_{iy}) \cdot z_{iy} \}$$

$t_i = 1$ ó 0 según se seleccione o no a i.

w_i = peso muestral aplicado a la unidad i, cuando es - seleccionada.

δ_i = 1 ó 0 según la unidad i seleccionada responda o no $(Pr(\delta_i = 1 | t_i = 1) = a_i)$

δ_{iy} = 1 ó 0 según la unidad respondiente i responda o - no a la característica y $Pr(\delta_{iy} = 1 | t_i = 1, \delta_i = 1) = a_{iy}$

w_{ai} = ajuste de ponderación para compensar la no res - puesta de una unidad.

y_i = valor observado de la característica y para la uni - dad i, como se definió antes en (4.6).

z_{iy}^I = valor imputado a la característica y por la no - iy respuesta a un ítem.

z_{iy} = valor imputado a la característica y por la no - respuesta a una unidad.

X. APPLICACIONES DE LOS AJUSTES POR NO RESPUESTA EN LA ENCUESTA DE FUERZA DE TRABAJO EN CANADÁ

Los procedimientos de ajuste que suelen usarse en encuestas muestrales y censos incluyen ajustes de ponderaciones, duplicación, sustitución por datos externos o históricos y ficheros calientes.

La utilización de estos procedimientos se ilustrará en relación con encuestas tan importantes como la de Fuerza de Trabajo en Canadá y otras grandes encuestas. Se darán ejemplos específicos y se seguirán cuestionarios completados parcialmente.

Una de las encuestas más importantes y llevada a cabo continuamente por la Oficina Estadística de Canadá es la de fuerza de trabajo en la cual se obtienen estimaciones mensuales de desempleo, empleo y de otras características. Los datos se publican en menos de tres semanas, después de haber encuestado áreas de 56.000 hogares, aproximadamente la tercera semana de cada mes. Es imposible contactar todos los hogares que se debería debido al horario estricto en la recogida y procesamiento de datos. En algunas casas no hay nadie durante toda la semana, en otras están ausentes siempre que van los entrevistadores y otras rehusan el ser entrevistados. Hay desde luego otras razones para la no respuesta pero son de menor importancia. La no-respuesta suele mantenerse alrededor del 5% excepto en los meses de verano que se incrementa al 7 y al 8% debido a las ausencias temporales por vacaciones.

Los ajustes por la no respuesta en los hogares se suelen llevar a cabo con los siguientes criterios: i) Se sustituye alrededor de un tercio de los que no responden con valores del mes anterior si se puede (con transformaciones adecuadas en ciertos campos para actualizar los datos del mes pasado). ii) Inflación por la inversa de la tasa de respuesta en las unidades de balance (la respuesta en este caso incluye a los valores de los que no han respondido en esta encuesta pero cuyos valores han sido sustituidos por valores que sí respondieron en la encuesta precedente y cuándo es posible esta sustitución). En el caso ii), la imputación de los 2/3 de no-respuesta que queda, es implícitamente la media de los que respondieron en la unidad de balance (unidades muestrales primarias, zonas rurales y urbanas pequeñas de unidades muestrales primarias o subunidades de grandes ciudades).

El ajuste por la no-respuesta a un ítem o de un rechazo en la fase de depuración previa se lleva a cabo de tres maneras, dependiendo del ítem o ítems con datos

que faltan o defectuosos, dependiendo del status de la respuesta y de las características de la unidad en la encuesta previa.

i) La respuesta omitida de un ítem se puede deducir sin ambigüedad del resto del cuestionario (una tabla de decisión aseguraría una respuesta única y consistente).

ii) La sustitución por la respuesta a un ítem de una encuesta previa está disponible y es adecuada según una tabla de decisión.

iii) Se aplica un fichero caliente cuando se tiene un conjunto de datos similares en la misma unidad muestral primaria, con la misma sucesión de preguntas (una de las seis posibles) y del mismo grupo de edad y sexo. Aquí el agrupar clases de ponderaciones puede requerir agrupar datos similares; generalmente se forman categorías de edad y sexo en lugar de unidades muestrales primarias o sucesiones de preguntas, para realizar agrupaciones con propósitos de ajuste cuando es necesario.

Para ilustrar los procedimientos de imputación de cuestionarios contestados parcialmente consideremos los siguientes ejemplos que corresponden a la encuesta sobre Fuerza de Trabajo en Canadá (LFS).

Ejemplo 1: (Fuerza de Trabajo en Canadá)

A cada persona de 15 años o más, dentro de los hogares seleccionados para la LFS se le pregunta la cuestión 80 (Q80). La semana pasada ha ido a la escuela o a la universidad?. Si ha ido, la respuesta se codifica como 1, si no como 2. A aquéllos que no han ido a la escuela no se les realizan otras preguntas relacionadas. Sin embargo, a aquéllos que van a la escuela se les realiza las siguientes preguntas: Pregunta 81(Q81) ¿Ha sido estudiante a tiempo parcial o total?. La cuestión 82 (Q82) es ¿A qué tipo de escuela fue? Se pueden dar las siguientes situaciones:

$Q80 \neq 1 \text{ ó } 2$ que es un error.

La relación entre las preguntas debe ser la siguiente:

- a) Si $Q80 = 1$, va a la escuela entonces $Q81 = 1 \text{ ó } 2$, tiempo parcial o completo y $Q82 = 1 \text{ ó } 2 \text{ ó } 3 \text{ ó } 4$, tipo de escuela

- b) Si Q80 = 2 no va a la escuela
 entonces Q81 = blanco
 y Q82 = blanco

Hay también entradas para otras preguntas que pueden estar relacionadas con Q80 y 82 y éstas son: Q14 y Q30: Va a alguna escuela como una razón para trabajar menos de 30 horas, Q58, Va a la escuela como contestación sobre lo que estaba haciendo en ese momento antes de buscar tra

bajo e ir a la escuela como una razón en Q64 por no haber trabajado la semana pasada. En todos estos casos el código es 3.

Como un ejemplo de un conjunto de condiciones específicas**, supongamos que no se dispone de la contestación a Q80, o es inconsistente con las demás. La siguiente tabla de decisión resume los pasos a seguir para imputar el valor apropiado para Q80.

	Regla de imputación						
	1	2	3	4	5	6	7
Q81 = 1 ó 2	S	N	N	N	N	N	N
Q82 = 1, 1, 3, ó 4	S	N	N	N	N	N	N
Q14, 36, 58 ó 64 = 3		S	N	N	N	N	N
Edad = 15 ó 16 años			S	N	N	N	N
Mes de la encuesta=julio o agosto*			S	N	N	N	N
Q80 en el mes previo = 1				S	N	N	
= 2					S	N	
= 1 ó 2						S	
Entonces Q80 en el mes actual	1	1	2	1	1	2	Fichero caliente

Aquí el fichero caliente significa búsqueda dentro de la misma unidad primaria, sendero o grupo de edad y sexo como antes se indicó.

Cada columna representa una sucesión separada de los pasos a seguir para llegar al valor a imputar a la pregunta de si una persona va o no a la escuela. S = Sf y N = No en el cuerpo principal de la tabla de decisión que se corresponden con las condiciones enunciadas.

Presentamos ahora la discusión detallada de la regla de imputación.

1. El valor imputado se basa en la lógica interna de cuestiones interrelacionadas. Así si Q81 y Q82 tienen respuestas válidas, se asigna a Q80 un valor consistente con ellos, por ejemplo si Q81 = 1 ó 2 y Q82 = 1, 2, 3, ó 4 entonces se impone 1 a Q80, o sea, *yendo a la escuela*.

2. Si no hay información disponible sobre Q81 y Q82, entonces se busca información relevante en algún otro lugar del cuestionario. Las otras cuestiones se preguntan dependiendo de cómo haya sido la secuencia particular seguida en esa encuesta. Si se ha indicado previamente que iba a la escuela entonces se codifica Q80 como tal por ejemplo si Q81 ≠ 1 ó 2 y Q82 ≠ 1, 2, 3, 4 pero Q14, 36, 58, 64 = 3, entonces se asigna 1 a Q80 o sea *yendo a la escuela*.

3.-4. Si no hay información disponible directamente seleccionada, entonces se usa información indirectamente relacionada con la pregunta que ha tenido problemas en la fase de depuración previa: tal es el caso del mes de la encuesta y la edad. Así, si una persona tiene 15 ó 16 años, lo probable es que durante los meses de agosto y septiembre no haya ido a la escuela y Q80 se codifica como 2, *no yendo a la escuela*. Mientras que los demás meses

*Para personas que no tengan 15 ó 16 años no importa que el mes de la encuesta sea Julio o Agosto para usar las reglas de imputación 5, 6 y 7.

**Excluye otras condiciones por ejemplo, que Q81 no sea 1 ó 2 pero Q82 sea 1, 2, 3, ó 4 o viceversa.

se ha podido codificar como 1, o sea, *yendo a la escuela.*

Estas imputaciones son ejemplos de casos raros por lo que se justifica la asignación del valor más probable aunque quizás prevé mejor teóricamente el fichero caliente si bien posiblemente más caro.

5.-6. Si no hay información disponible de los datos

del mes actual y la persona no tiene 15 ó 16 años, el siguiente recurso son los datos del mes previo. Siempre que se disponga de la respuesta el mes previo se transfiere el cuestionario del mes en curso.

El uso de los datos del mes pasado se justifica en base a que la correlación mes a mes de las características de ciertas estimaciones es muy alta.

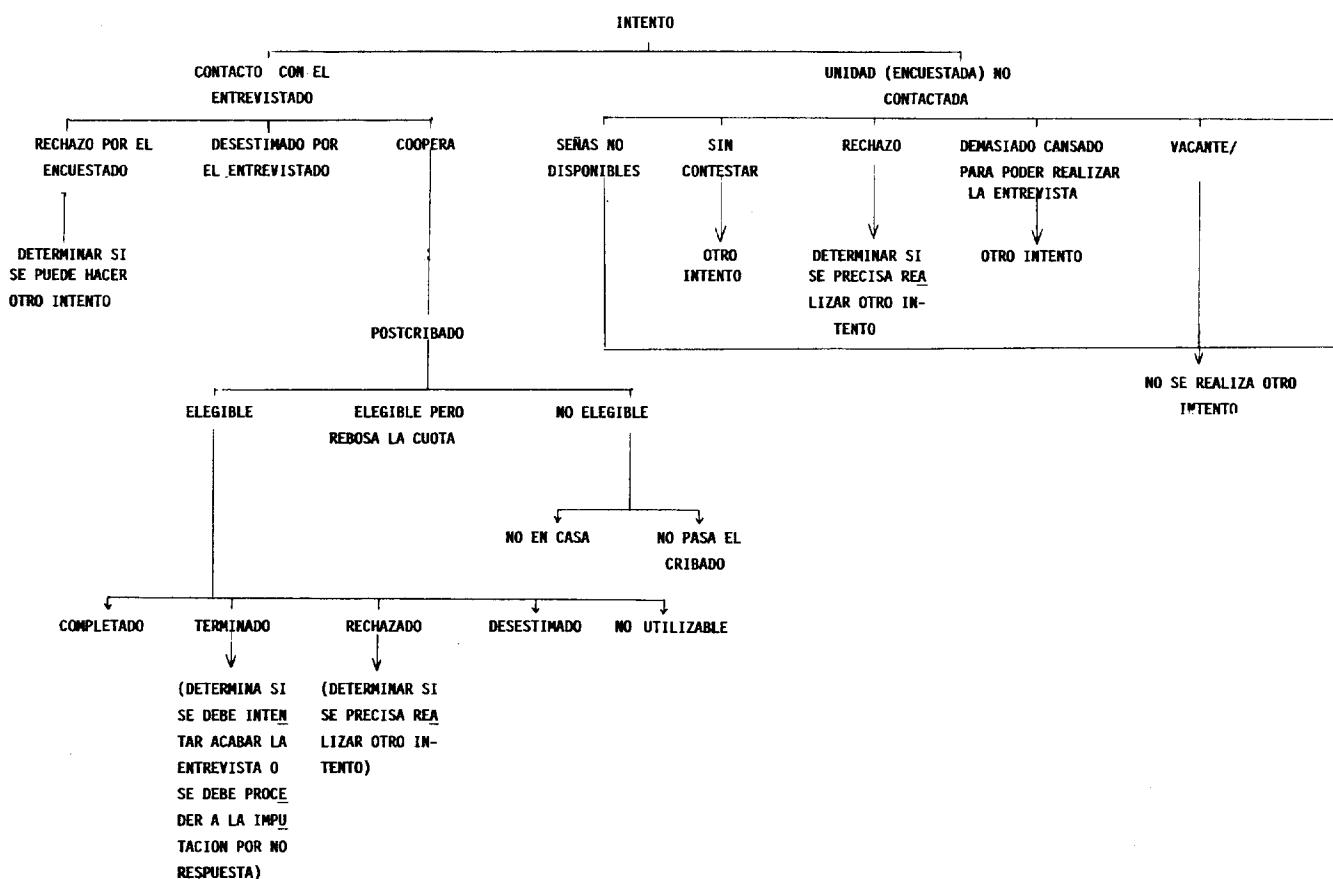
XI. INFORMANDO SOBRE LA NO RESPUESTA

Para controlar la no respuesta en el tiempo y para casos particulares es necesario tener un registro sistemático y continuo de respuesta/no respuesta.

Cuando se recogen datos de encuestas hay muchas situaciones que determinan la respuesta o la falta de ésta. Cada situación tiene una razón detrás de su respuesta/no

respuesta. Algunas unidades muestrales pueden ser rechazadas, otras vacantes y a otras no se les ha contactado. Así el examinar todos estos casos cuidadosamente, ayuda a tener para cada encuesta una información sobre la respuesta como la que se sugiere en la Tabla 13 para encuestadas personales y telefónicas.

TABLA 13
RESPUESTA EN ENCUESTAS PERSONALES Y TELEFÓNICAS



Basándose en la Tabla 13, se pueden construir otras tablas de contabilización cuyo propósito es mostrar cualquier reducción en el número de unidades de las que se obtienen respuestas en comparación con las unidades de

la muestra. Las unidades se clasifican en categorías por exclusión, inelegibilidad y diversos tipos de respuesta/no respuesta. Por ejemplo, para una encuesta particular una tabla de contabilización puede aparecer así:

TABLA 14
TABLA DE CONTABILIZACION

Nombre de la Encuesta	:	Encuesta sobre Renta
Poolación de la encuesta	:	
Población objetivo	:	
Forma de coleccionar datos	:	Entrevista personal
Tamaño muestral seleccionado	:	
Número de unidades contactadas	:	
Número de unidades elegibles	:	
Número de unidades vacantes	:	
Número de unidades nadie en casa	:	
Número de rechazos	:	
Rechazos por rechazos	:	
Número de ausentes temporales	:	
Otros tipos de no-respuesta	:	

Las tablas de contabilización se pueden clasificar por estratos, áreas de entrevistas, entrevistadores y así facilitar la dirección de la encuesta, lo que incluye el tomar decisiones sobre si se necesita apoyo adicional en ciertas áreas o reasignación de unidades a otros entrevistadores. Además las tablas de contabilización deben servir a dos propósitos fundamentales.

Primero, incluyen los ingredientes a partir de los cuales se pueden calcular las tasas de no-respuesta. El segundo objetivo es permitir mejorar la dirección de la encuesta. Por ejemplo, si se mantienen actualizadas las tablas, entonces la dirección de la encuesta podrá decidir quién y dónde se requieren esfuerzos especiales.