

6.5

A CASE STUDY ILLUSTRATING THE OBJECTIVES AND PERSPECTIVES OF FUSION TECHNIQUES

SUMMARY

Data collected through media audience surveys are used more and more as data banks for media planning. Media planners wish to be able to use data from various sources and of different kinds simultaneously, eg media audience data on the one hand, consumer panels or surveys data on the other hand. Fusions methods are part of a range of techniques which tend to solve problems of missing data in statistical files. The French programme called 'media-markets' (MM) provides a good case for applying these techniques. It suggests further developments at least in the French context.

Compared with initial and theoretical sampling plans, real field samples shows some shortfalls, eg not enough old people, not enough people living alone, not enough higher income level people, etc.

In the case of a postal survey, even if the response rate is high (eg 60%), the question is to know whether non-respondents would have answered in the same way as respondents.

A self-administered questionnaire accepted by people already interviewed face-to-face at home never obtains 100% completion. When the questionnaire is too long, you can divide the sample into sub-samples with different groups of questions for each sample; but you need the whole to be computerised as if everyone had given an answer to each question.

These different situations imply *problems of missing data*.

According to the objectives, various solutions can be provided for these problems. For correcting a bias of a statistical estimate, it is often possible to use statistical procedures - if needed, after having collected

some additional information - without changing file records. In other cases, it is possible to use the records of the survey data, directly after having corrected them by adding something replacing missing data. Techniques and methods considered below belong to this second type of procedure.

CONCEPTS AND WORDS

Before describing a specific case, we intend to discuss the need for clarifying concepts and words in relation to the problems encountered.

As far as problems or situations are concerned, Alexis Benguigui and Gilles Santini (1984) suggest the following list:

- the need for giving a figure in a 'nil' cell of a statistical table, when it is certain that the result is not nil;
- the need for eliminating 'non-response';
- the need for correct use of data banks with missing data (each statistical unit has not got the whole information required in an ideal situation);
- the need for estimating a local relationship between some variables, on the basis of non-exhaustive information.

Within a more limited framework, Lucien Boucharenc (1981; 1984) suggests that one distinguishes two broad categories of problems and procedures: the word *injection* should be used for situations where one wants to add information *within a given file*, and the word *fusion* when one wants to build a complete file by *putting together two files* which do not contain the same information.

This distinction has been adopted for several years, at least in France and Germany.

Other concepts can be defined. So the words *data connecting* (French: *rapprochement de données*), less precise than *injection* or *fusion*, could be used to cover a rather broader category of problems.

Sample weighting (French: *redressement d'échantillon*) is a very frequent procedure, very familiar to sampling and survey people, which consists of correcting some biases in statistical estimates due to an imperfect structure of the sample (with generally too few numbers for certain sub-groups of population, and sometimes too many for some others). This weighting can use various procedures.

Attributing information (French: *attribution*) is a basic operation which may be found in *injection* or *fusion* procedures. This word evokes different types of *methods* concerning *operations* or *procedures*, themselves used for solving certain categories of *problems*.

These methods can use statistical tools already known and developed in other fields, such as:

- concepts and methods of statistical estimation, and more broadly of econometrics, using weighting procedures, explanatory models with parameters, etc
- concepts and methods of data analysis: segmentation and cluster analysis, factor analysis; similarity, vicinity, distance analysis, etc.

Operations and procedures discussed in this paper have themselves led to the development of new concepts and the creation of new and specific words: in addition to *injection*, *fusion* and *attribution* already cited, other words are: *recipient* and *donor*; *alliance*, *marriage*, *liaison*, *crown* (around a type or group). These words can be found in the text below; they have been

originated by the authors cited in the bibliographical appendix.

Concepts and words concerning the types of variables considered are also used, more precisely to describe the *role* they play in the operations and procedures under review. Thus, the already cited authors considered *key* variables, *connecting* variables (French: *variables relais*), *common* versus *specific* variables, *direct* versus *composite* variables, etc.

As we hope the set of techniques cited in this paper will lead to further developments, we intend to stress from the beginning a strong need for setting up at the same time, statistical and computerised tools - which is for the most part achieved in western Europe now - and a theoretical framework which should be consistent, generally accepted and used at the international level.

THE CASE OF 'MEDIA MARKET' STUDIES IN FRANCE

The most recent and most important applied case of fusion techniques in France is the so-called 'media-market' (MM) CESP 1984 programme. Without relating the whole story of fusion techniques in France, it should be noted that an earlier important experiment had been carried out in 1982.

The main characteristics of this 1982 experiment are as follows:

- Three waves of a press and cinema audience survey were carried out according to usual CESP methods (a national representative sample of about 5,000 people for each wave);
- at the end of each interview of waves two (18 January to 9 March 1982) and three (15 April to 22 May 1982), (face-to-face home interviews) there was a self-administered questionnaire about consumption habits of 274 products and 3,583 brands of these products;

- the response rate of the MM questionnaire amounted to 72% (referred to as press-media interviews)*, an injection of information on consumption of products and brands was made for the 28% non-respondents;

- lastly, and to take account of the first press-cinema wave in media planning calculations, fusion of second and third wave file with the first wave file, information on products and brands were injected for each respondent on this first wave.

This experiment gave satisfactory results and it was decided to renew and enlarge it in 1984.

Figure 1 shows the general structure of the CESP 1984 programme:

- three waves for press and cinema audience survey, and three waves for radio and television audience survey;

- a sample of 1,000 people belonging to a high level of income and occupational status ('business and executives') is added to the second press wave.

- people interviewed for the second wave of the press-cinema audience survey (including the business and executives additional sample) together with people interviewed for the second wave of the radio-television audience survey, were asked to complete a self-administered questionnaire for 280 products and more than 3,000 brands of these products; two slightly different forms of the questionnaire were provided for men and women respectively.

Figure 2 shows the size of the various samples and sub-samples of computerised questionnaires more precisely.

* The real sample for press-cinema was 10,172 interviews. The MM questionnaire obtained 7,350 answers (ie 72.3% among which 7,311 were computerised (ie 71.9%).

FIGURE 1

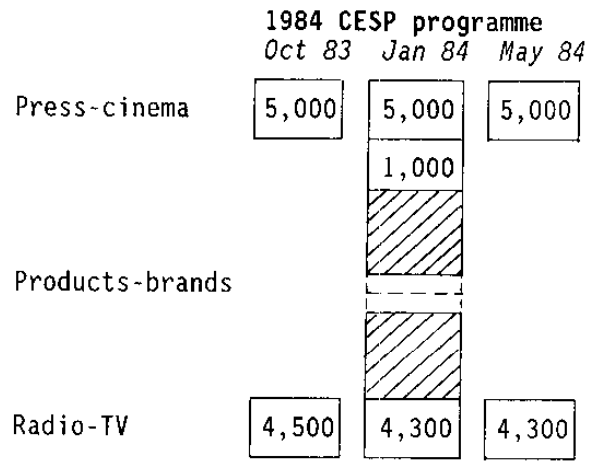


FIGURE 2

	Press	Radio TV	Press + RTV
V ₁	5,118	4,467	9,585
V ₂	6,250	4,265	10,515
(markets)	(4,956)	(3,433)	(8,389)
V ₃	5,100	4,430	9,530
Total	16,468	13,162	29,630

The response rate to the self-administered questionnaire has grown from 71.9% in 1982 to 79.7% in 1984 (on the basis of 100 collected press-cinema or radio-television questionnaires). This improvement has been gained, due to more experienced interviewers and with the help (as in 1982) of a little gift (approximate value \$5) as an incentive to answer.

Results of the MM study are planned on the basis of the whole file. After fusion (16,415 press interviews plus 13,162 radio-television interviews, ie 29,577 people as a whole) injections and fusions are carried out so that

complete information on the following is available for all these 29,577 people.

- *socio-economic variables*: directly available from the 29,577 interviews;
- *press audience variables*: available from press interviews, directly for 16,415 of the press sample, estimated through fusion techniques for 13,162 people in the radio-television survey;
- *radio-television audience variables*: available from radio-television interviews, directly for 13,162 of the radio-television sample, estimated through fusion techniques for 16,415 people in the press survey;
- *market variables*: available from self-administered specific questionnaire after the second waves of both press and radio-television surveys. These variables, available for 79.7% interviewed people in waves two have been estimated, either through injection (other interviews of both second waves) or through fusion (interviews of waves one and three) for other interviews.

Finally, 12 operations (two injections and 10 fusions, six vice-versa) have been carried out, according to the scheme shown in Figure 3.

These 12 operations have been carried out step by step, each step being thoroughly scrutinised and discussed by an ad hoc working group.

Specific weighting of each study have been kept up.

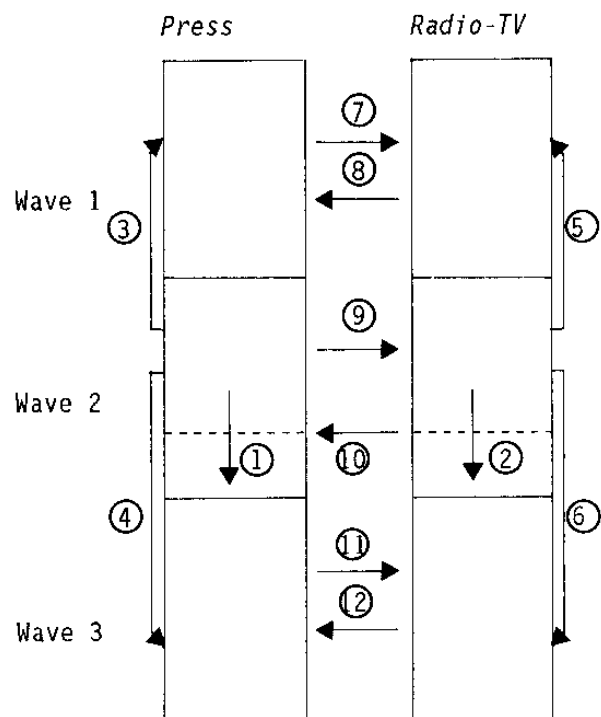
The process adopted to finalise the different phases of injections and fusions can be divided into four phases:

(1) Setting up of a factor system of reference on the basis of a factor analysis taking account of the whole samples concerned (both donors and recipients).

(2) Creating cells (through segmentation of the universe) within the factor system of reference in which the fusion operations will take place.

(3) 'Marriage' algorithm of people, and optimisation procedure (a 'fine' being given so that one given 'donor' should not be called too many times).

FIGURE 3



(1) and (2): 'injection' of missing data/products brands to non-respondents

(3) and (4): Simulating missing data/products brands for people interviewed in press waves 1 and 3 (inter-wave 'fusions')

(5) and (6): Simulating missing data/products brands for people interviewed in radio-television waves 1 and 3 (inter-wave 'fusions')

(7) to (12): Simulating missing data/media for each wave and vice-versa (inter-study 'fusions').

(4) Checking up of results, step by step, through comparison, control and validation of figures.

Appendix 1 describes the progress of these four phases in more detail.

EVALUATING RESULTS

The end of Appendix 1 gives a list of checks and comparisons which have been provided after injection and fusion operations. Other additional analyses have been completed. The limited frame of the written paper does not allow the elements of the various checks and controls to be given but significant examples were shown at the Salzburg meeting.

Experience and *a priori*, then a *posteriori* reflection, lead to the suggestion of several validity criteria and some type of constraints which seem appropriate.

(1) Sub-groups can be defined between which no exchange 'donor-recipient' can be authorised (eg men/women); these sub-groups constitute the cells defined above.

(2) A given 'donor' must not be called too often, in order to avoid cluster effects among recipients.

(3) The total of distances between donors and recipients must be globally minimum.

(4) Injection and fusion operations must not change the most important results of the study significantly (eg according to a chi squared test) whatever the file considered.

It is a matter for:

- marginal distributions of all variables
- some cross-tabulations
- some relationships between variables
- some global ratios.

In addition to operations aimed at constituting the final file, and in

order to check the stability of structures and results, it is possible to provide redundant operations, eg simulating the effect of injection and/or fusion procedures for variables which, in fact, are already known and available in the recipient file.

(5) Lastly, when possible (this is the most difficult criterion but perhaps the most useful) try to estimate correctly existing but unknown relationships between variables. Appendix 1 provides some results of evaluation procedures.

FURTHER DEVELOPMENTS AND PERSPECTIVES

At least in the case of France, we forecast a further development for fusion techniques as a consequence of two contrary trends, and as a solution to this contradiction. A first trend leads increasingly to separate audience survey vehicles for different media. Press and cinema are no longer in the same survey as radio and television. It is possible that radio and television will separate, at least partly (due to audience automatic metering systems which are now running for television and do not exist for radio). It is also possible that dailies on the one hand, and magazines on the other hand, will tend to have their own separate audience surveys.

These trends are relevant at the national level. Conversely, it is possible that audience surveys at regional and local level, which are now developing, will remain multimedia, for financial reasons among others, due to the adequate size of samples required at this level.

The second trend, due to extension and development in the use of data banks, is the wish to have at immediate disposal, through a unique computerised base, a maximum amount of data available for media planning.

We illustrate in Figure 4 various trends which lead to the development of fusion techniques in France.

FIGURE 4

Trends	Fusion needs
Survey vehicles tend to become different for different media	Inter-media fusion
Some surveys are not carried out annually	Up-dating of audience media data banks
Development of audience measurement using panels techniques	- Processing data as if the sample were constant, which is never the case - Permanent up-dating of a fixed sample according to a changing universe
Additional samples with partly specific questions and partly questions common to the main sample	Fusion of additional samples with main sample

Concerning the so-called media market field, there is no evidence that CESP 82 and 84 programme have to be continued. It could be less expensive, perhaps of comparable scientific validity and probably more useful and interesting as far as results are concerned, to study and carry out bridges through fusion techniques between audience survey files on the one hand and consumer surveys and panels files on the other hand.

Finally, we point out that it would be valuable if fusion languages and techniques are co-ordinated and accepted at international level. It is already the case for an important part of what is currently done in this field in western Europe; it would be good to enlarge the area of this co-ordination.

APPENDIX 1

Progress and evaluation of injection and fusion operations (CESP media-markets 1984 study)

Phase 1 determines a factorial system of reference as a basis for fusion.

It was decided to work on a factor system (correspondence analysis) to eliminate three frequent drawbacks in data analysis.

The redundancy between variables: the presence of dependent variables

The size effect: the scale effects due to the fact that some quality is present in a large or in a very small number of people

The outliers: to which most statistical methods are very susceptible.

In the factor system, the *similarity* between people is measured by the euclidian distance. It has been proved that this distance gets rid of the *random fluctuations* as well as of a part of the influence of the *way of coding* data.

The factor system can be built either from the recipient group alone or from the junction of the two groups which are to be fused together.

In the present case, systems of reference have been built from all the people interviewed in the second wave of each media study, separately for men and women.

Two sets of variables have therefore been considered:

- (a) socio-demographic variables
- (b) media audience patterns.

It has been considered necessary to select the same type of media variables for both studies.

National daily newspapers have been considered through boolean variables such as:

1 for regular or quite-regular readers
0 for others

and for magazines:

11 for regular or quite-regular readers
01 for other readers
00 for non-readers

302 variables have been so selected for referential system 'men' and 304 for 'women', with 55 socio-demographic variables, in both cases.

Eleven factor axes have been selected explaining 21.08% of the variance for men and 21.27% for women.

As far as the radio-television referential system is concerned, the same 55 socio-demographic variables as for press have been considered, and in addition:

- number of TV sets (0, 1, 2 or more)
- telecommand equipment
- number of radio sets (0, 1-3, 4 or more)
- FM disposal
- stereo disposal
- radio watcher disposal
- car radio disposal

This referential system has been developed for the Monday to Friday sample of the second wave; people belonging to the Saturday and Sunday sub-samples have been projected in a passive way within the Monday to Friday referential system.

The total number of active variables considered for radio-television was a little less than for press. Here too, 11 factor axes have been selected.

As far as inter-study fusions are concerned, the same system has been adopted. These fusions have been conducted first from radio-television to press and then from press to radio-television, and this for each of the

three waves, through the following variables:

- socio demographic variables, common to both studies
- general radio-television audience data, common to both studies
- market classifications (through cluster analysis) built with the two samples of respondents (second wave press and second wave radio-television) and thereafter transferred to people in waves 1 and 3.

Inter-study fusions were provided for each wave and it has been useful to consider the market data available in the second waves as additional variables for fusion.

Two classifications (men and women) have been developed from 211 consumption variables for men and 221 for women; the working group decided to consider 20 types.

Furthermore, two additional variables have been considered in the progress of inter-study fusion, at the level of the cells used as frames for the marriages:

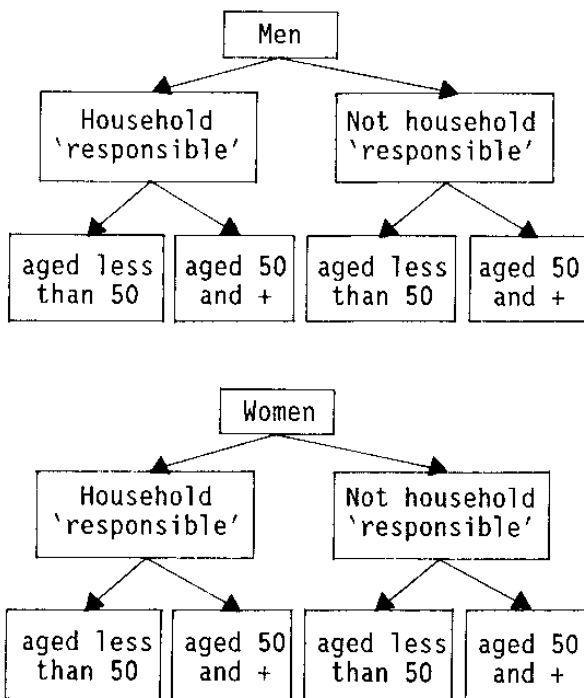
- active/non-active for women
- occupation of the head of household in two categories: self-employed, business executives, professionals, medium-level staff/others.

Phase 2 consists of creating cells within which the fusions will take place.

There are eight cells.

Six additional constraints have been considered in the case of radio and TV audience data, namely fusions allowed only:

- in the same geographical territory, France being divided into four broad regions
- in the same field of questions:
 - Monday to Friday audience
 - Saturday audience
 - Sunday audience.



After the factor analysis itself, the study has proceeded as follows:

calculation of the factors values (combination of the initial variables);
checking of the factorial system;
projection of the people from the two groups, recipients and donors.

Phase 3 consists of matching ('marrying') people.

As far as injections are concerned, each non-respondent to the market questionnaire has to find one respondent in the same second wave.

As far as inter-study fusions are considered (market information on waves one and three), donors used are only those who have the self-administered questionnaire; 'injected' people of these second waves do not contribute to these fusions.

The factor referential systems are considered as frameworks for deciding marriages. Therefore, everyone is first located within the referential system. Then, the objective is to find, for

each recipient, the most proximate donor in his vicinity.

NB 1: In order to allow the use of the algorithm of rapid search of the neighbour (according to the J Friedman technique), the data have been rearranged in a form called 'the K dimensional tree'.

NB 2: The people are married according to their proximity D calculated in the factorial system.

If a person is already married with Δ other people, this distance is distended (penalisation):

$$D^* = 1 - (1 - D)^\Delta$$

The six types of possible marriages are as follows:

- (1) Marriage at first sight (most proximate for each other)
- (2) Marriage with a childhood friend
- (3) Marriage by adultery
- (4) Marriage through attentiveness
- (5) marriage of convenience (through optimisation: after considering the 10 most proximate people)
- (6) Marriage of the invincible (confirmed bachelors, etc).

The chart shows definitions and procedures for cases 1 to 5.

Cases of marriages 1 to 4 are tried first. But, every person does not have a partner through these cases. For each unmarried person we then determine the 10 most proximate people and the cost of each possible marriage. (Cost = sum of the penalised distances.) The aim is to match one person and only one to each point of the system. The method used is derived from the Hungarian method.

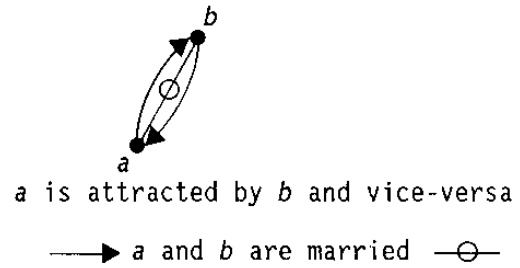
Finally, there remains a small percentage of invincibles (confirmed bachelors) who are married to their nearest neighbour. The links are then checked. The marriages involving too many people are broken and an improvement is attempted.

Third Phase

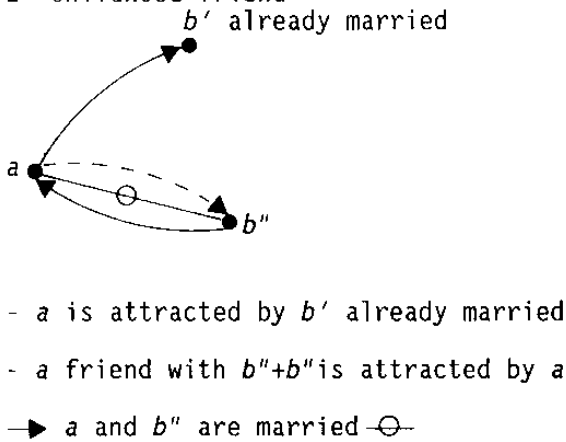
Marriages

- at first sight
- with a childhood friend
- by adultery
- through attentiveness
- indecision

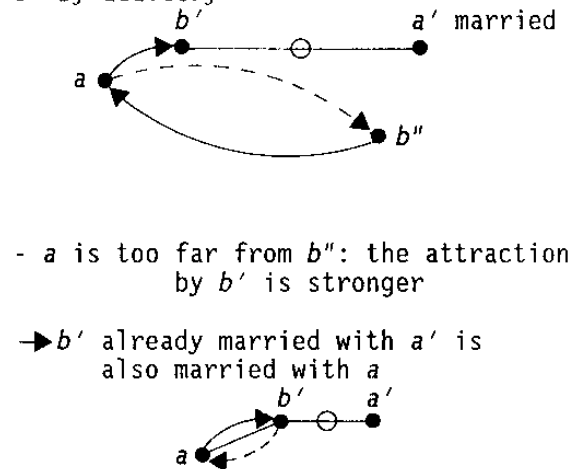
1 At first sight



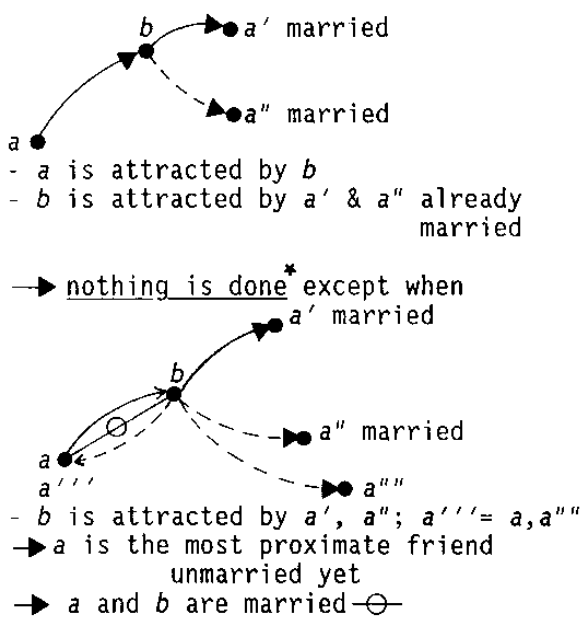
2 Childhood friend
 b' already married



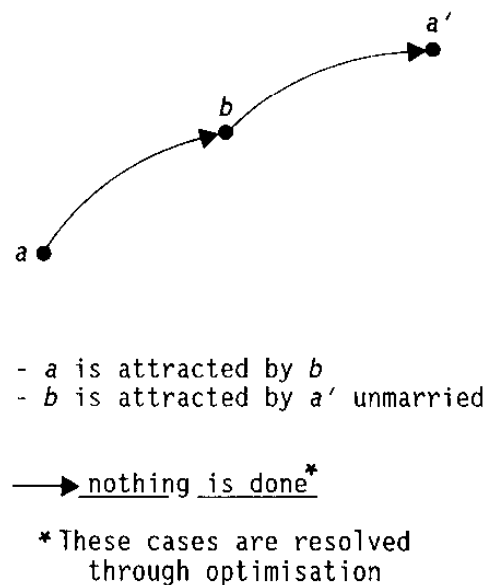
3 By adultery
 a' married



4 Through attentiveness



5 By convenience



Phase 4 consists first of examining the different types of marriages of inter-wave fusions and inter-study fusions.

We have established:

six tables for inter-wave fusions
(three waves x two studies)

six tables for inter-study fusions
(three waves x two studies)

The set of people 'to be married' is:

- non-respondents to the market questionnaire for inter-wave fusions
- non-respondents to the media questionnaires for inter-study fusions.

These tables give horizontally, the eight cells fixed at phase 2 within each of which the marriages have been concluded; and vertically, the various types of marriages, the number of people in each cell, and the average number of calls of a donor in each cell (average number of links).

Tables 1-5 illustrate that:

- The two most frequent types of marriages are type 1 (more proximate for each other), and type 5 (optimisation through the 'Hungarian method'). The number (or proportion) of marriages through optimisation (type 5) can be considered as an indicator of the quality of fusion (a small number being better than an important one).
- The average distances are not very different according to cells, except for small cells (containing few people).

As a whole, the methods considered have worked correctly.

Compared with intra-wave 2 fusions, the inter-wave fusions show an increase of the number of marriages through optimisation.

Press fusion:

- wave 1 : 43% of the marriages
- wave 3 : 41.3% of the marriages

Radio-TV fusion:

- wave 1 : 50.6% of the marriages
- wave 3 : 50.0% of the marriages

The origin of this increase lies in the fusions cells, which are larger. Therefore there is a greater rivalry between the 'donors' and the 'recipients'. This rivalry, linked to the proximity of the people among them, leads to a decrease of the number of strictly vice-versa neighbours.

Concerning the inter-study fusions, the number of marriages through optimisation varies from 43.6% to 50.4%, despite the constitution of clusters of people.

Besides that, chi squared analyses have been provided for each of three sub-groups: donors, recipients and donors joined to recipients. This is at the level of socio-demographic characteristics and of media audiences, for each wave of each study.

The following checks have also been considered:

(a) Breakdown of answers to 'market' questions after injection and fusion, for each of three files: press file, radio-television and press + radio + television file.

As a whole, no significant difference for market results have been found between the two studies (see Table 6).

(b) Comparison of press and radio-television audiences (reading within the relevant period for press; listening or watching habits for radio-television): directly from each specific study (the three waves being added for each study) on the one hand; after re-estimating through fusions, on the other hand. As a whole, a good consistency has been obtained between observed and simulated results (see Table 7 for press audiences).

Finally, after completion of these operations, CESP has a complete MM file for 29,577 people, as a base for all published results.

(c) Comparisons of press audiences cross-tabulated with radio-television audiences, before and after fusion, through chi squared analyses (see Table 8). As a whole a good consistency has been observed.

REFERENCES

Benguigui, Alexis and Santini, Gilles (1984) 'Rapprochement de données ou fusions d'enquêtes' Communication to IREP seminar

Bergonier, Henri; Boucharenc, Lucien and Irrmann, Philippe (1967) 'Fusion d'enquêtes. Application a l'évaluation de projets de plans media' *Revue Française du Marketing*. 22, 1

Boucharenc, Lucien (16 November 1981) 'Les techniques de fusion appliques aux études CESP' Memorandum

Boucharenc, Lucien (29 March 1984) 'Rapprochement de données ou fusions d'enquêtes' Communication to IREP seminar

CESP (1982) 'Etude media-marchés' First volume: general results

CESP (1984) 'Etude media-marchés' Tome 1: methodologie et depouillement du questionnaire

Santini, Gilles (29 March 1984) 'La méthode de fusion sur référentiel factoriel' IREP

Wendt, Friedrich (May 1976) 'Beschreibung eine fusion' Schriftenreihe band 21 Grüner and Jahr AG & Co, Hamburg.

This paper has been prepared with the help of Hugues Chavenon and Muriel Libs/CESP. The methodology is mainly due to Gilles Santini, IMS France.

TABLE 1
Press injection Wave 2

Name of cells		Type of marriages						Total	Average distance	Av. no. of links
		1	2	3	4	5	6			
HNJA	NBR	144	8	0	27	43	3	225	1,295	1.08
	%	64.0	3.6	0.1	12.1	19.2	1.4			
HNVA	NBR	91	7	1	15	33	1	148	1,273	1.11
	%	61.5	4.8	0.7	10.2	22.3	0.7			
HRJA	NBR	61	1	0	11	31	0	104	2,002	1.12
	%	58.7	1.0	0.1	10.6	29.9	0.1			
HRVA	NBR	71	14	3	6	44	10	148	1,571	1.20
	%	48.0	9.5	2.1	4.1	29.8	6.8			
FNJA	NBR	26	1	0	4	19	1	51	2,234	1.16
	%	51.0	2.0	0.1	7.9	37.3	2.0			
FNVA	NBR	9	3	0	0	13	1	26	2,305	1.37
	%	34.7	11.6	0.1	0.1	50.1	3.9			
FRJA	NBR	181	12	0	24	53	1	271	907	1.08
	%	66.8	4.5	0.1	8.9	19.6	0.4			
FRVA	NBR	161	22	14	19	96	8	320	978	1.26
	%	50.4	6.9	4.4	6.0	30.1	2.5			
Total	NBR	744	68	18	106	332	25	1,293		1.15
cells	%	57.5	5.3	1.4	8.2	25.7	1.9	100		

TABLE 2
Radio injection wave 2

Name of cells		Type of marriages						Total	Average distance	Av. no. of links
		1	2	3	4	5	6			
HNJC	NBR	91	3	1	14	19	0	128	1,281	1.10
	%	71.1	2.4	0.8	11.0	14.9	0.1			
HNVC	NBR	59	8	0	14	34	2	117	1,999	1.12
	%	50.5	6.9	0.1	12.0	29.1	1.8			
HRJC	NBR	42	2	0	5	16	2	67	1,617	1.05
	%	62.7	3.0	0.1	7.5	23.9	3.0			
HRVC	NBR	53	1	1	6	20	3	84	1,779	1.16
	%	63.1	1.2	1.2	7.2	23.9	3.6			
FNJC	NBR	13	2	1	4	10	4	34	3,470	1.26
	%	38.3	5.9	3.0	11.8	29.5	11.8			
FNVC	NBR	14	1	0	2	5	0	22	1,690	1.23
	%	63.7	4.6	0.1	9.1	22.8	0.1			
FRJC	NBR	104	5	3	11	21	0	144	1,264	1.10
	%	72.3	3.5	2.1	7.7	14.6	0.1			
FRVC	NBR	96	14	6	21	96	8	241	1,664	1.27
	%	39.9	5.9	2.5	8.8	39.9	3.4			
Total	NBR	472	36	12	77	221	19	837		1.16
cells	%	56.4	4.3	1.4	9.2	26.4	2.3	100		

TABLE 3
Inter-waves fusion

<i>Total cells</i>		1	2	<i>Type of marriage</i>				Total	<i>Av. no. of links</i>
				3	4	5	6		
Press									
Wave 1	Number	1,780	252	249	233	2,198	406	5,118	1.49
	%	34.8	4.9	4.9	4.5	42.9	8.0	100	
Wave 3	Number	1,785	259	269	272	2,083	379	5,047	1.51
	%	35.4	5.1	5.3	5.4	41.3	7.5	100	
Radio									
Wave 1	Number	1,231	183	277	197	2,261	318	4,467	1.64
	%	27.6	4.1	6.2	4.4	50.6	7.1	100	
Wave 3	Number	1,255	186	307	181	2,214	287	4,430	1.65
	%	28.3	4.2	6.9	4.1	50.0	6.5	100	

TABLE 4
Inter-studies fusion

<i>Total cells</i>		1	2	<i>Type of marriage</i>				Total	<i>Av. no. of links</i>	<i>Not used donors</i>
				3	4	5	6			
Press										
Wave 1	Number	1,538	226	334	233	2,490	277	5,098	1.45	1,243
	%	30.4	4.4	6.5	4.6	48.7	5.4	100		
Wave 2	Number	1,568	203	454	239	3,114	601	6,179	1.87	830
	%	25.4	3.3	7.3	3.9	50.4	9.7	100		
Wave 3	Number	1,532	224	316	222	2,422	331	5,047	1.58	1,238
	%	30.3	4.4	6.3	4.4	48.0	6.6	100		

TABLE 5
Inter-studies fusion

<i>Total cells</i>		1	2	<i>Type of marriage</i>				Total	<i>Av. no. of links</i>	<i>Not used donors</i>
				3	4	5	6			
Radio-TV										
Wave 1	Number	1,558	207	216	258	2,035	193	4,467	1.42	1,965
	%	34.9	4.6	4.8	5.8	45.6	4.3	100		
Wave 2	Number	1,568	206	240	237	1,860	154	4,265	1.40	3,197
	%	36.8	4.8	5.6	5.6	43.6	3.6	100		
Wave 3	Number	1,532	194	244	251	2,028	181	4,430	1.48	2,045
	%	34.6	4.4	5.5	5.7	45.8	4.0	100		

TABLE 6

<i>Products or brands</i>	<i>Fre- quency</i>	<i>Press file</i>	<i>Radio TV file</i>	<i>Media markets file</i>
<i>Toilet products</i>				
Shampoo	(1)	91.6	92.6	92.1
	(2)	37.3	39.1	38.3
- Dop	(1)	24.3	22.9	23.6
Dop	(3)	14.0	13.3	13.6
Lipsticks	(1)	61.8	62.1	61.9
	(2)	23.0	23.6	23.3
- Yves Rocher	(1)	17.3	16.8	17.0
Yves Rocher	(3)	10.9	11.2	11.0
<i>Drinks/tobacco</i>				
Beers	(1)	50.4	48.8	49.6
- Kronenbourg	(1)	21.6	20.3	20.9
Kronenbourg	(3)	13.0	13.0	13.0
Champagnes	(1)	58.7	58.9	58.8
- Mercier	(1)	18.6	17.6	18.1
Mercier	(3)	12.2	12.2	12.2
Cigarettes	(1)	35.5	35.1	35.3
more than 10 per day		17.2	16.6	16.8
- Blonde with a filter	(3)	11.7	10.8	11.2
<i>Equipment</i>				
- Records	(1)	50.5	48.9	49.7
1 to 4	(1)	24.7	23.4	24.0
Varieties	(3)	33.2	31.1	32.2
- Racquet	*	16.9	17.6	17.2
- Watch (1000f)	*	10.6	8.4	9.5
- Hot-plate	*	22.1	22.5	22.3
<i>Leisure services</i>				
- Plane flights	(1)	11.9	11.2	11.6
- Books	(1)	57.5	56.2	56.9
of which 6+	(1)	30.7	29.6	30.2
- Credit cards	*	13.7	11.9	12.8
- Mail order	(1)	45.1	47.5	45.4
<i>Food</i>				
- Tea	(1)	54.7	55.0	55.0
Tea once a day		21.3	18.6	20.0

TABLE 6 continued

<i>Products or brands</i>	<i>Fre- quency</i>	<i>Press file</i>	<i>Radio TV file</i>	<i>Media markets file</i>
<i>Food</i>				
Lipton	(1)	31.9	29.0	30.5
- Rice	(1)	91.4	92.7	92.5
Rice once a week		41.4	42.0	41.6
Uncle Ben's	(1)	45.8	45.9	46.1
Uncle Ben's	(3)	31.5	33.2	32.5
<i>Frozen products</i>				
Findus	(1)	65.1	63.6	64.9
Findus	(1)	39.8	37.0	38.6
Findus	(3)	26.7	25.0	26.1
<i>Dog food</i>				
Pal	(1)	25.5	25.0	25.6
Pal	(1)	8.8	7.6	8.3
Pal	(3)	4.9	5.0	4.9
<i>Detergents</i>				
- Powder for dish-washer	(1)	25.2	24.5	25.0
Sun	(1)	19.0	18.2	18.7
Sun	(3)	16.8	15.7	16.4
<i>Cleaning material for window panes</i>				
Ajax vitres	(1)	83.7	86.8	85.9
Ajax vitres	(1)	58.3	62.4	60.8
Ajax vitres	(3)	48.1	51.7	50.3
<i>Equipment</i>				
- Electric coffee pot	*	64.6	65.3	65.6
Moulinex	*	34.4	34.8	34.9
- Electric frying pan	*	20.9	22.5	22.0
Seb	*	12.3	13.5	13.0
- Mixer	*	56.2	55.9	56.4
Moulinex	*	48.5	48.0	48.6
<i>(1) Last 12 months; (2) More than once a week; (3) The most frequent * Ownership</i>				

continued

TABLE 7

Comparison between
 Press 3 waves cumulated ←————→ Media-markets Press + Radio-TV fused

Weeklies - Concerning the 39 publications studied, the discrepancies between the data* of the press file and of the media-markets file are:

- 0 for 12 titles
- 1% for 11 titles
- 2% for 14 titles
- 3% for 2 titles

Monthlies - Concerning the 66 publications studied, the discrepancies between the data* of the press file and the media-markets file are:

- 0 for 24 titles
- 1% for 33 titles
- 2% for 8 titles
- 4% for 1 title

* AIR in % total adult population

TABLE 8

Comparison using the Chi² method between the press titles' penetration rates cross by the radio-TV listening before and after fusion

Tests on the last period reading

Radio : 1.1 crossing of the press audience data by the radio listening for 9 TP and 2 universes (women or men):
 → 117 titles x 9 TP radio = 1053 measures for each universe
 of which 10.6% are significant for the men sample
 8% are significant for the women sample

continued

TABLE 8 continued

1.2 crossing of the press audience data by the radio listening (4 stations + radio in general) during the whole day and for each universe:
 → 117 titles x 5 stations = 585 measures for each universe
 of which 5.8% are significant for the men sample
 4.1% are significant for the women sample

Television : 1.3 crossing of the press audience data by the television listening for 6 TP and 2 universes (women or men):
 → 117 titles x 6 TP = 702 measures
 of which 8% are significant for the men sample
 12.4% are significant for the women sample

1.4 crossing of the press audience data by the television listening (3 channels + television in general) for the whole day and for 2 universes (men or women):
 → 117 titles x 4 channels = 468 measures
 of which 11.3% are significant for the men sample
 13.7% are significant for the women sample

Tests on reading habits (regular readers)

Radio : 2.1 crossing of the number of regular readers by the radio listening of 9 TP for 2 universes (men or women):
 → 117 titles x 9 TP = 1053 measures

continued

TABLE 8 continued

of which	12.7% are significant for the men sample 6.4% are significant for the women sample
	2.2 crossing of the number of regular readers by the radio listening (4 stations + radio in general) for the whole day and for 2 universes (men or women):
—————>	117 titles x 5 stations = 585 measures
of which	7.5% are significant for the men sample 3.1% are significant for the women sample
<i>Television :</i>	2.3 crossing of the number of the number of regular readers by the television watching of 6 TP for 2 universes (men or women):
—————>	117 titles x 6 TP = 702 measures
of which	6.8% are significant for the men sample 9.2% are significant for the women sample
	2.4 crossing of the number of regular readers by the television watching (3 channels + TV in general) for the whole day and for 2 universes (men or women):
—————>	117 titles x 4 channels = 468 measures
of which	10% are significant for the men sample 17.5% are significant for the women sample.
