Rolf Speetzen
Axel Springer Verlag AG, West Germany

8.4 THE ART OF MODELS – ASCRIPTION IN GERMANY

INTRODUCTION

Germany is the country where ascription has been practised in numerous ways for a comparatively long period. Ascription should stand here for all mathematical approaches and models to align data from one source with data of a different source, as well as processing the empirical results into probabilities and the like.

Why the need to ascribe and align data from different sources? The reason can be seen in the increasing number and variety of media to be included in the multi-media analyses, breaking all the limits of the respondents to answer questions which are irrelevant to them. On the other hand we have the need of the data users to tap just one source for target group and intermedia information, to avoid arriving at different schedules by using different data bases.

While some countries tried to streamline the interview and stay single-source, the Germans followed a different route. It started when some German publishers tried to highlight individual markets, markets which were the stronghold of their titles. Unfortunately, the media data differed somewhat from the 'official' MA data. Where the empirical sciences failed, the models had to take over.

THE EARLY ATTEMPTS

The first model in the early 70s was called 'Anpassung'. That means 'adjustment' and is a very basic model, just used to match the individually established media information with the MA. The first attempts were based on a very limited number of criteria, such as sex and age.

The basic adjustment model was developed to a more complex technique in the course of time. The more recent models are based on the 'minimum distance technique'. For each demographically defined target group, the official MA requirements are precisely reconstructed, but, under the restriction of minimising variations necessary for the adjustment within the broadest audience, thus aiming at changing the empirically established information as little as necessary, at the same time distributing the media probabilities evenly and according to their share within the different readership segments of the broadest audience.

Another early model is called 'transformation'. This is again extremely simple. It multiplies the empirically established media user scores with a weight to arrive at the exact MA score. The disadvantage of transformation can be seen in the fact that only the AIR is ascribed to the MA, whereas all other criteria, such as broadest audience, cume values and demographics remain unchanged, that is unadjusted.

Primitive or simple as the early stages were, they showed us a way of managing the data inflation and satisfying the user's need. They showed us a solution through abandoning the single-source approach, through splitting the extremely lengthy interviews into separate units and then merging the individual data sets into a comprehensive data base, embracing all necessary market and media information. However, it was a long way to go from adjustment to the partnership model of the MA, until we used a new technique: fusion.

FUSION

I would not have been able to describe the process of the German fusion without the excellent report of the German MA written by Peter Beike and Friedrich Wendt. My description is therefore based on (or to stay in terms, aligned to) the official MA document.

Fusion in Germany does not mean the complete merger of two data sets into one comprehensive one. Fusion means, just one piece of information of one data set will be ascribed to another. Maybe a more precise name would be infusion or injection. Anyway, fusion seems to have hit the charts, so let us stay with it.

What data sets are being fused? In Germany we have a continuous TV Meter Panel, the GfK Panel, and the almost official Media Analysis, the MA, which is split into two parts: MA Electronic, some 19,000 interviews and MA Print, 18,000 different interviews.

The TV stations run, control and finance the meter panel. And they pushed the idea of taking the media currency, the TV viewing probabilities, from the panel and merging those data with the MA data, where TV viewing is established by day-part questioning, with much lower ratings.

So the Germans fuse first TV meter panel data of 7,000 panellists into the MA Electronic based on 19,000 ad hoc face-to-face interviews, then they fuse the enriched MA Electronic into MA Print based on 18,000 ad hoc face-to-face interviews to arrive at one data base for intermedia planning.

Fusion is structured into four steps:

(1) Preparation of the data, till they are ready for the transfer

(2) Creating homogeneous sub-groups

(3) The actual data transfer process

(4) Delivery and fusing of the transfer data.

Data preparation

The first step, data preparation, is a very crucial step. Here it has to be decided whether the individual data sets are generally suitable for matching. In the German case, as just said, data from a continuous meter panel were fused into data, based on ad hoc face-to-face interviews.

The universes are different, too. The MA universe is more comprehensive, representative of the adult German population. The GfK universe should be representative for TV households.

TV viewers according to GfK are people who live in households with at least one TV set. In the MA we find viewers who do not own a set, but view elsewhere. And in addition the MA reports non-viewers.

For data preparation the TV viewers had to be separated into terrestrial TV viewers with a limited choice of channels and cable TV viewers. Another preparatory step was the split of the viewing within the panel into two half years and the split of the MA Electronic into two parts, first two sweeps and last two sweeps (Table 1).

On the one hand this split should reflect the different viewing pattern of the seasons, on the other hand it should level off the data base sizes: GfK: terrestrial 5,500 panellists, MA: 9,000 in each half! Cable TV had to be considered in one lot, due to the limited sample size. Last but not least, we had to analyse men and women separately. So the MA fusion, is based on six runs, ie six fusions!

Table 1

Data bases: sample sizes

	Terrestrial TV viewers		Total population		Cable TV viewers
	Sweeps		Sweeps		Sweeps
	1+2	3+4	1+2	3+4	1 - 4
GfK					
Men	2,629	2,514			713
Women	2,880	2,761			744
Total	5,509	5,275			1,457
MA Electronic					
Men	3,998	3,909	4,068	3,961	888
Women	5,036	4,853	5,122	4,944	967
Total	9,034	8,762	9,190	8,905	1,855
MA Print					
Men			3,965	3,828	744
Women			4,671	4,649	850
Total			8,636	8,477	1,594

Source: MA 88 Documentation

A very crucial issue is the calculation of the viewing probabilities. The panel records actual viewing, the MA needs probabilities for media planning purposes. We therefore had to transform actual viewing into p-values before the fusion. The process is fairly easy. It is the standard procedure, creating an average profile for the reporting period by calculating an average daily rate for each individual.

The final step in data preparation is the choice of the common criteria. It has to be defined whether and which common criteria can be established in the two sets to be fused. Really common criteria means criteria which are established under comparable conditions with a comparable distribution.

Table 2 shows the number and the scope of the criteria. Obviously, there are a lot of criteria for the match. The scope runs from socio-demographics, household equipment, leisure time activities to media criteria.

Homogeneous sub-groups

Fusion means in simple words that for each recipient the most similar donor should be found. Staying in simple terms we just have to compare each recipient with each donor over all relevant criteria.

Thinking in real terms: 9,000 recipients and 5,500 donors mean some 50 million

Table 2

Available criteria

COMMON CRITERIA	Number of Criteria
Terrestrial homes:	
GfK/MA electronic	196
MA EI/MA print	278
Cable homes:	
GfK/MA electronic	204
MA EI/MA print	278
SCOPE OF CRITERIA	
Socio-demographic	93
Household equipment	109
Appliances	36
Additional criteria	44
Leisure time	48
Coordinates	12
TV	111
Radio	11
Dailies	64
Other media	18
Sweep	1

Source: MA 88 Documentation

comparisons over a lot of criteria. But this is just one reason not to do it this way, comparing individuals. Another and even more important reason is the fact that by matching individuals in some cases the less important criteria could dominate the really relevant criteria; for example, a match in demographics could dominate the media behaviour.

The MA fusion is based on segments by sex and broadcasting areas. Within these 24 segments (12 terrestrial broadcasting areas by sex) homogeneous sub-groups are constructed by cluster analyses. These sub-groups function as cells for matching donors and recipients on the condition that the structure of the clusters is similar on both sides, and the sizes of the clusters stand in a reasonable relation.

The process starts by clustering all donors, thus constructing archetypes which show maximum homogeneity within types and maximum discrepancy between the different clusters. Criteria for cluster construction are all common characteristics, common to donors and recipients.

The next step transfers the donor clusters into the recipient sample. This sounds easy, but in a number of cases the cluster structure cannot be matched, the distribution of the variables within the recipients being different from the distribution within the donors. Now the newly formed recipient types are transferred back to the donor side and then the donors are once more iteratively assigned to the new clusters thus resulting in maximum correspondence between the two samples.

Sometimes the assignment of cases to certain clusters is not unambiguous. To avoid an arbitrary decision for one or the other, the MA fusion uses aureoles, or halos. Aureoles are constructed around the donor clusters like the rings and moons around Saturn. The aureoles consist of cases which are primarily assigned to another cluster but show great affinity also to the orbited fixed star. This reservoir is used to find better matches in some cases, better than the matches of primary clusters would allow.

An extremely important factor is the relation of the size. Any recipient cluster needs a donor cluster of adequate size. If the recipient cluster is too large, then the most similar donor clusters have to be grouped for the match.

It could be the other way round, as well. The donor clusters could be larger. In this case the obvious solution would be to pick those donors which show the maximum similarity with the recipients within the cluster in question. That is wrong, because it could lead to a biased donor choice, resulting in a distortion of the data to be transferred. A better way of using the donor supply is to choose at random. The unconsidered donors will be assigned to the aureole of that cluster and could be considered for a later improvement of the matching.

As a result of all preparatory steps we have now:

(1) Similar clusters

(2) Based on common characteristics

(3) With reasonable relations in size

(4) And an additional reservoir of suitable cases in the orbits.

Data transfer

Now the actual fusion process can start. The criteria from the donor sample can be transfered into the recipient sample. At the end we will have one set of data, like single-source, but established in two separate surveys.

The concept of the fusion is fairly simple:

– to identify matching donors and recipients by using common criteria.

– the match is used to transfer the relevant information, here viewing probabilities, from donor to recipient.

– after the fusion, donor and recipient will be separated again, each remaining intact in their own sample, with the exception that the recipient has gained additional criteria, the p-values.

One can conclude from this concept, that fusion is just infusion!

It is not necessarily the most similar donor that has to be matched with his twin recipient, but the mean of all distances between the pairs has to be minimised.

If one always matched the most similar pairs, some donors would not find a recipient at all and ordinary donors would have a multiple chance. Minorities among the donors would not be represented within the recipients. Therefore, the average distance between all pairs should be minimum.

To make sure that the same donor will not be matched with too many recipients, each donor will be penalised after each ascription, meaning the distance unit will be weighted with the factor square root of ten. Thus the donor will only be tapped again if no donors with smaller distances can be found.

What data were transfered this way?

Fusion 1 Meter Panel into MA Electronic: Just probabilities of viewing commercial air-time slots.

Fusion 2 MA Electronic into MA Print: The same probabilities of viewing plus commercial radio probabilities.

All other criteria of the recipients remain unchanged. The MA sample structure has not changed through the increase in information. That is an advantage, bearing in mind that panels are difficult to balance!

Results of the fusion

Let us skip the technical details of test runs etc and turn to the results. The first result is that we find donor as well as recipient surpluses. The main reason for the donor surplus lies in the fact that the samples show a different regional disproportion. The MA Electronic has to serve regional broadcasters; MA Print has to serve national magazines as the main sponsors.

The total donor surplus is just over 3,000 cases, leaving 29,000 real donor potential. Why potential still? Well, not all of the 29,000 donors were actually tapped for data transfer, due to the

Table 3

Donor surplus: total fusions

Donor potential	Donor surplus regional disproportion	Real donor potential
14,773 Men	1,439 Men	13,334 Men
17,418 Women	1,733 Women	15,684 Women
32,191 Total	3,173 Total	29,018 Total

Real donor surplus

504 Men	= 3.8 % of real donor potential
385 Women	= 2.5 % of real donor potential
889 Total	= 3.1 % of real donor potential

Source: MA 88 Documentation

different cluster structures within donor and recipient samples. Overall we find 889 donors who were not tapped at all, that is 3.1% of the real potential, more men than women (Table 3).

In the majority of cases it is the other way round. We find more recipients than donors. Ideally the technique employed should make sure that each donor is at least tapped once and that no donor should be assigned excessively.

Table 4 is one example of the MA report, where all six tables are shown. This illustrates that multiple assignments are the exception. In fact, bearing in mind that the recipient sample is some 50% larger than the donor sample, each donor should ideally be tapped 1.5 times. Therefore the assignment distribution is better than it looks at first sight, avoiding a distortion of the distribution of the specific criteria within the recipient sample.

Table 4

Distribution of assignments

None and multiple assignments	Men		Women	
	Donors	recip-ients	Donors	recip-ients
None	203	-	157	
1 Assignment	1,114	1,114	1,159	1,159
2 Assignments	605	1,210	1,049	2,098
3 Assignments	274	822	348	1,044
4 Assignments	117	468	101	404
5 Assignments	12	60	65	325
6 Assignments	4	24	1	6
	2,629	3,998	2,880	5,036

Base: GfK/MA Electronic, sweeps 1 + 2, terrestrial viewers

Source: MA 88 Documentation

For judging the quality of the fusion we need a comparison of the common criteria in recipient and donor sample (Table 5). The main differences appear between the meter panel and the MA. Between the two parts of the MA the number of significant differences in the common criteria is much smaller. The reason is clear: the sample structure of the panel is, let us say, TV minded!

But the 64,000 dollar question is: How did the media scores come through from donor to final recipient sample? In the MA report we find *all* comparisons of the media criteria before and after fusion. Conclusion: the matches were successful. However, there are some exceptions, especially in the small sub-group of cable homes. But compared with last years' results, based on a very small number of cable homes,

Table 5

Comparison of common criteria in donor and recipient sample before fusion

	Number of significant differences at 5 % level Men	Women
GfK -> MA Electronic 1 + 2 sweep, non cable		
131 Chi-square tests	86	92
72 T-tests	30	28
GfK -> MA Electronic 3 + 4 sweep, non cable		
131 Chi-square tests	84	87
72 T-tests	38	34
MA El -> MA Pr. 1 + 2 sweep, non cable		
176 Chi-square tests	36	35
102 T-tests	24	28
MA El -> MA Pr. 3 + 4 sweep, non cable		
176 Chi-square tests	41	56
102 T-tests	11	14
GfK -> MA El 1 – 4 sweep, cable		
131 Chi-square tests	76	85
72 T-tests	10	18
MA El -> MA Print 1 – 4 sweep, cable		
176 Chi-square tests	25	35
102 T-tests	16	17

Source: MA 88 Documentation

MA El: – MA Electronic
MA Pr: – MA Print

we find an enormous improvement. So, with the increasing cable density, 100,000 homes per month, this problem will vanish in the future.

Table 6 seems to prove the MA conclusion. In MA's words: the data at the final stage are an image of the donor scores! However, the difference in the universes can only account for one percentage point of this difference, but we see relative differences of up to 4%.

The real touchstones for the success of the fusion are cross comparisons before and after fusion for a combination of media and common criteria. This will make it evident whether the relations of common criteria will remain or change when they are crossed with specific criteria such as TV viewing.

The changes in the significant differences are crucial. If the number of significant and non significant differences before crossing remains after crossing with the specific criteria, then this verifies that the original relations are maintained after fusion. To judge the quality of the fusion, we therefore have to establish the number of such changes. An increasing number would indicate a bad match, the fused set would be distorted. A decreasing number, as the majority of the MA shows, demonstrates that differences in common criteria between donors and recipients vanish with the crossing with the media criteria. MA views this as an indicator for the improvement of the data structure. Anyway, fusion seems to smooth the data.

For explanation let us pick two examples from Table 7: GfK fused into MA Electronic, cable households, and GfK into MA Electronic first half, terrestrial, both based on women. The specific criteria are the commercial time slots on German TV stations. They have been crossed with 131 common criteria, resulting in 1,703 and 4,752 matrices respectively, which means 1,703 recipients and 4,752 chi square tests between donors and recipients. Of the total tests, the

Table 6 TV Reach after fusion

Time slots	GFK TV house holds only %	MA electronic all households %	MA print all households %
ARD: MEN			
18.00 - 18.30 hours	14.7	13.9	13.9
18.30 - 19.00	18.3	17.6	17.5
19.00 - 19.30	19.7	18.9	18.9
19.30 - 20.00	21.3	20.5	20.5
Average half hour	18.5	17.7	17.7
ARD: WOMEN			
18.00 - 18.30 hours	12.6	12.0	11.7
18.30 - 19.00	17.4	16.7	16.3
19.00 - 19.30	21.2	20.3	19.9
19.30 - 20.00	23.3	22.5	22.1
Average half hour	18.6	17.9	17.5
ZDF: MEN			
17.30 - 18.00 hours	14.3	13.3	13.2
18.00 - 18.30	15.9	15.0	14.9
18.30 - 19.00	20.7	19.7	19.7
19.00 - 19.30	21.7	20.8	21.0
Average half hour	18.2	17.2	17.2
ZDF: WOMEN			
17.30 - 18.00 hours	14.6	14.0	13.6
18.00 - 18.30	17.3	16.6	16.2
18.30 - 19.00	22.1	21.2	20.9
19.00 - 19.30	21.6	21.0	20.7
Average half hour	18.9	18.2	17.8
H 3: AVERAGE HALF HOUR			
Men	0.9	0.8	0.8
Women	0.8	0.7	0.7
RTL +: AVERAGE HALF HOUR			
Men	1.0	0.7	0.7
Women	0.9	0.7	0.7
SAT 1: AVERAGE HALF HOUR			
Men	1.0	0.7	0.6
Women	0.9	0.6	0.6

Table 7

Cross comparisons

	Men		Women	
	Significant		Significant	
	before yes / after no	before no / after yes	before yes / after no	before no / after yes
GFK -> MA EI 1 + 2 SWEEP NON CABLE 1,703 Comparisons, of which	116	111	150	28
GfK -> OMA EI 3 + 4 SWEEP NON CABLE 1,703 Comparisons, of which	128	97	227	43
MA EI -> MA PRINT 1 + 2 SWEEP NON CABLE 4,752 Comparisons, of which	445	235	383	258
MA EI -> MA PRINT 3 + 4 SWEEP NON CABLE 4,752 Comparisons, of which	612	221	662	230
GfK -> MA EI 1 – 4 SWEEP CABLE 1,703 Comparisions, of which	157	106	130	174
MA EI -> MA PRINT 1 – 4 SWEEP CABLE 4,752 Comparisons, of which	267	480	549	259

majority of comparisons showed significant differences before the fusion which disappeared after the data transfer. Fusion has smoothed the structure.

The other way round, we find 174 significant differences after fusion which were not significant prior to the fusion. A quote from the MA 87 documentation states that in these cases the fusion has accurately expressed differences which were not existent before the transfer (sic). The MA also concludes that the balance indicates very little shifting in the relations, which may be true for the majority of the terrestrial homes. But the exception are the Cable TV viewers, due to the small sample sizes. Do we not know that God is always on the side of the larger battalions.

Matching

So much for fusion. Not everybody in Germany is in favour of the fusion technique. Two reasons account for this: costs of the process and the black box, the part not under control. Therefore the German Consumer Analysis, the equivalent of the TGI, followed another route to create maybe the most comprehensive data base, the super brain for media planning in Germany. They use Matching. Actually, the Germans call it injection, but as I said fusion is really injection. I prefer Matching, but I would not quarrel about names.

What does Matching mean? It is a technique through which single or few criteria from one

sample will be aligned with another. It is therefore a direct successor of the early attempts. We have seen that fusion marries the sample globally, regardless of the individual media. Although the global result of the fusion is a perfect match, if successful, in practice it can distort individual media probabilities. Matching avoids this by aligning the criteria of donor and recipient *separately* for each medium. Separate means, if 100 titles have to be brought in alignment, we have to use 100 individual matches, title by title.

Particular emphasis has to be put on the linkage criteria, reading frequency, broadest audience and time filters. They have to be designed to match in both samples. Therefore matching requires a lot in the early stages of data collection: survey and questionnaire design, fieldwork and the like. Matching requires that the assumptions are right, that we compare like with like and not apples with pears. Matching requires a need to think first, then start the job.

Let us have a look at the matching process. The first question: what direction should be ascribed? We bring the Consumer Analysis in line with our Media Standard, the MA. Therefore we have to pick from the MA the broadest audience of the first title to be matched, break down the broadest audience by reading probability classes, which are constructed by tree analysis, then establish the demographic profile, all done separately for male and female respondents.

In the next step, we have to reconstruct the cells within the Consumer Analysis for demographic criteria compatible to the working variables of the MA.

The problem in reconstructing can be seen in the different sizes of the broadest audience and different cell structures. The different sizes can be corrected by opening or narrowing the time filters. The different cell loads can be corrected by modifying the probabilities within the cells iteratively.

This rather strict procedure of adjusting the readers according to reading probabilities and demographics guarantees a close alignment of the results, not only for the broadest audience, but also for the average issue readership and the cume. However, the limited size and structure of some sub-groups can result in more or less severe deviations.

The latest creation in the German terminology is taken from the sky-labs and is called Andockung = docking. It is invented by the agencies and clients and is not really a model or technique of its own. It should merely illustrate the need of data linkage in view of the limits of respondents and individual surveys. And it is directed more towards my last example: Matching, the preferred method of the advertising industry in Germany, preferred, because it is based on transparent procedures, whereas the main part of fusion remains in a black box.

I should try to summarise what I have said so far. Technically, fusion is a wonderful method of enriching one data base with the information of another, a different one. It does not melt two complete data bases into one, it blows up the information in the recipient data base. But one has to be sure whether all preconditions are really matching, if the information to be transfered is really of the same kind and quality as can be expected from the recipient side. The minimum requirement, in my view, is that both sets are established on a representative random basis. If we compare panels with sophisticated MA samples, we compare apples with pears.

The German TGI approach, matching, fulfils all this. It is designed carefully to match coverage data from different sources. Sample, questionnaire, fieldwork and the like are very similar. But then again it is easier, because the match is done for data that should be alike, the

reach of magazines on both sides. With fusion, the TV viewing probabilities are a completely new piece of information, although within the MA Electronic we establish TV viewing by questions about the day spent yesterday as well.

Segmentation: the miracle of reading probabilities

When the Germans changed their Readership Analysis LA from a static model, just establishing and reporting the average issue readership, to a dynamic model, considering AIR, broadest audience plus the cumulative values based on the claimed reading frequency, they experienced different results for the different units AIR and K1. The AIR is based on the recency question, K1 is calculated from the frequency claims and should in theory be equal to the AIR, if the respondents answer both questions correctly. But other countries have experienced those differences as well.

After some discussion in the early 70s, the Germans thought of a solution or convention and aligned the calculation from the frequency claims to the recency results, assuming that the recency results were the harder currency. The alignment resulted in a new technique, the segmentation or tree analysis. The segmentation (known as AID run in America) selects on the basis of readership, frequency and demographics, and creates discrete and mutually exclusive cells. Each of these cells in turn has an element of AIR readers which can be related to the broadest audience, derived from the time filter, to establish reading probabilities for that particular cell.

In practice, this works perfectly well for national media on a national basis. That means, in general we notice close matches of recency and frequency based AIRs. However, if we go by target groups or regional units, we notice severe deviations. Let me demonstrate this by showing the figures for the regional units of *Bild-Zeitung*, by no means a small paper (Table 8).

Table 8

Bild-Zeitung AIR before and after segmentation

Regional unit	AIR before (LpN) %	AIR after (LpA) %	Index before = 100
Hamburg	39.1	38.5	98
Berlin	7.3	8.0	110
Hannover	30.7	31.0	101
Düsseldorf	27.3	26.7	98
Ruhr-Ost	22.8	23.8	104
Ruhr-West	22.5	23.0	102
Köln-Bonn	14.3	14.6	102
Aachen	24.8	25.2	102
Frankfurt	27.4	27.2	99
Stuttgart	17.7	17.3	98
München	19.0	15.6	82
Total Units	24.4	24.5	100

Source: MA 1988

The indices demonstrate very clearly the fact that the small gains in the majority of the regional units are offset by a tremendous loss in one particular unit, in München. There is no reasonable explanation for this. Despite the fact that the reading probabilities for *Bild-Zeitung* are particularly high in München (all cells containing readers in München showed an above average reading probability) astonishingly, the result of the tree analysis showed particularly low audience levels. Clearly, the system is flawed. In spite of this, the MA holds up this system as the Gold Standard. There are possible solutions, one of which would be to conduct the segmentation on a regional basis. However, the appropriate

committees of the MA have found it unnecessary to move immediately. Anyway, this is not merely an academic problem. Let me demonstrate the fact by the figures in Table 9.

That means that, by moving away from the recency based AIR, the harder media currency, to the frequency based AIR (it should be the other way round), *Bild* loses 70,000 claimed yesterday readers overnight, not through sampling errors or wording of the readership question, but solely by means of an inadequate computer model.

Table 9

Readers per day

	LpN Recency based	LpA Frequency based
Bild-München	430,000	360,000