Worldwide Readership Symposium

Vienna, October 2007

TOUCHPOINTS: INTEGRATION FOR MULTI-MEDIA PLANNING

Steve Wilcox and Noel O'Sullivan, RSMB Audience Research Ltd

---

1. INTRODUCTION

With the developments in all forms of media communication there has been an increasing demand to understand consumers' behaviour across a range of communication channels and, in particular, the various interactions between different forms of media.

As in most countries, in the UK the key audience currencies have been measured separately, each to a high level of technical expertise. However, after consultation with media agencies the IPA's Media Futures Group recognised that there was a need for cross-media information to inform the campaign planning process.

The key requirements were:

(i) to measure cross media interactions.

(ii) to develop a multi-media planning tool.

(iii) to harness the power and integrity of the separate industry single source survey measurements.

The IPA TouchPoints project creates two separate databases:

TouchPoints Hub Survey

This database was created from a sample survey of 5,000 adults undertaken by TNS.

Each respondent completed a detailed questionnaire about their media usage, attitudes to media, shopping habits, lifestyle and attitudes. In addition, each was given a PDA on which they were required to register their location, activity, companions, mood and media usage for every half hour across a 7 day period.

TouchPoints Integrated Planning Database

Although the Hub Survey is a single source survey of cross media consumption it doesn't have the granularity of the individual industry surveys, definitions of exposure are different and it delivers information for only seven days. The integrated planning database is a more complex database designed to provide data at a more fragmented level such that the user can evaluate campaigns across a large number of media channels.. This is the first UK, industry available consumer centric planning tool, constructed by integrating the industry currencies for national newspapers and magazines, regional press, TV, radio, posters, cinema and TGI onto the specifically designed IPA TouchPoints Hub Survey. Currencies have also been simulated for on-line, SMS and direct mail using data from the hub survey. The integration process was undertaken by RSMB.
The IPA TouchPoints Hub Survey was launched in March 2006 and the Integrated Planning Database in October 2006.

This paper will concentrate on the integration methodology and the planning model. In particular we will explain how the value of the hub survey has been optimised in creating the all important media imperative hooks. We will also explain the methodologies used to derive personal probabilities and the associated multi-media reach and frequency model.

We will concentrate primarily on the process of appending national and regional press data onto the planning database. However, reference is also made to other media integrations as the ultimate aim was to ensure the correct multimedia relationships for respondents on the database.

2. DEFINING THE DATA INTEGRATION REQUIREMENTS

One of the most important principles in a data integration exercise is to select techniques which recognise the end objectives and complement methodologies or modelling used in the application of the data. As with all areas of statistics, no single solution is universally correct for all data integration exercises.

In this case the end objective is to create a multi-media schedule planning tool across TV, national and regional press, radio, outdoor (posters), cinema, internet, direct mail and SMS text. Each media is planned by channel, day-part, title, or whatever level of granularity has been agreed by the IPA and its sponsors, to create a set of media events. The user defined input for each media event is the number of times it is used in an advertising schedule or the required audience size or GRPs (gross rating [percentage] points). The output is a multi-media reach and frequency analysis, showing the unique contribution of each component media and their combined effect. The marginal reach and frequency analysis for each media must be comparable with results obtained from planning applications of the respective currency database. The emphasis is on planning and therefore prediction of normal rather than specific behaviour.

3. THE HUB SURVEY

The basic principle of a data integration exercise is to simulate the results of a single source survey through a model which is driven by the information common to a series of independent survey databases. For GB's media trading currencies, the usable common information is largely limited to geography and demographics. The success of any integration exercise will depend upon the ability of this common information to explain systematic variations in the interactions between consumptions of the different media. In practice each media trading currency database allows us only to assess the ability of common demographic information to explain systematic variations in the consumption of each individual media separately, not their interactions. Whilst such demographic based integrations can provide acceptable solutions (at least as good as sophisticated demographic targeting), there is no real basis for validation (by definition the "real" single source data is not available) and there will always be a concern that differences in behaviour go beyond that which can be explained by demographics.
The TouchPoints solution is the hub survey which provides reasonably extensive single source media consumption data, but without the granularity nor time span of the media trading currencies. The hub survey forms a respondent level database into which the individual media currencies can be integrated using respondent level data fusion, profile matching or calibration, dependent upon the structure of each survey database and as summarised below:

Television	-	Fusion from BARB
Magazines and National Newspapers	-	Fusion from NRS
Radio	-	Fusion from RAJAR
Regional Press	-	Profile matching from JICREG
Posters	-	Calibration from POSTAR
Cinema	-	Calibration from CAA admissions
Internet	-	TouchPoints
Direct Mail	-	TouchPoints
SMS	-	TouchPoints
Product Usage	-	Fusion from TGI

The key is creative use of the TouchPoints demographic and media data to create a powerful modelling link to each of the media currencies. By definition, the media component of the link must be limited to the individual media which is to be integrated into the hub (e.g. the National Readership Survey can be integrated using only a demographic and readership link). In the case of data fusion, the link will be based upon a set of demographic and media imperative variables which can be constructed in both the hub survey and the media currency database and whose power can be judged in terms of cross-media interactions using the hub survey.

3.1 Re-Engineering the Hub Survey

An important step in the integration process is the re-engineering of the TouchPoints sample to provide a robust hub survey. As a minimum, the hub survey needs to be weighted to universe profiles derived from a large random probability sample. Further, the TouchPoints sample is relatively small compared to the media currencies: a standard fusion would use only a small proportion of the currency survey sample and its effectiveness would be severely reduced. The solution was to fuse the TouchPoints survey onto the BARB Establishment Survey which generates a 50,000 adults sample in six months of fieldwork. This does not give any precedence to television because the only interest is in the demographic and geographic data. The result is a large, high quality random sample featuring demographics and TouchPoints single source media data. The other candidate survey was of course the NRS but the ES was preferred because it yields around three times the sample.

4. DATA FUSION

The currencies for television, radio, magazines and national press have been integrated using respondent level data fusion. Fusion is driven by a set of variables which can be found in the two surveys to be fused, often called common variables or hooks. For example, if there is a respondent in the hub survey who is a heavy TV viewer but light radio listener, then the fusion process will try to match them separately with a heavy TV viewing BARB panel member and a light radio listening RAJAR diary respondent. The hooks available for each media fusion comprise demographics, geography and a set of media imperatives.

4.1 Multimedia Input Variables

In order to determine the relative importance of the characteristics to be matched, data relating to patterns of multimedia usage are needed.

The hub survey delivers a wealth of media consumption for each respondent from the weekly diary data. From this activity behaviour across time slots was generated for:

Travelling
SMS texting picture messaging
Radio (by station)
Television (by channel)
Reading (by title)
Internet (by activity)
Cinema
Direct mail.

In order to make optimum use of these data the following procedures had to be applied:

Advertising Expenditure

The data was weighted to reflect the advertising expenditure for each medium.

Factor Analysis

A factor analysis was applied to the data. This is a statistical multivariate technique that transforms the data by removing all correlations from the dataset to give a smaller dataset of independent factors that fully describe the patterns of multimedia consumption for each respondent. This enables the data to be analysed correctly without over-representing correlated activities.

4.2 Fusion Hooks

The list of linking variables used for the TouchPoints/NRS fusion is listed below.

Sex (Critical)	Housewife (Surrogate)
NRS/ES Segment (Critical)	Marital Status
Age	Media Imperatives
ACORN Weighting Factor	Multi Channel Home
Chief Income Earner	Number of Children
Ethnic Group	Presence of Children
Full Time Job	Size of Household
Head of Household (Surrogate)	Social Grade
Home PC	Terminal Age of Education

The following sections expand on the development of the more complex linking variables.

4.3 NRS/ES Segment

Although weighting is incorporated in each of the main industry sources to make the data representative of the population, each sample survey has its own built-in geographical disproportionality in the sampling frames due to its different objectives; for example, the National Readership sampling frame is designed to deliver more AB respondents, with particular wealthy ACORN areas oversampled.

Each of the industry surveys contains their own geographical segment definitions which help to describe the areas of this disproportionality. It is important that differentially represented groups are critical linking variables in the fusions i.e. recipients will be forced to match to a donor in these groups. For this reason, for each fusion an interlaced geographical segment definition (based on the respective area segments of the two surveys) was defined in order to account for the geographical disproportionality in each structure. In each fusion this geographical segment was set as a critical linking variable.

For the fusion of the RAJAR and NRS respondents onto the hub survey an interlaced structure of their geographies with the Establishment Survey geographical segment (i.e. the geography of the re-engineered database) was used.

4.4 Surrogate Variables

One problem that continually faces fusion practitioners is to find common variables between surveys and ensure that they are consistently defined. Factual information such as age is straightforward (as long as you have the necessary detail). However, many classifications can have different definitions (e.g. due to wording of the respective questionnaires) or 'out of sync' definitions (e.g. Presence of Children 0-15 vs. Presence of Children 0-18). In many cases, the latter may still be used on the basis that it is an important discriminator and the slight inconsistency in the matching is better than having no matching at all. For the former, in many cases these variables are dropped as the inconsistencies are too significant.

For the TouchPoints fusions there were inconsistencies in Head of Household and Housewife classifications across all the major industry surveys. Given that it was felt these classifications could provide a useful distinction between individuals we developed surrogate classifications.

Using, the Establishment Survey definition of these variables we performed a CHAID analysis. A CHAID analysis is a statistical procedure that selects a set of classifications and their interactions that optimally predict the dependent measure, in this case the summary of cross media consumption. The developed model is a classification tree that shows the partitions that explain the dependent measure.

From this analysis the surrogate Head of Household classification was defined as anyone in the following groups:

Adult in household where number of adults = 1
Male, Chief Income Earner
Female, Chief Income Earner, not working full time
Male, not Chief Income Earner, not working full time, aged over 44 years old

Similarly, the surrogate Housewife classification was defined as anyone in the following groups:

Adult in household where number of adults = 1
Female, Aged 21-26, Number of adults = 2
Female, Aged 27-70
Female, Aged 71+, Household size = 2

From the Establishment survey, more than 90% of the respondents had the same Housewife or Head of Household definition based on these criteria. Exact matching is not necessary; the above is a consistent demographic classification across surveys that can be used as a linking variable and which is designed to correlate highly with these original classifications thus preserving the majority of any discrimination peculiar to these groups.

4.5 Media Imperatives

A media imperative is a summary of each respondent's consumption of the media to be fused. To be used as a hook, we must be able to construct the media imperative in both the hub survey and the currency survey to be fused, and be reasonably confident that they are measuring the same thing. For example, it is possible to calculate hours of viewing by time segment, by day of week, by channel group in both the TouchPoints Hub and the BARB panel.

These patterns of viewing must be summarised to form a usable set of fusion hooks. In order to avoid subjectivity in this process, a principle components analysis was used to construct the media imperatives. A principle component is a linear combination (like a regression model) of hours of viewing by time segment, day of week and channel group which maximises the diversity between individuals. A relatively small number of principle components explain the majority of the systematic variation between individuals.

For the BARB and RAJAR fusions the diary activity could be used in this analysis as this respective information could be recreated on their respective currency surveys. For the NRS fusion, however, this was not the case as time spent reading publications is not collected on the NRS. For this reason, the NRS style TouchPoints self-completion questionnaire information had to be used; note also that this was based on the NRS measure of frequency of readership and therefore had a lower level of fragmentation than radio and television data. In order to create more powerful measures, publications were collated into 17 larger more meaningful groups (e.g. [highest] frequency of readership of any quality daily newspaper).

The principle components were constructed in the TouchPoints Hub Survey, giving a functional model. This principal components analysis yielded 6 principal components for the NRS Fusion. Given a particular respondent's frequency of readership the value of the principle component was calculated for each TouchPoints respondent and each NRS panel member, using the same functional model.

This process was controlled to allow for differences due to survey effects in both overall levels and variation in viewing levels.

5. THE FUSION PROCESS

The principle of the data fusion process is to find a respondent in the media currency (donor) survey who has the same demographic and media imperative profile as a particular respondent in the TouchPoints hub (recipient) survey. When a match is found, this donor's media currency data is then assigned to the TouchPoints recipient and replaces their TouchPoints media data.

5.1 Importance Weights

A large number of demographic and media imperative hooks (about a dozen of each) were used in the matching process. Inevitably it is not possible to find exact matches across all hooks. Where compromises have to be made it is necessary to give precedence to the more important hooks. Therefore we need to quantify their relative importance or discriminatory power through analysis of variance.

A key feature of this particular fusion exercise is that because we have the single source hub survey, we can evaluate the hooks in terms of the true object of the fusion, i.e. volumes and patterns of consumption across all media as measured by the TouchPoints half-hour diary. In this respect, for example, the importance of a television viewing based media imperative is tempered by its relative inability to explain variations in consumption of publications. Demographic hooks have a chance to gain their rightful place in the hierarchy.

Obviously there is a separate set of hooks for each media fusion and therefore a separate set of importance weights. For each fusion a multivariate analysis of variance technique has been used to consolidate the patterns of consumption across all media to construct a single importance weight for each fusion hook.

5.2 Distance Measurement

RSMB's data fusion algorithm uses a variation of Mahalanobis' Distances to quantify the similarity between recipients and potential donors. This allows for correlations and differences in scale between the hooks. There is also a cohesive piece of statistical theory which justifies the formula used to incorporate the importance weights into the distance measurement.

The analyses needed to calculate the importance weights are extensive and the computation of Mahalanobis' Distance is intensive. Some fusion practitioners have argued that this sophistication is unnecessary. This may be true if there are only a few demographic hooks but an equivalent to Mahalanobis' Distance is required when, as in this case, there are many.

The optimisation routine for pairing donors with recipients makes a trade off between the closeness of their hook profiles (as measured by Mahalanobis' Distance) and the donor frequency distribution (the number of times each potential donor is used). The greater the control of the donor frequency distribution, the more likely we are to preserve the media currencies in the integrated database.

5.3 Calibration

In the final integrated database, at a total level, all probabilities are calibrated to the published levels achieved in their respective surveys. However, given the objective of the fusion in trying to match similar respondents we would expect the fused results to be similar to the original level and less calibration would be needed. For RAJAR and NRS the level of calibration was low; for the BARB fusion the level of calibration required was higher at up to 4% and this may have been due to differences between sample structures for BARB and TouchPoints.

To illustrate the low levels of calibration needed, the following gives a summary of the Average Issue Readerships percentages for all publications, along with a random sample of publications, for industry versus the pre-calibrated integrated TouchPoints fused data:

Publication Title	NRS %	Fused %	Difference	Index
All Publications (c.250)	2.68	2.69	0.01	100
Publication A	18.51	17.64	-0.87	95
Publication B	15.38	14.23	-1.15	92
Publication C	2.71	2.95	0.24	109
Publication D	2.71	2.60	-0.11	96
Publication E	1.17	1.17	0.00	100
Publication F	1.16	1.18	0.02	102

Source: NRS 12 months ending June 2005
BARB Establishment Survey 6 months ending June 2005
IPA TouchPoints 2005

6. EVALUATION

The IPA commissioned an independent data integration expert, Ken Baker, to conduct an independent evaluation of the integrated database. The following is based upon the author's interpretation of Ken Baker's appraisal document and is not a comprehensive summary.

In such an evaluation, there are two key questions to address:

- Has the integration methodology made the best possible use of the information available?
- Has the integration succeeded in recreating the true relationships between the media which are the subjects of the integration?

We will concentrate on the second of these questions, which is the acid test of the integration, in relation to the TV, readership and radio results where data fusion was used.

6.1 Regression to the Mean

Data fusion will recreate the true relationship between any two variables X and Y from independent surveys, if and only if:

- The other variables which donor and recipient surveys have in common are sufficient to fully explain this relationship.

- The fusion algorithm can match respondents exactly on all common variables.

Any violation of these two conditions may result in the loss of efficiency known as regression to the mean. Consider the hypothetical example shown in the table below:

	All Respondents	**Average Issue Readers of Newspapers X (Real)**	**Average Issue Readers of Newspaper X (Fused)**
Average hours spent watching TV station Y per week	6.50	8.20	7.40
Average hours spent watching TV station Z per week	3.40	2.80	3.10

In both cases the fused data has regressed about halfway towards the mean for "All Respondents", i.e. the fused data is closer to the sample mean than the real data.

To judge the efficiency of the fusions, it is possible to compare fused (predicted) media interrelationships with the media interrelationships as observed in the TouchPoints hub survey. This comparison was made for around 25,000 cross tabulations of TV viewing, radio listening and press readership.

At the risk of over summarising an extensive document, after taking due account of the effects of sampling variability, the overall estimate of regression to the mean is 7%, i.e. 93% of discriminatory power has been retained by the fusions.

7. PROFILE MATCHING

Profile matching was used to integrate the JICREG regional press currency into the hub survey. JICREG data is not available as a respondent level database. Instead the number of non-readers and the average issue readership are reported for each of a number of demographic groups.

The process for ascription for each title was as follows:

Within each of twelve mutually exclusive and exhaustive demographic groups (based on age/class/gender combinations) and within each title's circulation area, a random selection of hub survey respondents who had read a local newspaper were classified as readers of the title. This sample was selected so that the weights accumulated to that publication's audience reach. This sample was then assigned a probability of reading equal to the average issue readership for that demographic group. Readership of a local newspaper was garnered from the relevant NRS field that had been fused to that hub survey respondent.

For some publications, the criteria for this had to be relaxed in order to achieve a big enough sample of readers in the hub survey.

The following was the hierarchy for increasing the available sample:

i) include NRS non-readers of a local newspaper.
ii) expand to whole postcode districts that encompass the postcodes of newspaper's circulation area.
iii) expand further to postal towns of the newspaper's circulation area.

The probabilities were calibrated further to account for any discreteness issues.

All of the above relaxing of constraints are legitimate because the aim of the ascription is to ascribe a probability to a hub survey respondent who has similar multimedia behaviour to a reader in that publication's circulation area. We believe the compromises made do not contradict this underlying principle.

Note that even including smaller publications, the AIR levels from the integrated database were within +/-1,000 for 97% of the demographic segmentations across for all publications (c.10,000 audiences).

8. REACH AND FREQUENCY PLANNING

At the end of the data fusion process, the integrated database is a large representative, sample of the population. The information available for each respondent is:

- Demographic, geographic and geodemographic classifications;
- A full Target Group Index product usage and ownership record;
- A media usage record from each of the integrated media currencies which reflects the levels and patterns of that respondent's media consumption as measured by TouchPoints.

The primary application of the integrated database is multi-media schedule reach and frequency analysis. It is recognised that the database will also be used for non-commercial applications, but the reach and frequency application is discussed here for illustration of the thought processes involved.

For television, in theory it would be possible to use the BARB panel's long term viewing records so that actual contacts with a TV schedule of commercial spots could be counted for each individual. However, all other media currencies have only short-term measurements of exposure and depend upon probability expansions to estimate longer-term contacts with a schedule.

Further, each media currency has a different probability model (the National Readership Survey has several) for the expansion. This partly reflects the fact that the inputs to the models are different for each medium. For example, the NRS readership data is recency and frequency whilst the RAJAR radio data is from a one-week quarter hour diary. A common denominator is required which embraces all media. The solution we adopted is personal probabilities. Not only do we believe this to be the best statistical solution, it is also probably the only practical way of getting systems into the market place.

If a person does or doesn't make contact with a specific media event (e.g. a TV spot, yesterday's issue of a newspaper, a radio station in a particular quarter hour) then their personal probability is 1 or 0. However, this is not indicative of that person's probability of making contact with the equivalent event on another day or in another week. As a basis for estimating that person's long-term contacts it is useless because it doesn't allow that they might change from 0 to 1 in the course of a schedule. What we need is their underlying probability of making contact with each media event, a number between 0 and 1. Then if that media event is repeated a number of times, we can use a Binomial expansion to estimate a particular person's probability of making 0, 1, 2, 3, … contacts. Simple probability theory allows estimation of each person's "personal frequency distribution" within and across media. These respondent level frequency distributions are then aggregated to form the full sample multi-media reach and frequency analysis.

The decision to use personal probabilities means that the database is transparent to the bureaux, guaranteeing a level of consistency in the market place.

8.1 Calculating Personal Probabilities

The methodology used to calculate personal probabilities varies by media according to the survey data available.

For TV, long-run data is available for each panel member and individual personal probabilities were calculated as a twelve week average for each time segment and channel.

For radio, national press and magazines, segmentation analyses are used to split the sample into homogeneous groups separately for each station or title.

For regional press the personal reading probabilities were allocated as part of the profile matching process.

For outdoor and cinema, personal rates of exposure were allocated in the calibration process, these to be used in conjunction with a Poisson rather than Binomial expansion.

Where necessary and to tidy up random distortions in the data fusion processes, the final stage is to calibrate personal probabilities to force consistency with the original media currencies for a number of key demographic groups.

The press related models are now discussed in more detail.

8.2 National Readership Survey

The success of the segmentation approach depends upon being able to split the sample into a number of homogeneous groups within which each individual has the same underlying probability of reading. Our approach was to conduct an independent segmentation for each title. The segmentation was obviously based upon demographics but significantly improved by incorporating the NRS claimed frequency question – this is the key to preserving the all important duplications between titles. The following is a quote from a paper by Jerome Green[1]:

”Assigning probabilities separately by frequency class is the vital step which helps preserve intermedia correlations. The correlations are not perfectly maintained, however, as will be shown”.

Of course we have done our own evaluation of how well the correlations (title duplications) were maintained.

[1] Reference

Jerome D.Greene: Personal Media Probabilities. 1970 Journal of Advertising Research

---

For each title, the segmentation analysis resulted in up to 50 segments dependent upon sample size. Having identified a homogeneous segment, the personal probability for each individual is numerically equal to the average issue readership for the group. The duplication between two titles for a particular individual is calculated by multiplying their probabilities. Individuals’ results are then summed to generate an overall average. These can then be compared with “real” duplications from the original NRS AIR data. Here’s an example for two reasonably large titles:

Duplication Between Titles

	Title 1 AIR	Title 2 AIR	Duplication
Actual	4.3	11.8	0.73
Random	4.3	11.8	0.51
Segmentation	4.3	11.8	0.71

Source: NRS 12 months ending June 2005

The actual duplication is 0.73. A useful benchmark is derived by making no segmentation at all and simply multiplying the total sample average probabilities (100 x 0.043 x 0.118) – effectively an assumption of randomness. The segmentation delivers a duplication which is very close to the actual, only 3% under. For all titles the average difference is only 4% under, very much a second order effect.

In theory, the effect on an under-estimation of the duplication will be an over-estimation of schedule reach, but we expect this to be a small effect. To demonstrate the order of difference found in practice, here are the results of ten newspaper and magazine schedules provided by KMR Software.

Number of Titles	Number of Insertions	GRPs		Reach	
		NRS	TouchPoints	NRS	TouchPoints
3	3	30.1	30.1	25.4	24.7
3	6	60.2	60.2	31.0	30.5
4	4	20.1	20.1	18.6	18.5
4	8	40.2	40.2	23.7	23.7
3	10	90.7	90.7	39.0	40.0
6	6	20.5	20.5	14.4	14.3
6	12	41.0	41.0	18.7	18.6
5	5	9.2	9.2	7.4	7.3
5	10	18.4	18.4	9.7	9.5

Source: NRS 12 months ending June 2005
IPA TouchPoints 2005

As a final stage in the segmentation analysis, the personal probabilities are calibrated to the NRS average issue readership figures to tidy up differences between the hub and NRS survey sample profiles. This explains why the GRPs match exactly. On average, for these 10 schedules, the reach is actually under-estimated but only by 0.08 of a percentage point or 0.4%.

8.3 Regional Press

As explained in section 7, JICREG regional press currency data is not available as a respondent level database and has been integrated using profile matching. This means that the personal probabilities (AIRs) are already provided in the form of a segmentation. Therefore the JICREG currency and TouchPoints data and reach and frequency models are already aligned.

9. SUMMARY

The TouchPoints Integrated Planning Database has been constructed by integrating the industry media currencies into the specifically designed IPA TouchPoints Hub Survey.

The data integration techniques used are selected to be compatible with the modelling requirements of a multi-media schedule reach and frequency planning tool.

The success of a data integration depends upon the information which is common to the independent surveys. The hub survey extends this beyond demographics to patterns of media usage.

With a single source hub survey, this common information can be ranked in importance in terms of discrimination of cross-media behaviour.

There is also a real basis for evaluation of the integrated database in terms of regression to the mean.

Consumption data from the different media currencies have been converted to personal probabilities to form a common base for reach and frequency modelling.