DEVELOPING A CROSS PLATFORM AUDIENCE CURRENCY FOR GREAT BRITAIN

Steve Wilcox, RSMB and Katherine Page, National Readership Survey, GB

---

1. Introduction

Fusion is an increasingly popular solution to the pressing market requirement for combined print/website audience data. It is clearly 'achievable': a number of fusion solutions are already in place in different markets around the world. However, as the objective is an additional currency we must show that a transparent, unbiased and consistent methodology can be developed.

What are the issues? Is fusion equally 'fit for purpose' in all cases? How can it be made better? This paper presents the results of development work fusing National Readership Survey (NRS) readership and UKOM/Nielsen website audience data.

The development process undergone in the UK demonstrates a number of key issues and practical solutions in respect of managing considerable differences in the structures of the two samples to be fused and preserving website/print duplications, particularly for the 'long-tail' of relatively small websites.

2. Background

Like many other national readership surveys NRS has been tracking the readership of publisher websites for some time. Early experiments included collecting data for 60 or so websites via a self-completion questionnaire that respondents were asked to complete after the main interview had ended.

Asking respondents about the websites they have visited has considerable limitations in respect of what respondents can be expected to recall. Websites visited regularly, or linked to brands the respondent is already familiar with, perhaps because they read the print publication, are more likely to be claimed than one-off or infrequent visits and sites reached in an indirect manner. NRS was very aware of this when looking at the website audience data generated by the self-completion questionnaire. In a market already claiming that there were "too many conflicting numbers" from the various site-centric counts and user-centric panels, here was yet another set of numbers.

With this in mind, in 2010 NRS Ltd released a specification for a test of data fusion between NRS print estimates and an established source of website audience estimates, in effect either the UKOM/Nielsen panel or ComScore panel. After an intensive tendering process, RSMB were appointed to carry out the fusion, and UKOM/Nielsen to supply the website audience estimates.

The objectives of the project are:

- To enable NRS users to assess website audience data alongside and in combination with standard NRS data, providing:
  - A measure of the net deduplicated reach of each NRS publication and its associated websites.
  - A comparison of the respective print publication/website audience profile.
- To enable users to carry out reach and frequency analysis based on a combination of print publications and publisher websites.

It is important that the methodology should be transparent, and the results consistent and reproducible. Much of the focus of this paper is on how best to achieve these aims.

There have been two main tests at the time of writing:

- A 'fold over' test based on the NRS, focusing on data for 30 newspaper and magazine websites to establish how 'good' a fusion could be, and the contribution made by media consumption data to fusion quality

- A fusion between the NRS and the UKOM/Nielsen online panel to address different research designs and techniques, and deal with the practical issues of currency preservation and calibration.

A further test is in progress to refine certain aspects of the methodology, assess the degree of consistency between separate fusions and allow the data bureaux to test analysis of reach and frequency.

The intention is to commence publication of a fused database on a regular basis from early 2012, assuming the final phase of development work and testing proves satisfactory and the NRS stakeholders are convinced of the commercial value of the database.

3. Method

The choice of data integration methodology is dependent upon the ultimate application of the integrated database and the structure of the component surveys. In this case, the NRS also requires that the product is sufficiently clear and transparent to all bureaux who will be delivering analysis systems to the industry. Respondent level data fusion satisfies the objectives and requirements and in our opinion is overall the most effective solution to a data integration requirement for the purposes of mixed media reach and frequency analysis:

- The product looks like a single source survey dataset.
- Each respondent's record contains demographics, readership and website usage.
- The integrity of the readership and website usage records are both retained.

In principle data fusion is a simple concept. If two individuals (one from the NRS and one from the UKOM/Nielsen panel) have the same demographic profile (fusion hooks), then we assume that they are the same person and merge their print and online data records. In practice the execution is necessarily complex:

- There are practical issues related to the structure and data collection methodologies of the input surveys.
- The fusion hooks probably need to be extended to incorporate top-line website usage classifications.
- Optimisation of the fusion process has to be based upon the rigorous application of coherent statistical theory.

It is not the intention of this paper to provide yet another description of the data fusion process. Instead we will focus on the methodological issues which relate to this particular data fusion.

3.1 NRS vs. UKOM

It is important to recognise that the NRS and UKOM are two quite disparate studies with different survey designs, sample structures and data collection techniques.

The NRS is a single sample survey which is representative of the Adult population in Great Britain. Data collection is a one-off, in-home event, using double-screen CAP1 and controlled by an interviewer.

The UKOM measurement is based upon a panel operated by Nielsen. Data collection is via "electronic" monitoring of each panel member's website usage on a continuous basis. The data is made available on a monthly basis as an accumulation of pageviews across the month.

A complication is that the panel has two components designed to separately represent and measure the "at work" and "at home" on-line universes. Nielsen then use a combination model to create a single respondent level dataset which provides a representative measurement of combined "at home" and "at work" on-line activity. This respondent level dataset has a somewhat artificial structure and does not constitute a true representative sample of the on-line active universe either in terms of the demographic profile or the distribution of combined at home/at work website usage. This doesn't matter for the UKOM measurement currency because the modelled dataset is designed to produce a representative measurement of the on-line audiences. However, it creates a challenge for the data fusion process which is designed to integrate two sample structures which are both representative of the same active universe. In particular, the modelled website usage record for each respondent is either artificially high or low. The compromise is to take the combination sample structure at face value, but to use the unmodelled website usage data for the construction of fusion hooks (see section 3.2) and the modelled website usage data to create fused readership/on-line respondent level audience records.

Table 3.1.1 shows some key survey profiles in terms of unweighted and weighted sample counts.

Table 3.1.1 Survey Profiles

	NRS		**UKOM**	
	Sample	Weighted	Sample	Weighted
Active Universe	25,559	37.1m	25,020	35.2m
Men	47%	51%	49%	52%
ABC1	67%	61%	54%	63%
15-34	32%	40%	43%	35%
Work F-T	48%	51%	47%	55%
Heavy Internet	21%	21%	34%	21%
Light Internet	24%	24%	12%	24%

NRS: 12 m/e March, 2011
UKOM: March, 2011

The unweighted samples show large profile differences for Social Grade, Age and Weight of Internet Usage. While unbalanced sample profiles will reflect each survey's sample design, an unconstrained fusion approach may struggle to accommodate such profile differences, leading to a greater potential for currency distortion.

As expected, the weighted sample profiles are broadly the same, although there are some differences for Age and Working Status. A constrained fusion approach is designed to work with weighted samples and may be a better solution for this fusion.

These arguments are developed in later sections of this paper.

The Active Universe is defined to be the number of people who have used the internet in the last month. The NRS universe will define the base for the fused database and this is 3% higher than the UKOM universe. The consequent expectation is that the universe differences alone will contribute a 3% increase in the UKOM currency in the fused database. In addition, profile differences will contribute a relative inflation of the 16-34 UKOM currency and a relative deflation of the Working Full-Time UKOM currency. Of course there will be other factors in the fusion process which will affect the overall achieved level of currency preservation.

3.2 Fusion Hooks

The success of the fusion is dependent upon the degree to which the available information common to the two surveys – the fusion hooks – explains, or correlates with, systematic variations between individuals in terms of readership, website usage and their duplications. This is known as conditional independence. If the fusion hooks get nowhere near to satisfying the assumption of conditional independence, then it doesn't matter how good the actual fusion algorithm is – the fusion just won't work.

The demographic hooks available for this fusion are shown in table 3.2.2.

Table 3.2.2 Demographic Fusion Hooks

Standard Region
Sex
Social Grade
Actual Age
Working Status
ACORN Geodemographic
Household size
Presence of Children 2-10
Presence of Children 11-15

These hooks provide a reasonable description of lifestage and socio-economic status and in our opinion provide a sound base for the fusion. However, they miss a lifestyle component which is so often important in understanding media choices. In line with the experience of the IPA's TouchPoints project in the UK, this fusion was always likely to benefit from a set of directly related media hooks. Relevant data from the NRS survey include recall questions covering recency and frequency of usage for total internet and thirty individual publisher websites.

3.3 Media Hooks

There is some disproportionality in the Nielsen panel by overall weight of internet usage. This is controlled in the panel weighting with a ten break weight of usage classification. This classification has been made available by Nielsen and is an obvious fusion hook. It is replicated in the NRS sample as follows:

- A personal probability of daily internet usage is calculated for each respondent based upon claimed recency and a sample segmentation derived from demographics and claimed internet frequency.
- The NRS sample is sorted from heaviest to lightest internet usage.
- A ten break weight of usage classification is created, such that the NRS weighted profile matches the UKOM/Nielsen panel weighted profile.

The implicit assumption behind this process is that the NRS weighted profile and the UKOM/Nielsen weighted profile are both representative of the same active internet universe. The fact that the NRS and UKOM/Nielsen metrics have different definitions and that the NRS respondents may under or over claim is of second order importance in constructing a classification based upon ranking each sample from heavy to light.

The UKOM/Nielsen panel provides a one month snapshot of respondent level website usage. This includes the average daily reach for each of the thirty websites itemised in the NRS survey. Expressed as a proportion, this average daily reach value can be thought of as a normative personal probability of using the website on a typical day. An equivalent personal probability can be constructed for each NRS respondent, again based upon claimed recency and a sample segmentation derived from demographics and claimed internet frequency. The challenge in using these thirty personal probabilities as fusion hooks is that the percentage of each sample with a positive value (the monthly reach) and the average value differ between the two surveys (see table 3.3.1). Given the differing data capture techniques (recall vs. electronic) these differences are to be expected; of course, this is one of the key reasons for the data fusion of the UKOM currency.

Table 3.3.1 Monthly and Daily Reach

Website	Monthly Reach %		Daily Reach %	
	NRS	UKOM/Nielsen	NRS	UKOM/Nielsen
1	12.9	6.8	1.35	0.49
2	9.6	13.1	1.85	1.09
3	9.5	17.6	0.73	1.23
4	8.4	2.6	0.66	0.11
5	8.3	6.7	1.08	0.74
6	6.9	12.0	1.15	0.93
7	5.3	14.7	1.26	1.59
8	4.0	5.6	0.58	0.29
9	2.9	1.0	0.25	0.06
10	2.6	1.9	0.45	0.15
11	2.3	1.2	0.19	0.06
12	2.1	4.1	0.24	0.21
13	2.0	4.1	0.26	0.24
14	1.9	3.3	0.26	0.17
15	1.8	2.0	0.32	0.02
16	1.4	1.1	0.09	0.05
17	1.4	1.7	0.34	0.2
18	1.0	0.5	0.11	0.03
19	0.7	0.6	0.15	0.07
20	0.7	1.5	0.08	0.08
21	0.6	0.2	0.10	0.01
22	0.5	0.4	0.09	0.02
23	0.5	0.9	0.07	0.05
24	0.3	0.4	0.05	0.02
25	0.3	0.4	0.04	0.04
26	0.2	0.4	0.07	0.03
27	0.1	0.7	0.02	0.06
28	0.1	0.9	0.05	0.07
29	0.0	0.4	0.0	0.03
30	0.0	0.3	0.0	0.02

NRS: 12 m/e March 2011

UKOM: March 2011

For each website separately, the following process is used to create a like-for-like hook value in the two surveys:

- Sort each sample from highest to lowest probability and lay them side by side.
- Equalise the weighted sample counts of non-users by setting the smallest probabilities to zero in one of the surveys.

- The NRS has a larger but still manageable set of discrete probabilities. Use these to create percentiles of the weighted sample.

- Segment the weighted UKOM/Nielsen sample into the same percentiles. Because the UKOM sample has only a small set of discrete probabilities (maximum is the number of days in the month) only a proportion of panel members with the same UKOM probability will be required for each segment. In practice it doesn't matter which ones we take because we only want their UKOM probability.

- Calculate the average UKOM probability for each segment and assign this to all respondents in the equivalent NRS segment.

- All UKOM/Nielsen panel members retain their original panel probability.

The implicit assumption is that the two weighted samples are both representative of the same probability distribution for each website. The weakness of this approach is that the NRS or UKOM/Nielsen respondents who have had their positive probabilities converted to zero, may be fused "at random" to the large pool of non-users in the other survey. However, this damage is limited by the control of all the other hooks and the fact that only the lightest users have been deleted. Overall, 56% of monthly users in one or other of the two surveys have been converted to non-users, but this is a long tail of light users who only account for 18% of the average daily reach. The key objective is satisfied in that the heaviest website users are identified in the hook in both surveys.

3.4 Factor Analysis

Among the 30 publisher websites used as fusion hooks, early results showed that the fusion worked better for the larger penetration websites. This is to be expected because the calculation of importance weights is designed to evaluate explanatory power in terms of all publications and websites, therefore larger penetration hooks will dominate. It may well be that the set of 30 publisher websites currently available is not optimum in terms of their correlations with all 155 publisher websites available from the UKOM panel. Therefore, with a look to the future, we used a factor analysis to understand the correlation structure within the 155 publisher websites leading to the identification of a more parsimonious set of publisher websites to include in the NRS survey.

A factor analysis is essentially a data reduction technique. The underlying assumption is that there exists a (small) number of unobserved latent factors that account for the correlations amongst a (large) number of observed variables (155 websites). The analysis partitions the variability into independent variables with each subsequent variable explaining a diminishing amount of the variability. We can then look at the correlations between each factor and every one of the original observed variables (155 websites). In an ideal world, each factor would have a perfect (positive or negative) correlation with a different set of publisher websites and zero correlation with the rest. For example, an ideal solution would be 30 factors each having a perfect correlation with a different 15 or 16 of the 155 websites. Then we could pick one website from each factor to include as a fusion hook, knowing that this takes care of all the other perfect correlated websites in the factor. In practice, of course, it is not so clear cut.

The factor analysis was based upon the UKOM/Nielsen panel for June 2010, filtered on all panel members who used at least one of the publisher websites in the month. Table 3.4.1 shows a frequency distribution of the numbers of publisher websites accessed at all in the month.

Table 3.4.1 Number of Websites accessed in a month

1	37%
2	21%
3	13%
4	8%
5	6%
6-10	12%
11+	3%
Average	3.2

UKOM: June 2010

This indicates that there is a long tail of websites with relatively low penetrations and therefore low correlations with any other websites. With this caveat, the outcome of the factor analysis was strong with 26 factors (17% of the original variables) explaining 70% of the variability amongst the 155 websites. This is the result for All Adults.

The first factor explains 12% of the variability and has moderately high correlations in the range 0.3 to 0.6 with an eclectic mix of 48 magazine websites. This is a useful finding because it means we can control a large proportion of magazines effectively and indirectly with a prudent selection of a handful of these websites. All other factors are highly correlated with small groups or individual newspapers or magazines. Separate factor analyses of Men and 16-34s feature a strong motoring magazine website factor and a factor correlating highly with a range of national newspapers, the latter indicating that 16-34s surf a broad spectrum of newspapers.

From these factors, 33 of the highest correlating websites have been chosen for inclusion in the future NRS surveys. As a final check we have calculated the multiple correlation between each of the 155 publisher websites and the combined effect of the 33 fusion hooks. In table 3.4.2 we have segmented the websites according to the size of the multiple correlation with the fusion hooks.

Table 3.4.2 Website Correlation with Fusion Hooks

Segment	Number of Websites	Sites per Person	Multiple Correlation	
			Range	Average
Fusion Hooks	33	2.4	100%-100%	100%
High Correlation	44	0.7	25%-44%	36%
Low Correlation	78	0.1	0%-25%	15%
Total	155	3.2	0%-100%	82%

UKOM: June 2010

Obviously the 33 websites which are the fusion hooks have a perfect correlation with the fusion hooks! In terms of the numbers of sites accessed in a month, these account for the majority of the publisher website activity. Amongst the remaining 122 websites, the average multiple correlation is 25%. We have used this figure to segment the websites. The next 44 websites account for nearly all the remaining website activity and have a reasonably healthy average multiple correlation of 36%. The remainder shows the expected tail, having weak correlations with the fusion hooks but contributing little to the total website activity. The lack of correlation is most likely a small sample issue – in fact a proportion of these websites do not pass the probable reporting sample threshold.

In part this factor analysis has demonstrated a significant challenge to the fusion from extreme sampling error in the tail of low penetration websites. However, it is also encouraging that a relatively small number of website hooks is able to explain a considerable amount of variability among all the publisher websites.

3.5 Constrained vs. Unconstrained Fusion

Although it is not our intention to provide another description of the data fusion process, it is important to explain our decision to adopt a constrained fusion approach. First it must be noted that with both approaches RSMB use an adjusted version of Mahalanobis' Distance to evaluate the hook differences between potential fusion pairings of NRS respondents and UKOM/Nielsen panel members. The distance measurement takes account of correlations between fusion hooks and accounts for scale differences by incorporating importance weights from multivariate analysis of readership behaviour and website usage. In our opinion this rigor is essential when using more than a handful of fusion hooks.

A simple illustration of **Unconstrained Fusion** is shown in Figure 3.5.1.

Figure 3.5.1 Unconstrained Fusion

NRS

Person	Weight	Class	Internet	Title
R_1	1500	ABC1	Heavy	Yes
R_2	1500	C2DE	Light	No

UKOM/Nielsen

Person	Weight	Class	Internet	Website
D_1	500	ABC1	Heavy	Yes
D_2	1500	ABC1	Light	No
D_3	1000	C2DE	Light	No

Fused Sample

Person	Weight	Class	Internet	Title	Website
R_1	1500	ABC1	Heavy	Yes	Yes
R_2	1500	C2DE	Light	No	No

In this approach the NRS is defined to be the recipient survey and the UKOM/Nielsen panel is defined to be the donor survey. The recipient sample and its survey weights form the sample base for the fused dataset. In this respect the fused dataset will exactly replicate all analyses of the original NRS survey data.

For every person in the NRS sample, the fusion algorithm will select the person with the most similar hook profile from the UKOM/Nielsen panel. Panel members can be used more than once or not at all to maximise donor survey effective sample size and UKOM currency preservation.

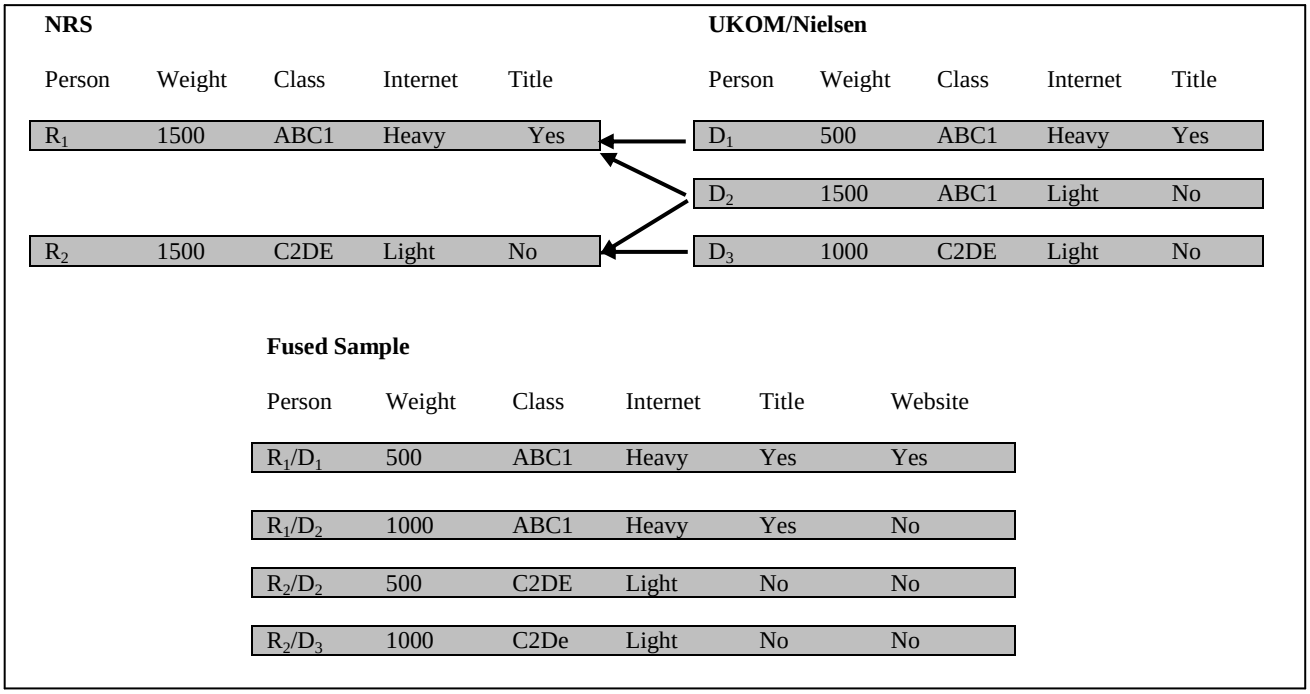
In this example, panel donors D_1 and D_3 match best with NRS recipients R_1 and R_2 respectively. Panel donor D_2 is not used.

The fusion prediction is that 100% of title readers also use the website and this looks like the best estimate.

The NRS readership currency is preserved by definition – 50% read the title. However, the fusion has changed the website reach from 16.7% to 50%. There's a double "hit" from first choosing only one ABC1 donor panel member and second because they have a low panel weight.

The **Constrained Fusion** addresses this currency preservation issue by forcing every panel donor and their survey weight to be used exactly once. A simple illustration is shown is Figure 3.5.2.

Figure 3.5.2 Constrained Fusion



The key to this approach is that fragments of weighted panel donors are fused to fragments of weighted NRS respondents. The constraint ensures that in the new sample, the weight fragments will sum to both the original NRS respondent weights and the original UKOM/Nielsen panel weights. In this way, both measurement currencies will be preserved exactly at the top-line.

As before, panel donors D_1 and D_3 will match best with NRS recipients R_1 and R_2 respectively. In each case the panel donor weight is less than the NRS recipient weight. The shortfall is addressed by splitting each NRS recipient in two to form four "virtual" respondents with smaller weights. This creates two additional NRS recipients who share panel donor D_2 who has not yet been matched.

Compared to the unconstrained fusion, the prediction now is that only 33% rather than 100% of title readers also use the website. This does not look like the best estimate because a heavy internet user has been matched with a light internet user.

As explained, the beneficial trade-off is that the All Adults website currency of 16.7% has been preserved.

But even for currency preservation this is not necessarily a perfect solution. In the fusion we have matched an ABC1 panel member with a C2DE NRS respondent. We can only use one of the social grade classifications and the decision is the NRS version. This results in a distortion of the UKOM website currency for ABC1s from 21% to 33%.

In the real world of larger sample sizes, the distortions are not as severe as these simple examples, and to a degree we can use calibration to close currency gaps. However, the fundamentally disparate sample structures and weighting of the two surveys have led to the unconstrained fusion creating an unacceptably large calibration requirement. In fact for monthly reach, which is a key audience metric, calibration is not possible because monthly reach is a 0/1 variable for each person.

Therefore we have opted for a constrained fusion approach. In addition we have defined Sex and Social Grade to be critical variables for the fusion (i.e. always fuse Men with Men and ABC1 with ABC1). This obviously ensures perfect currency preservation for Men, Women, ABC1s and C2DEs in terms of percentage reach and GRPs (pageviews). It doesn't guarantee perfect currency preservation for higher groups (due to survey universe profile differences) nor for other demographics. However, an acceptable level of currency preservation has been achieved with a modest level of calibration.

Finally, whilst it is true that the constrained fusion approach is not so good at predicting readership/website duplications, the compromise is not significant.

4. Validation

There were two key phases to the validation of the fusion.

The first was a foldover test within the NRS itself, designed to set a benchmark for how good a print/on-line fusion can be. Essentially this was an evaluation of fusion predictions of readership and website duplications, but also demonstrated the benefit of collecting individual website hooks in the NRS survey.

The second was a real fusion between the NRS and UKOM, designed to develop and test procedures to address the practical issues of fusing two disparate surveys and maintaining the UKOM currency. This test was also designed to evaluate the relative merits of unconstrained and constrained fusion.

4.1 Foldover Test

The concept of the foldover test is straight forward. Thirty individual websites have been measured on the NRS survey from Quarter 4 2009. From this single source database we can calculate print average issue readership, website reach and their duplication. For example:

10% of adults read The Daily Mail yesterday

3% of adults used dailymail.co.uk in the last 4 weeks

1% of adults did both (duplication)

32% of dailymail.co.uk users were also average issue readers of The Daily Mail

We then "pretend" that the Quarter 1 2010 NRS survey is the UKOM panel and fuse this onto the Quarter 4 2009 NRS survey, replicating the methodology described in section 2. The Quarter 1 2010 website usage data is carried across so that the fused data contains:

Quarter 4 2009 sample

Quarter 4 2009 demographics

Quarter 4 2009 print average issue readership

Quarter 4 2009 website usage

Fused website usage

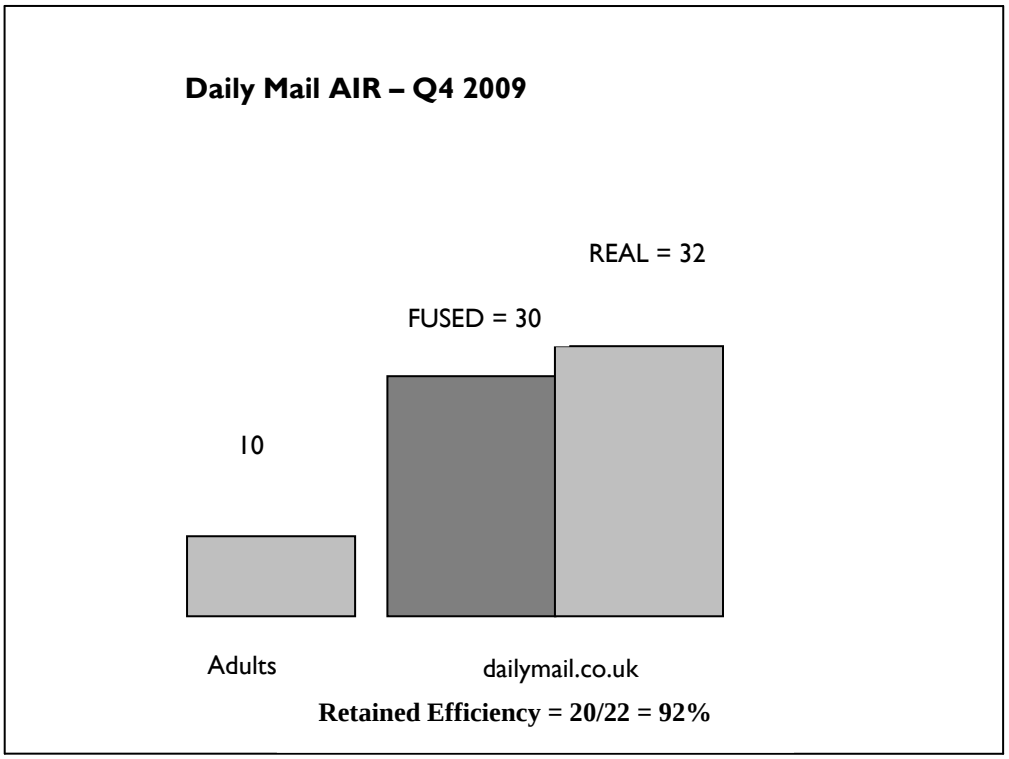
Then we can compare real and fused duplications. Following the examples for the unconstrained fusion method:

% of dailymail.co.uk users who were also average issue readers of The Daily Mail

- **Quarter 4 2009 = 32%**
- **Fused = 30%**

The fusion prediction is close to the real duplication. If the fusion failed completely, which effectively means that donors and recipients were matched at random, then we would expect the fusion prediction of the duplication to be 10% - the same as all adults. The level of "failure" is known as regression-to-the-mean. But we think that's a negative way to look at things. The complement of regression-to-the-mean is known as retained efficiency, the level of "success". Figure 4.1.1 shows the calculation of retained efficiency for our Daily Mail example.

Figure 4.1.1 Retained Efficiency



The retained efficiency measures how well the difference in readership between users and non-users of the website has been retained by the fusion, in this case 92%.

This measurement of retained efficiency is very much a statistical diagnostic. It's fully justified but doesn't necessarily reflect key applications of the fused database. In this respect, the duplication is a second order component of the combined net reach of print and on-line. Website unique reach increment is a more relevant context:

10% of adults read The Daily Mail

2% of adults used dailymail.co.uk but did not read The Daily Mail; the website unique reach increment

12% of adults read The Daily Mail and/or used dailymail.co.uk

In the context of website unique reach increment over the associated print title, for all 30 websites collected by the NRS survey, the average retained efficiency is 93%.

Figures 4.1.2 (newspapers) and 4.1.3 (magazines) shows a scatter plot of the website unique reach increment for each print title and its associate website.

Figure 4.1.2 Newspapers/Websites – Index Net Reach/Publication Reach

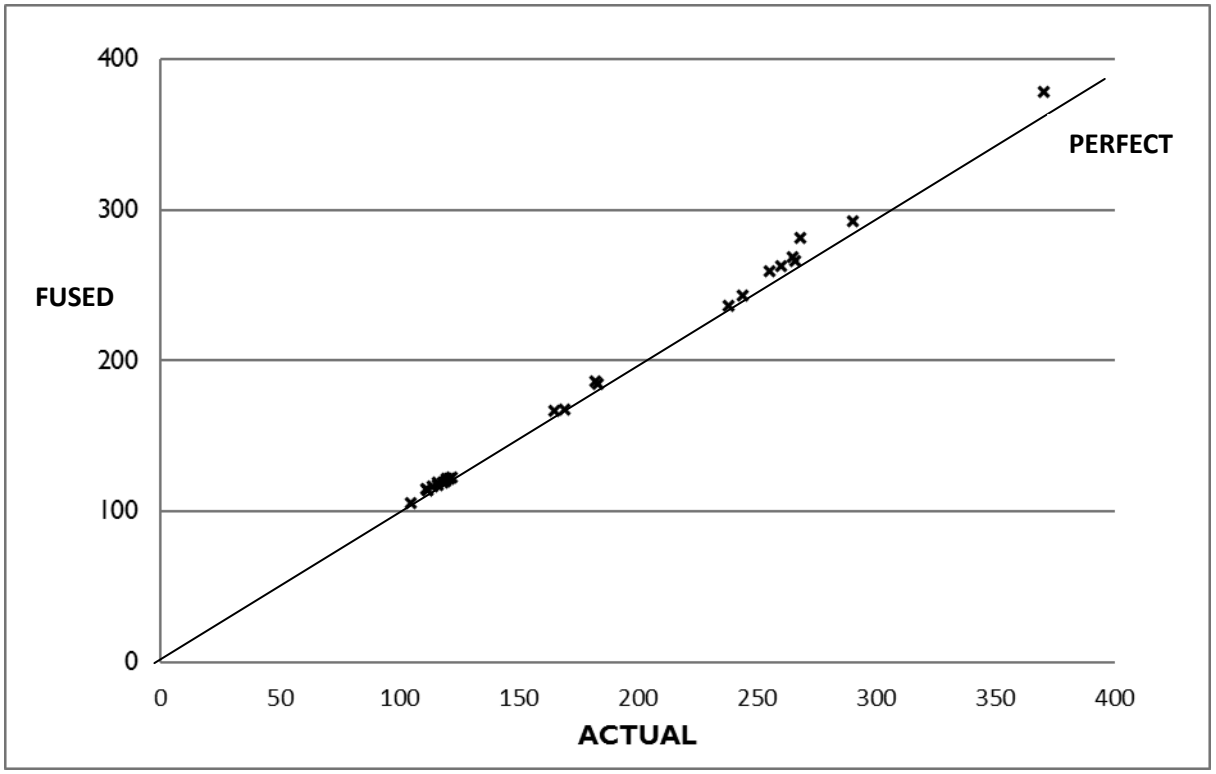
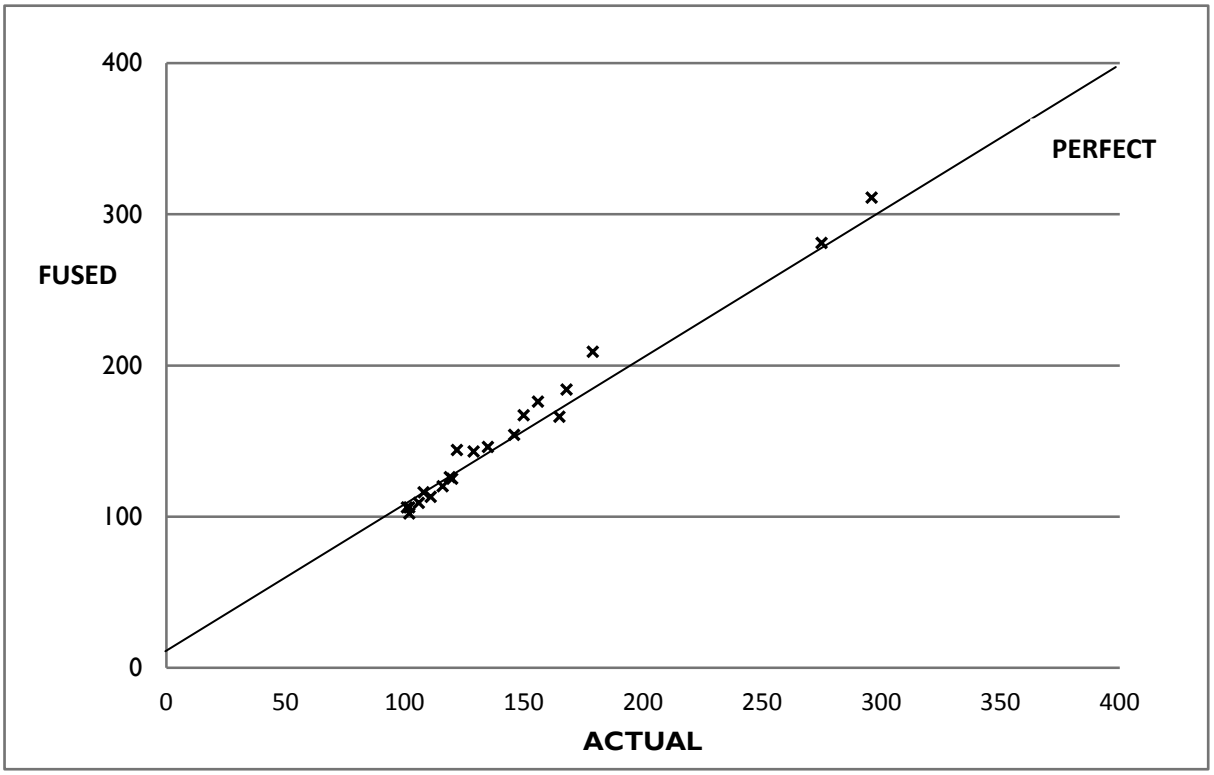


Figure 4.1.3 Magazines/Websites – Index Net Reach/Publication Reach



The scatter plots contain a 45˚ line to indicate what the relationship would be for a perfect solution. The fused and actual indexes are very close, with a slight tendency to over-estimate unique reach.

Key objectives of the foldover test were to evaluate the contribution of the website hooks and to compare the performance of unconstrained fusion with constrained fusion. The different retained efficiencies are summarised in table 4.1.4.

Table 4.1.4 Retained Efficiency

	Unconstrained	Constrained
With Website Hooks	93%	90%
Without Website Hooks	78%	78%

The contribution of the website hooks is significant. As expected, the constrained fusion suffers a loss in retained efficiency, but this is only a small sacrifice against the potential practical benefits.

4.2 NRS/UKOM Fusion

Following the methodologies described in section 2,a test was carried out to fuse UKOM/Nielsen panel data (November 2010)with NRS data for the year ending September 2010. The operation of the fusion was subjected to extensive diagnostic evaluation, but this is not reported in this paper. The key findings relate to the preservation of the UKOM currency and the effectiveness of calibration routines. Note that as the recipient survey, the NRS currency is preserved exactly.

Both unconstrained and constrained fusions were run in each of 22 critical cells defined in terms of gender and 11 geographical regions. Table 4.2.1 shows a top-line comparison of fused and actual website audiences in terms of daily, weekly and monthly reach for All Adults.

Table 4.2.1 Index Fused/Actual

Adults	Daily Reach	Weekly Reach	Monthly Reach
Unconstrained	125	118	114
Constrained	100	100	100

As expected, the constrained fusion preserves the website currency at this top-line level. This will be more or less the case for every website which contributes to these averages.

However, the unconstrained fusion has significantly distorted the UKOM currency. Moreover, this average hides a significant range of currency distortions across websites. Basically, because the unconstrained fusion does not use the UKOM/Nielsen panel weights, it is unable to create a database which is representative of the website weight of usage distributions. No amount of tweaking the fusion set up will resolve this situation.

A possible solution is calibration. Values for individual respondents can be scaled up or down to create the correct overall averages, although the level of calibration required here is probably not acceptable. However, the fundamental problem is that monthly reach cannot be calibrated because it is a yes/no or a 0/1 variable. This effectively rules out unconstrained fusion.

Whilst we know that constrained fusion can preserve the top-line UKOM currencies, this will not necessarily be the case for lower level demographics which do not feature in thedefinitions of the fusion critical cells. Table 4.2.2illustrates the currency distortions for key age and sexual class groups.

Table 4.2.2 Index Fused/Actual

Constrained	Daily Reach	Weekly Reach	Monthly Reach
Adults	100	100	100
Men	100	100	100
Women	100	100	100
Age 15-34	104	104	104
Age 35+	97	97	98
ABC1	113	110	109
C2DE	77	82	85

It may look like the obvious solution is to expand the number of critical cells to embrace region, gender, age and social class. This is not necessarily the right solution:

- Constrained fusion will preserve the UKOM currency in each critical cell. However, differences in demographic profiles between the NRS survey and the Nielsen panel could cause distortion in the All Adults currency.
- If critical cells are too small, then discreteness issues may lead to a poor fusion performance for small websites in terms of print and on-line duplications.
- Critical cells override the hierarchy established within the full set of fusion hooks.

The compromise has been to introduce social class (ABC1 vs. C2DE) but to reduce the geographic component from 11 to 6 regions. The number of critical cells remains around the same at 24. The revised currency comparisons are shown in table 4.2.3.

Table 4.2.3 Index Fused/Actual

Social Class Critical	Daily Reach	Weekly Reach	Monthly Reach
Adults	100	100	100
Men	100	100	100
Women	100	99	100
Age 15-34	102	102	103
Age 35+	98	98	98
ABC1	101	100	100
C2DE	98	98	99

The compromise has worked in that the social class issue is more-or-less resolved. A small currency distortion has been introduced into the weekly reach for Women but this is not important. The results for Age actually show a marginal improvement.

As a final step in the process, the on-line metrics have been calibrated to the gender, age and social class UKOM currencies. The results are presented in table 4.2.4 and at this level do not make much improvement. However, these averages hide more important improvements in individual website audiences.

Table 4.2.4 Index Fused/Actual

Calibrated	Daily Reach	Weekly Reach	Monthly Reach
Adults	100	100	100
Men	100	100	100
Women	100	99	100
Age 15-34	99	98	103
Age 35+	101	101	98
ABC1	100	100	100
C2DE	98	99	99

Our conclusion is that the extent to which the currencies can be controlled is limited by the differences between the NRS and UKOM universe profiles. Of course there is more distortion for individual websites, but overall the results are considered to be fit for purpose.

5. Sense of Data

While the standard fusion diagnostics are an important guide to the success of the fusion, for publishers unfamiliar with the concepts of regression-to-the-mean or retained efficiency, the key is the credibility of the net reach data produced and, unavoidably, the level of duplication between each publication and associated website.

Unfortunately we have no external source of benchmarking or validation.

Many publishers have commissioned their own studies into the net reach of their multi-platform brands, but like the website data collected within the NRS interview, these are based on recall, i.e. what readers tell us about websites they also visit, or what website visitors tell us about publications they also read, the latter usually by means of relatively low response "pop-up" surveys.

There are also a number of studies in the public domain which use recall data to estimate duplication in order to arrive at net reach, for example the Portfolio Audience project for regional newspapers conducted by JICREG[i] and the Financial Times Daily Global Audience Project[ii].

However, as discussed earlier, the difficulty with assessing duplication by recall is that it is likely to over-estimate level of duplication between publication and website.

What makes comparisons even more difficult that the degree to which duplication is over-estimated will vary for different website/publication combinations, depending on the saliency of the website, the proportion of infrequent visitors who are missed by recall and so on.

This makes it difficult to compare even by ranking levels of duplication (leaving aside the absolute levels of duplication). Taking the largest newspaper sites as an example, where we know the fusion has worked well, there may be some hints of consistency with the duplication data collected by recall by both the NRS and (separately) IPA Touch Points Hub Survey.

Table 5.1 Comparison of the proportion of yesterday website visitors also reading the newspaper in print yesterday

	UKOM/NRS fusion	NRS (Jul10-Jun11)	Touch Points 3 Hub (2010)
	%	**%**	**%**
dailymail.co.uk	16.6	21.5	34.5
guardian.co.uk	17.7	24.1	34.3
telegraph.co.uk	11.1	20.5	32.8
thesun.co.uk	30.5	41.3	42.6
independent.co.uk	5.2	15.4	14.6
mirror.co.uk	8.2	30.5	NA
metro.co.uk	11.7	35.8	73.9
timesonline.co.uk	10.8	25.5	NA
ft.com	9.2	27.7	26.3

For instance, The Sun website consistently exhibits more duplication with The Sun newspaper than some of the other newspaper websites do with their print siblings. However these hints are not sustained very far down the list of websites, and for the magazine websites the duplication comparisons are even more various.

Given this lack of comparative data, RSMB have analysed what proportion of the duplications observed in the fused database are better than random.

Table 5.2 Percentage of website/print duplications better than random

	% of duplications better than random	Index on random
All duplications	84	
Newspaper website duplications	91	197
Magazine website duplications (UKOM sample 40-174)	70	269
Magazine website duplications (UKOM sample 175 - 399)	91	219
Magazine website duplications (UKOM sample 400+)	80	163

This shows that for 84% of the website/print combinations duplication was better than random, with an index twice that of random. There is some difference by size of magazine website, not surprisingly given that preserving duplication is expected to be more difficult for the smaller sites.

These outcomes are not perfect, given that we have 16% of combinations for which duplication is no better than random. However, they do provide further evidence that the fusion is working as well as expected to preserve duplications, despite the levels of duplication being lower than those publishers are accustomed to see from recall methodologies.

6. Application of the Fused Database

The key application of the fused database is cross-media print and on-line reach and frequency analysis. The NRS respondent records are preserved in their entirety and supplemented by four on-line metrics which have been fused across from the UKOM/Nielsen panel. Each respondent has the following single source record:

NRS survey i.d.

UKOM/Nielsen panel i.d.

NRS/UKOM weight

NRS demographics

NRS readership data (including recency and frequency)

UKOM monthly pageviews

UKOM monthly reach

UKOM average weekly reach

UKOM average daily reach

The use of the NRS/UKOM composite weight is transparent in the construction of a reach and frequency analysis – it is treated as any other survey weight. Then because the composite weights sum to the original NRS survey weights for each survey respondent, analysis of the fused database will exactly replicate analysis of the original NRS survey. The same principle is true for the on-line data but limited by the extent to which the fusion preserves the UKOM currency.

Because the constrained fusion database is a construct which fragments and increases the number of "respondents", a "respondent" count will not indicate the true analysis sample size. In this respect it is necessary to construct a fusion weight for each "respondent":

$$\text{Fusion Weight} = \frac{\text{NRS/UKOM composite weight}}{\text{Original NRS weight}}$$

Then a summation of these fusion weights will provide a reasonable assessment of the sample count. It is acknowledged that this will not be perfect for analysis filters based upon the website usage.

Finally it has to be recognised that there is significant development required of the NRS bureaux to incorporate these new data and sample structure into their proprietary planning models and systems.

7. Next steps

One of the key requirements of the project is that the results should be consistent and reproducible. With that in mind we are now conducting a second fusion of NRS and UKOM/Nielsen data, using different data periods to those used in the first test. One of the key objectives of this second test is to assess the degree in volatility in the net print/website estimates between fusion updates (and the extent to which this can be attributed to the fusion rather than volatility of the estimates themselves).

Bearing in mind all the practical issues described in this paper, NRS Ltd will also use data from this second test to make an assessment of what the cut-off should be in terms of the sample size of websites to have fused data published. Nielsen publishes estimates for websites with a sample of 40 or more respondents, but NRS Ltd may recommend a higher cut-off, depending on the degree of volatility observed. We have already decided that the fusion should only be carried out on a full 12- month NRS database in order to maximise the sample sizes available (some of the larger print titles have estimates published on a six-month database).

With stakeholder approval, the first release of data will be in early 2012, with quarterly updates, and the possibility of moving to monthly updates in the future.

However, more development work is likely to be needed along the way. The Nielsen panel data in the UK is evolving to a hybrid methodology, and the implications of the way in which this is implemented will require consideration. In addition, publishers are keen to include estimates for other platforms – mobile and tablets in particular – at the earliest opportunity.

8. Conclusions

NRS Ltd does not believe it can arrive at realistic measures of net publication/website reach by recall alone, and is not considering the possibility of introducing passive measurement of publisher websites due to the investment involved. In this context fusion is the best, if not the only, way of providing a currency of net audience estimates for publications/websites. However, fusion is not a panacea.

Many of the specific issues encountered relate to the practicalities of how the respective NRS and Nielsen databases are set up, particularly where they cannot be directly aligned in respect of the home/work samples and social grade profiles.

The focus of the project has been very specifically on reproducing print/website duplication as far as possible. While there are considerable challenges in achieving this, we have no doubt that the commercial value to publishers and agencies of a unified database created in this way is well worth the endeavour.

[i] Regional Newspaper Readership and Integrated Website Audience Data. Dick Dodson, Roger Holland. WWRS 2009.
[ii] http://fttoolkit.co.uk/2011mediakit/index.php