MAKING BEST USE OF BRAND TARGET AUDIENCES

Pete Doe, Nielsen
Dr. James H. Collins, GfK MRI

---

1. Introduction

In the last few years the use of integrated and fused databases has proliferated, enabling planners and media buyers and sellers to go beyond standard demographics to focus on brand targets as well as to understand cross-media effects more precisely.

This greater precision and complexity comes with a need for greater understanding of the capabilities and limitations of these databases. The overarching question we need to answer is: does planning and (potentially buying) using brand targets lead to better business decisions than simply using demographics?

Sub-topics within this broad organizing question include:

a) How predictable are people, their consumer behaviors and attitudes, data from period to period?
b) What range of gains in efficiency might be expected using brand targets?
c) Do different methodological approaches influence the utility of brand targets? For example, how do statistical methods such as data fusion compare with single source data?
d) What analytic approaches best identify media vehicles that truly improve efficiency?

This is a very broad subject and this paper does not attempt to definitively answer every question. Instead, it outlines a framework of key concepts and dimensions important when assessing these issues, presents some empirical data to demonstrate these concepts and finally suggests guidelines for working with brand targets.

It should be noted that this is not a new topic: In November 1970 John Dimling, working at National Association of Broadcasters (NAB), published a paper assessing brand ratings consistency[1]. Its primary conclusion was that product use ratings in a later period were better predicted using demographic ratings than product use ratings from a previous period. This finding was based on 1968/1969 single–source TV diary and product use information and although the world has changed in many ways since then the challenge remains – how can we make best use of brand target ratings? In this paper we demonstrate that gains in media efficiency can be obtained using brand targets and that there are statistical strategies we can employ to help realize these gains.

2. Illustrating the Issues

The best way to illustrate the various issues is with a simple example:

We have a brand target X and two TV shows, each with 10 episodes in both period 1 and period 2 (20 episodes of each in all). Table 1 gives the average rating estimates for period 1:

Table 1

		Period 1	
		Show 1	Show 2
A	Persons 18+	12	2
B	Women 25-54	6	4
C	Brand Target X - Persons 18+	10	3
D	Brand Target X – Women 25-54	7	3
	Index - C/A	83	150
	Index - C/B	167	75
	Index - D/A	58	150
	Index - D/B	117	75

The first question is: what is our target? Is it Persons 18+ in the brand target or Women 25-54 in the brand target? An obvious argument is that if you have a brand target why would you interlace it with a demographic? However, for various reasons (e.g., execution related logistical issues, brand objectives, established practices) brand targets often are interlaced with demographics.

So let us assume that our target is D – the brand target interlaced with Women 25-54. How do we assess the relative value of the programs? There are various considerations:

a) Take the target ratings at face value and determine that Show 1 has a relative value of 7, Show 2 a relative value of 3.
b) Look at the relative strength of the target rating against the total Persons 18+ audience for the show. In this case (Index D/A). In this case Show 1 is weak and Show 2 is strong.
c) Look at the relative strength of the target rating against the total Women 25-54 audience for the show. In this case (Index D/A). In this case Show 2 is weak and Show 1 is strong.

Clearly we need more information - for example, the cost and availability of spots in each of the shows, and perhaps additional qualitative information such as engagement scores or some assessment of program/brand synergy. For simplicity let's assume that the cost of the shows is in proportion to the women 25-54 ratings – i.e. in relative terms, Show 1 costs $6 and Show 2 costs $4. Then, based on these data, the best return on investment in pure GRP terms is to buy Show 1 only – buying 100 Women 25-54 GRPs in Show 1 will deliver 117 Brand Target Women 25-54 GRPs, while the equivalent for Show 2 would only yield 75 GRPs.

So let's proceed on this basis – let's say we buy 100 women 25-54 GRPs in Period 2, all in Show 1, expecting to get our 117 GRPs (leaving aside considerations of reach where a mix of Show 1 and Show 2 would probably be advisable). So what happens in period 2? Table 2 shows this.

<u>Table 2</u>

		Period 1		Period 2		Period 2/1 Index	
		Show 1	Show 2	Show 1	Show 2	Show 1	Show 2
A	Persons 18+	12	2	11.5	2.0	96	100
B	Women 25-54	6	4	5.8	3.8	97	95
C	Brand Target X - Persons 18+	10	3	8.5	4.0	85	133
D	Brand Target X – Women 25-54	7	3	6.0	4.2	86	140
	Index - C/A	83	150	74	200	89	133
	Index - C/B	167	75	147	105	88	140
	Index - D/A	58	150	52	210	89	140
	Index - D/B	117	75	103	111	89	147

The Persons 18+ and Women 25-54 audiences are fairly consistent, but the brand target ratings are less predictable. The brand target by demo index on demo (D/B) has shifted so that, while Show 1 is still positive for the brand target, Show 2 has switched from being negative to positive relative to the demo. Armed with this knowledge we would have been better buying Show 2 than Show 1. In fact, had we looked at Index D/A we would have come to that conclusion from the Period 1 data in this case – though we could find other examples that would support the D/B index approach.

So what has happened here? Why did we choose the wrong show? There are various things we should consider to mitigate the risk of this sort of decision making and these are discussed in the next section.

3. Factors to aid Decision Making

3.1. Audience Consistency

We define consistency as the similarity of results over two periods. Inconsistent results between periods may be caused by real change, statistical fluctuations or a combination of both. This section explores statistical reasons for variation.

AIR/Ratings Consistency
Consider a single event that is consistent in the population between periods 1 and 2, e.g. an Average Issue Readership (AIR) of *p*. We measure the two periods with consistent effective sample sizes *n* in each period.

We would expect the samples both to return the same result i.e. *p*, though we know in practice that there will be some variance around this. The standard error of the observed difference between the two periods is as follows:

$$SE = \sqrt{\frac{2p(1-p)}{n}} \qquad \text{(equation 1)}$$

It is worth noting that in the case of panel measurements, there are other effects on consistency that are obtained from longitudinal measures, both through sample consistency and averaging of measurements over time (e.g. considering a TV show or a website's average use). We can therefore define and include a design effect d:

$$SE = d\sqrt{\frac{2p(1-p)}{n}} \qquad \text{(equation 2)}$$

For individual measures from two periods measured by independent samples, d = 1. In other cases, d can vary depending on a number of factors, such as:

- The difference in time period – leading to less consistency in a panel's composition and perhaps greater real change in the population.
- Clustering effects
- Averaging of results

In Section 4.1 we provide some estimates of d based on empirical data.

A variation measure can be defined as the SE of the difference divided by the expected value p, giving essentially a coefficient of variation:

$$\text{Variation } V = SE/p$$

$$V = d\sqrt{\frac{2p(1-p)}{np^2}} \qquad \text{(equation 3)}$$

Because p values are often fairly small relative to 1 (i.e. media ratings are typically small due to fragmentation) we can approximate and simplify by approximating 1-p as 1: doing this we get

$$V \cong d\sqrt{\frac{2}{np}} \qquad \text{(equation 4)}$$

This shows that the consistency is inversely proportional to the product of the sample size and the size of the vehicle being measured – this can be estimated by the sample size of those reading a magazine or viewing a TV program (or any other simple media exposure). Section 4.1 explores the variability of d for demographics and brand targets.

3.2 Index Consistency

When assessing brand target audiences a common approach is to assess the index of the brand target audience against a benchmark such as Persons 18+ or another suitable demographic. However, indices can fluctuate wildly for small viewing events measured with small samples. So how reliable is an index and how can we determine for a given audience and brand target index how variable and actionable the index is?

Strictly speaking it is not possible to calculate a sampling error on an index of two normally distributed variables (the distribution created is a Cauchy distribution with infinite variance) but there are reasonable ways of assessing the reliability of indices.

One simple way is to assess the standard error of the difference between the benchmark audience and the brand target audience with the assumption that the two measures come from independent samples. Although this assumption is usually not true - the brand target audience is usually a subset of the benchmark audience – the mathematics are much simpler if we assume independence and the difference in results is small, particularly for lower penetration brand targets, and conservative – the correlation between the estimates is ignored and the standard error is larger.

For single vehicles we have:

Benchmark Audience = p_1, effective sample size = n
Target Audience = p_2, effective sample size = m

Let n = cm i.e. target penetration is 1/c
Let $p_2 = kp_1$ i.e. target audience index is 100k

Then the standard error S of the difference between p1 and p2, assuming independence, is:

$$S = \sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}}$$

With manipulation we get:

$$S = \sqrt{\frac{p_1(1+ck) - {p_1}^2(1+ck^2)}{cm}} \qquad \text{(equation 5)}$$

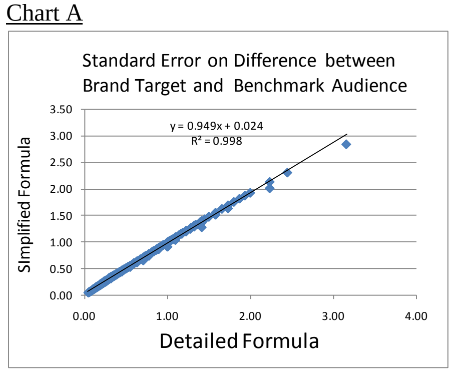
We can use a much simplified version with little loss of accuracy in most scenarios, the exception being cases of higher penetration targets and high rated vehicles, where the assumptions of independence fall down anyway.

Simplified version:

$$S \cong \sqrt{\frac{p_1 k}{m}} = \sqrt{\frac{p_2}{m}} \qquad \text{(equation 6)}$$

The closeness of this approximation to the more detailed version is shown in Chart A below, for a combination of scenarios (330 in total) as follows:

Chart A



Benchmark audience values: 0.1, 0.5, 1, 2, 3, 5
Target Index: 70, 80, 90, 100, 100, 120, 130, 140, 150, 200, 400
Target Penetration (%): 1, 2, 5, 10, 20

Note that we can apply an adjustment to this formula to account for averaged vehicles – e.g. series averages for TV shows – if we have an assessment of the design factor d. In this case we have:

$$S \cong d\sqrt{\frac{p_2}{m}} \qquad \text{(equation 7)}$$

With this, we can assess the index levels that are likely to indicate real differences. If we require the difference between the benchmark audience and the target audience to be significant at the 95% level, using a one tailed test, we get this condition:

$$zd\sqrt{\frac{p_2}{m}} < p_2 - p_1 \qquad \text{for } p_2 > p_1$$

Where:

p_2	=	kp_1
k	=	index/100
m	=	Brand Target Effective Sample Size
d	=	design effect
z	=	Significance Value (e.g. 1.645 for one-tailed test, 95% significance)

Re-stating this and setting the equation to equality to get the minimum index required for significance we have:

$$zd\sqrt{\frac{kp_1}{m}} = p_1\,|k-1| \qquad \text{(equation 8)}$$

We can solve this for k or z, depending on our analysis requirements.

For z we get

$$z = \pm\frac{(k-1)}{dk}\sqrt{kmp_1}$$

Equivalently

$$z = \pm\frac{(k-1)}{dk}\sqrt{mp_2} \qquad \text{(equation 9)}$$

For k we get:

$$k = \frac{2mp_1 + z^2d^2 \pm zd\sqrt{4mp_1 + z^2d^2}}{2mp_1} \qquad \text{(equation 10)}$$

Example: Brand target effective sample size m = 400, Benchmark audience = 1, i.e. p = 0.01
Assume z = 1.645 and d = 0.6, so zd = 1

Then

$$k = \frac{8 + 1 \pm \sqrt{16 + 1}}{8} = 0.61 \text{ or } 1.64$$

This means that if we have an index of less than 61 or greater than 164 the difference is likely to be significant.

Table 3 extends this to other sample sizes and audience levels for indices greater than 100, demonstrating the obvious fact that higher ratings and larger sample sizes are more stable and as a result, smaller indices are more likely to be significant in these cases.

Table 3

	Significant Indices for Brand Targets							
Audience	Brand Target Sample Size							
(AIR/Rating)	100	200	300	400	500	1000	2000	4000
0.1	1192	685	514	427	373	262	200	164
0.3	514	337	275	242	222	177	150	133
0.5	373	262	222	200	186	156	137	125
1	262	200	177	164	156	137	125	117
2	200	164	150	142	137	125	117	112
3	177	150	139	133	129	120	114	110
5	156	137	129	125	122	115	111	107
10	137	125	120	117	115	111	107	105
20	125	117	114	112	111	107	105	104

4. Empirical Validation

We created a series of analyses based on a) U.S. National People Meter (NPM)/GfK MRI fused databases comparing various targets with differing penetrations, and b) single source GfK MRI data. Sections 4.1 and 4.2 assess the fused data and Section 4.3 assesses the GfK MRI single source data.

a) NPM/GfK MRI Fused Data Analysis Dimensions

Periods: September, October and December 2010 and March 2011

Targets:

Target	Brand Target	Average Effective Sample Size	Source
A18+		23161	NPM
A18-49		13199	NPM
F25-54		6600	NPM
A18+	Automotive	1987	NPM/MRI Fusion
A18-49	Automotive	1247	NPM/MRI Fusion
A18+	NPM Automotive	3392	NPM
A18-49	NPM Automotive	2183	NPM
A18+	CPG	1359	NPM/MRI Fusion
F25-54	CPG	379	NPM/MRI Fusion
A18+	Financial	4648	NPM/MRI Fusion
A18-49	Financial	2704	NPM/MRI Fusion

"NPM Automotive" refers to a single-source classification. It is similar but not identical to the Fused Automotive classification. Apart from this, the sources of the brand target data were:

September 2010: NPM 2010 Q3 / GfK MRI Spring 2010 Survey

October and December 2010: NPM 2010 Q4 / GfK MRI Spring 2010 Survey

March 2011: NPM 2011 Q1 / GfK MRI Fall 2010 Survey

This allowed us to include an assessment of the consistency between different fusions, with six different combinations of periods:

September-October, September-December, September-March, October-December, October-March and December-March

Viewing: Average ratings for all national programs on these networks:

ABC	CMT	NICKELODEON
CBS	DISCOVERY	SCIENCE
NBC	ESPN	SPEED
FOX	FOOD NETWORK	SYFY
A&E	HISTORY	THE WEATHER CHANNEL
ADULT SWIM	LIFETIME	TURNER
BBC AMERICA	MILITARY	VH1
CABLE NEWS NETWORK	MTV	

For the consistency analyses we restricted our analyses to programs that aired in all four periods – 483 programs in total.

b) GfK MRI Single Source Data Analysis Dimensions

Periods: Spring 2010 and Fall 2010

Targets: 16 targets were created from 4 demos plus 3 brand targets (the same as used in the fusion analysis) by these 4 demos:

Demos: Adults 18+, Adults 18-49, Adults 18-34, Women 25-54

Brand Targets: Automotive, CPG, Financial

Sample sizes are given here.

	Spring 2010				Fall 2010			
	All	Auto-motive	CPG	Financial	All	Auto-motive	CPG	Financial
Women 25-54	7070	1020	372	1344	6846	995	353	1227
Adults 18-49	14412	2157	1053	2796	13869	2084	1036	2474
Adults 18-34	6435	918	672	1171	6128	884	636	1077
All	26342	3536	1394	5072	25890	3496	1380	4647

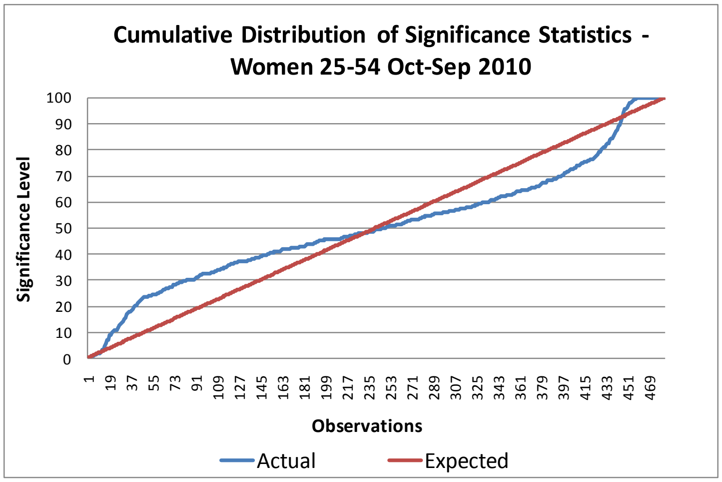
Readership: Average Issue Readership for these publications:

AARP The Magazine	ESPN The Magazine	O Oprah Magazine
Allure	Esquire	Popular Mechanics
Automobile	Martha Stewart Living	New Yorker
Better Homes & Gardens	Maxim	US News & World Rpt.
Consumer Reports	Newsweek	Vogue

4.1 Audience Consistency

We assessed audience consistency by comparing the differences in average program ratings between periods and calculating the basic standard error using equation 1 given in Section 3.1. The difference in ratings divided by this standard error estimate gave a Z-statistic for each program. Plotting the cumulative distribution of significances implied by these Z-scores allows us to assess the extent to which this differs from the expected linear distribution of significances. Chart B below illustrates this, showing that for Women 25-54 there were a larger number of expected observations at the less significant levels – about 90% of observations occurred within significance levels of 20% – 80%. This demonstrates that the program ratings between September and October were more consistent than we would expect to see from two independent samples.

Chart B



Then we estimated the design effect *d* by fitting the distribution of Z-scores given by equation 1 to the expected distribution of the 483 different viewing statistics. This was achieved by finding the least–squares solution for factored Z-scores. In this case the best solution for d was 0.54. Chart C shows that the line obtained by fitting a distribution with d = 0.54 matches the expected distribution closely (correlation = 0.999), except for outliers at the ends of the distribution. These outliers – there were 6% significant at the 99% level – were typically high rating prime time broadcast shows that were finding their feet in the new season. Table 4 shows that these programs increased ratings by 41% on average between September and October (against 9% overall).

Chart C

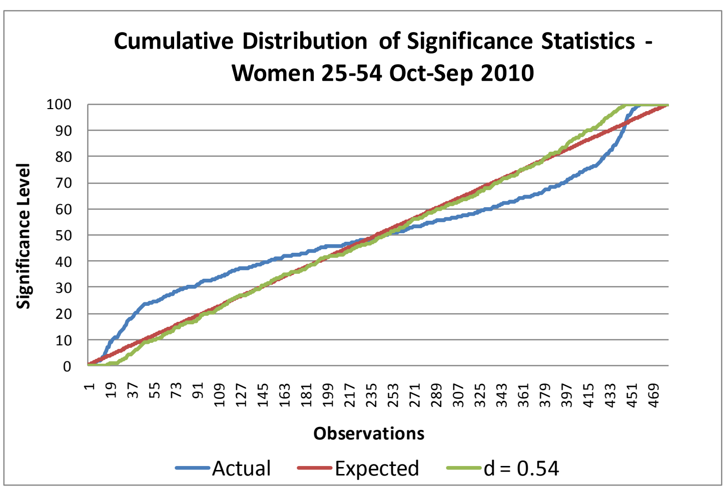


Table 4

	Number of Shows	Average Ratings Sep	Oct	Index
Outliers	27	2.73	3.86	142
All shows	483	0.60	0.65	109

Looking at the various periods and targets we obtained design effects as shown in Table 5. Some key insights are given below (note that the smaller the d value, the more consistent the results compared with two independent samples in the two periods):

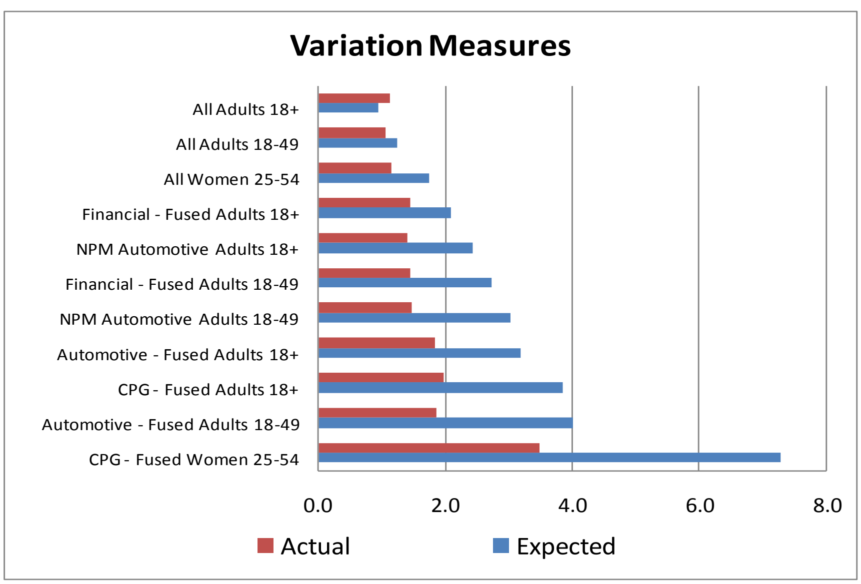
a) Consistency is typically better for shorter intervals, ranging from d = 0.5 for one month to d = 0.7 for 5 or 6 months.
b) Demographics are relatively less consistent than brand targets – presumably because they are, by definition, less targeted. The average brand target design effect was 0.54 compared with 0.76. There is insufficient information here to distinguish between brand targets.
c) Associated with b above, the smallest design effects are seen for brand targets interlaced with demographics.
d) There is no significant difference between the single-source and fused automotive design effects.
e) There is perhaps some evidence of a "new fusion" effect as design effects for consistency involving March fused data show an increase – December-September (3 months) is 0.51, while March–December (also 3 months) is 0.61.
f) Design effects mostly indicate an improvement on the assumption of sample independence (i.e. d<1). For adults 18+, d> 1 for all but October-September. This is related to clustering of persons' viewing in homes – this is more pronounced for all adults than for sub-groups. For this reason Women 25-54 have a lower d value than Adults 18-49 as Women 25-54 are less clustered in homes than Adults 18-49.

Table 5 Design Effect Estimates

		Adults 18+					Adults 18-49				Women 25-54		Average		
Months Apart	Period	All	Finan-cial - Fused	CPG - Fused	Auto-motive - Fused	NPM Auto-motive	All	Finan-cial - Fused	Auto-motive - Fused	NPM Auto-motive	All	CPG - Fused	Un-fused	Fused	Total
1	Oct 10- Sep 10	0.81	0.51	0.47	0.51	0.47	0.64	0.41	0.43	0.40	0.54	0.39	0.57	0.45	0.51
2	Dec 10 - Oct 10	1.23	0.65	0.42	0.52	0.55	0.82	0.46	0.40	0.44	0.62	0.35	0.73	0.47	0.59
3	Dec 10 - Sep 10	1.39	0.61	0.50	0.54	0.55	0.88	0.52	0.43	0.50	0.65	0.46	0.79	0.51	0.64
3	Mar 11 - Dec 10	1.25	0.71	0.58	0.65	0.60	0.91	0.59	0.52	0.50	0.72	0.62	0.80	0.61	0.69
5	Mar 11- Oct 10	1.25	0.86	0.56	0.63	0.64	0.93	0.62	0.52	0.54	0.71	0.54	0.81	0.62	0.71
6	Mar 11- Sep 10	1.27	0.81	0.55	0.62	0.66	0.96	0.59	0.50	0.54	0.72	0.52	0.83	0.60	0.70
	Average	1.20	0.69	0.51	0.58	0.58	0.86	0.53	0.46	0.49	0.66	0.48	0.76	0.54	0.64

Of course, design effects are not the whole story – it is worth noting that the lowest design effect – 0.39 for CPG in October-September, is also associated with the smallest sample size and therefore the least consistent results. Chart D shows the relative variation of the various targets for ratings of 1, using equation 4 in section 3.1, with d values as given in Table 5.

Chart D



In summary, the targets with the least variation inevitably have large sample sizes, but the reduction in variation due to design effect is larger for the smaller groups. What this does not tell us is how these different variation values affect the actionability of the data: although it is clear that if we plan using the CPG Women 25-54 target we will clearly expect less consistency than for the other groups, the key issue is to what extent we can find consistent differences in this group that will deliver improvements over standard demographics.

4.2 Index Consistency

Using the same dataset, we tested the validity of the theory presented in Section 3.2. For each brand target we calculated three Indices of viewing for each program:

- Brand Target 18+ indexed on Adults 18+
- Brand Target by Demo indexed on Demo
- Brand Target by Demo indexed on Adults 18+

Then we applied equation 9 from section 3.2 to these Indices, using the estimated design effects that are given in Section 4.1, to test the hypothesis that the index is significantly different from 100.

Example: Fused Financial Target Adults 18+ indexed on Adults 18+, September 2010.

Program X Brand Target Rating (Adults 18+) = 0.33, Adults 18+ Rating = 0.52, Index = 63

We have:

$$z = \pm \frac{(k-1)}{dk}\sqrt{mp_2} \qquad \text{(equation 9)}$$

Effective Sample = m = 4700, Design Effect d = 0.51, k = Index/100 = 0.63, p_2 = 0.33/100 = 0.0033

This gives us z = 4.4 and we conclude that this index is significantly different from 100 at the 95% level. In fact, the October index was 61, supporting this conclusion and meaning that the data are actionable.

Appendix A presents results for all combinations of index types, periods and targets, showing the percentage of programs in these four groups:

Group 1: Index in Period 1 is significantly different from 100 and Period 2 Index is above or below 100 in line with Period 1

Group 2: Index in Period 1 is significantly different from 100 but Period 2 Index is inconsistent with Period 1 – i.e. moves from above 100 to below or vice versa

Group 3: Index in Period 1 is not significantly different from 100 and Period 2 Index is above or below 100 in line with Period 1

Group 4: Index in Period 1 is not significantly different from 100 and Period 2 Index is inconsistent with Period 1 – i.e. moves from above 100 to below or vice versa

These groupings give us an indication of actionability – how many indices remain consistently above or below 100 and how many of those are predicted via the application of the theory.

In each case, "significantly different" refers to a 95% confidence interval.

Key findings:

a) Overall about a third of programs fall into Group 1 – they index significantly different from 100 in one period and are consistent in the next period. This percentage varies from 40% to 18% - the lower penetration CPG category being the least actionable as expected, although there is a clear improvement in performance for this group when the index uses CPG x women 25-54 indexed on all adults 18+.

Table 6: % of Programs by Group

		Group 1	Group 2	Group 3	Group 4
		Significant	Significant	Not Sig	Not Sig
Target	Index Type	Consistent	Inconsistent	Consistent	Inconsistent
Automotive - Unfused	Target 18+ on All 18+	40	1	45	14
Financial - Fused	Target 18-49 on All 18-49	35	1	45	18
Automotive - Unfused	Target 18-49 on All 18-49	35	2	46	18
Financial - Fused	Target 18+ on All 18+	34	1	48	17
Automotive - fused	Target 18-49 on All 18-49	33	2	49	16
Automotive - fused	Target 18+ on All 18+	32	2	51	15
Automotive - fused	Target 18-49 on All 18+	30	2	46	22
CPG - Fused	Target W25-54 on All 18+	29	2	46	23
Financial - Fused	Target 18-49 on All 18+	29	2	46	23
Automotive - Unfused	Target 18-49 on All 18+	28	2	46	24
CPG - Fused	Target 18+ on All 18+	25	3	43	30
CPG - Fused	Target W25-54 on All W25-54	18	3	44	36
Average	Target 18+ on All 18+	33	2	47	19
Average	Target Demo on All Demo	30	2	46	22
Average	Target Demo on All 18+	29	2	46	23
Total	All Options	31	2	46	20

Although the largest group is Group 3 these programs account for only a small percentage of the ratings difference as the example in Table 7.1 demonstrates. With Group 1 at 86% in September and 82% in October, the significance test clearly identifies the majority of the ratings differences that follow through into period 2.

Table 7.1 Financial Target Ratings Differences by Significance/Consistency Group

		% of total rating differences between Target 18+ and All 18+	
	% of Programs	Sep	Oct
Group 1	42	86	82
Group 2	1	1	0
Group 3	42	11	14
Group 4	15	2	4

b) Overall the data suggests that equation 9 works well. We would expect 5% of significant differences to be inconsistent (i.e., Group 1 would account for 95% of Group 1 and Group 2 programs). This is what happens (See Table 7.2), though the consistency tails off slightly for the six month interval.

Table 7.2 – Consistency of Significant Index Differences

	Sep/ Oct	Oct/ Dec	Sep/ Dec	Dec/ Mar	Oct/ Mar	Sep/ Mar	Average
Interval (Months)	1	2	3	3	5	6	3.3
% significant and consistent	95	96	95	94	94	92	95

There is some variation by Target and Index type with the CPG target being the least consistent – perhaps because it has the smallest sample size and perhaps because there are fewer actionable programs as a result, leading to more inherent variability in the Group 1 and 2 figures.

Table 7.3 Group 2 as a % of Groups 1 and 2

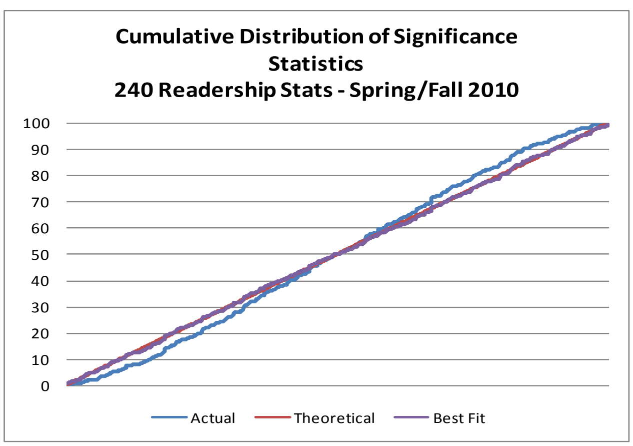
Target	Index Type	% Significant that are consistent
Financial - Fused	Target 18+ on All 18+	98
Automotive - Unfused	Target 18+ on All 18+	97
Financial - Fused	Target 18-49 on All 18-49	97
Automotive - Unfused	Target 18-49 on All 18-49	95
Automotive - fused	Target 18+ on All 18+	95
Automotive - fused	Target 18-49 on All 18-49	94
Automotive - fused	Target 18-49 on All 18+	94
CPG - Fused	Target W25-54 on All 18+	93
Financial - Fused	Target 18-49 on All 18+	93
Automotive - Unfused	Target 18-49 on All 18+	93
CPG - Fused	Target 18+ on All 18+	89
CPG - Fused	Target W25-54 on All W25-54	86
Average	Target 18+ on All 18+	95
Average	Target Demo on All Demo	94
Average	Target Demo on All 18+	93
Total	All Options	95

4.3 GfK MRI Data

4.3.1 Readership Consistency

Readership values for the 15 publications across the sixteen targets varied from 0.3% to 32%. When we examine the distribution of differences between the Spring 2010 and Fall 2010 data we obtain close correspondence with the predicted distribution given by equation 2 as shown in Chart E. Adjusting the distribution using a least squares method gives us a correlation of 1.000, based on a design effect of 1.22. This is equivalent to an effective sample size multiplier of 0.67. This reflects the effect of the survey sample design which features a geographically disproportional structure.

Chart E



4.3.2 Index Consistency

We calculated indices on adults 18+ for each magazine for each demo by brand target combination – 180 statistics in total. We then applied equation 9 to the indices to assess significance and count the number falling into each of the four groups described in Section 4.2. Table 8 shows the aggregated results. The percentage of significant results at the 95% level is slightly lower than for the viewing data discussed in Section 4.2 (25% vs 33%). This is to be expected given the smaller sample sizes and in addition we would not expect the same results anyway given the different media and methods. Table 8 also shows that the percentage of group 1 and 2 that is found in group 2 (i.e. what percentage of significant indices in Spring were consistent in direction in Fall) does not equal the significance level. However they are within the range we might expect given the relatively small number of observations.

Table 8

	Signi-ficant?	Consis-tent?	Significance Level 95%	90%	85%	80%	75%	70%
% of Indices								
Group 1	Y	Y	22	31	39	46	49	54
Group 2	Y	N	3	3	4	7	11	12
Group 3	N	Y	51	43	34	28	24	19
Group 4	N	N	24	23	22	19	16	14
% of Significant in Group 2			89	90	90	86	82	82
% of Difference in AIR								
Group 1	Y	Y	47	57	66	71	73	78
Group 2	Y	N	6	7	9	11	14	15
Group 3	N	Y	35	25	16	11	9	4
Group 4	N	N	12	11	9	7	5	3

5. Alternative Planning Scenarios

In this section we explore whether planning using brand targets can lead to better delivered audiences for the target. We assessed the four National People Meter brand targets (three from the fusion, one single source) interlaced with the target demographic.

As a benchmark we selected the top 100 programs from the 483 in the analysis database using the index of the demo on Persons 18+ as the selection criterion. We then looked at the GRP delivery in the subsequent period for the brand target and the demographic. Second, we selected programs based on the significance estimator given in equation 9, taking the 100 most positive programs in terms of significance of index. We did this in two ways, first using brand target by demo indexed on demo and second, brand target by demo indexed on Persons 18+. We then looked at the GRP delivery in the subsequent period for the brand target and the demographic and compared this with the benchmark.

These analyses were conducted for each of the six intervals available to us. Appendix B gives the full results.

Key findings:

a) Using brand target based program selection delivered significant gains. Overall gains in efficiency (defined as the improvement in the index of brand target ratings on demo) ranged from 19% to 44%. Table 9.1 averages the performance of each brand target across the six intervals. Indexing on demo gave better gains than indexing on Persons 18+.

Table 9.1: Delivery Gains achieved by Planning using Brand Targets

	Planning using Demo GRP Delivery			Planning Using Brand Target x Demo GRP Delivery			Indices		
	Demo	Brand Target	Index	Demo	Brand Target	Index	Demo	Brand Target	Index
1. Brand Target by Demo indexed on Demo for Planning									
Financial - Fused	34	24	72	31	31	97	94	127	135
Automotive	34	24	71	32	33	101	96	138	143
Automotive - Fused	34	22	67	37	35	97	109	158	144
CPG- Fused	65	91	141	60	101	168	92	110	120
Average	41	40	98	40	50	124	97	123	127
2. Brand Target by Demo indexed on 18+ for Planning									
Financial - Fused	34	24	72	34	30	87	102	124	122
Automotive	34	24	71	37	35	95	109	146	134
Automotive - Fused	34	22	67	35	32	92	104	143	138
CPG- Fused	65	91	141	66	111	167	103	122	119
Average	41	40	98	43	52	121	104	129	124

Note: Demo benchmark is adults 18-49, except for CPG brand which uses women 25-54

b) The gains were fairly consistent over the intervals, except for longer-term planning of the smallest brand target - the CPG brand with programs selected in September or October 2010 for March 2011.

Table 9.2: Delivery Indices by Brand Target and Interval

	Sep-Oct	Oct-Dec	Sep-Dec	Dec-Mar	Oct-Mar	Sep-Mar	
Interval (Months)	1	2	3	3	5	6	Average
1. Brand Target by Demo indexed on Demo for Planning							
Financial - Fused	138	135	131	134	134	140	135
Automotive	142	145	154	138	138	138	143
Automotive - Fused	158	143	149	136	143	141	144
CPG- Fused	124	132	124	123	72	67	120
Average	141	139	140	133	122	121	136
2. Brand Target by Demo indexed on 18+ for Planning							
Financial - Fused	122	118	117	122	114	139	122
Automotive	132	133	141	128	131	138	134
Automotive - Fused	147	139	140	135	130	141	138
CPG- Fused	119	126	117	121	127	67	119
Average	130	129	129	127	126	122	128

6. Conclusion

Brand Targeting can deliver significant gains over standard demographics. However, there can be consistency issues for smaller sample sizes and the increased precision of brand targeting can sometimes be offset by the variability associated with the audiences. This paper outlines statistical approaches that help mitigate this risk and deliver improved efficiency in planning and execution. There is more work that could be done to investigate other brands, demos, and media, but the real examples examined here fit the theoretical framework outlined in Section 3 quite well, suggesting that the reliability estimation processes outlined here and the strategy for selecting media based on these estimates are valid and may be more widely applicable.

Appendix A: Actionability Statistics

Appendix A presents actionability statistics based on brand target indices, evaluating the percentage of 483 analyzed programs in terms of significant Indices (brand target versus demo, using equation 9 in Section 3.2) for one period, and the subsequent period's outcomes in terms of consistency.

Target	Index Type	Sig 95%	Consis-tent	Sep/Oct	Oct/Dec	Sep/Dec	Dec/Mar	Oct/Mar	Sep/Mar	Average
Financial - Fused	Target 18+ on All 18+	Y	Y	42	40	36	35	27	25	34
Financial - Fused		Y	N	1	0	0	1	0	1	1
Financial - Fused		N	Y	42	45	47	43	52	56	48
Financial - Fused		N	N	15	14	16	20	21	18	17
Financial - Fused	Target 18-49 on All 18-49	Y	Y	46	42	38	31	31	25	35
Financial - Fused		Y	N	2	1	1	1	2	0	1
Financial - Fused		N	Y	38	46	45	46	42	53	45
Financial - Fused		N	N	14	11	16	23	25	21	18
Financial - Fused	Target 18-49 on All 18+	Y	Y	37	39	33	20	28	17	29
Financial - Fused		Y	N	3	2	2	1	3	1	2
Financial - Fused		N	Y	43	45	46	48	48	49	46
Financial - Fused		N	N	17	14	19	31	21	33	23
Automotive - fused	Target 18+ on All 18+	Y	Y	36	40	35	30	32	21	32
Automotive - fused		Y	N	3	2	2	2	1	0	2
Automotive - fused		N	Y	45	49	47	53	53	60	51
Automotive - fused		N	N	16	10	16	14	14	19	15
Automotive - fused	Target 18-49 on All 18-49	Y	Y	37	45	38	21	36	20	33
Automotive - fused		Y	N	4	2	2	2	2	1	2
Automotive - fused		N	Y	45	45	44	58	47	58	49
Automotive - fused		N	N	14	9	15	19	15	21	16
Automotive - fused	Target 18-49 on All 18+	Y	Y	37	39	34	24	29	18	30
Automotive - fused		Y	N	3	2	2	1	2	1	2
Automotive - fused		N	Y	43	45	46	47	48	49	46
Automotive - fused		N	N	17	14	18	28	21	31	22
Automotive - Unfused	Target 18+ on All 18+	Y	Y	45	43	40	48	37	27	40
Automotive - Unfused		Y	N	1	2	1	2	1	1	1
Automotive - Unfused		N	Y	41	46	45	38	46	52	45
Automotive - Unfused		N	N	13	10	13	12	16	20	14
Automotive - Unfused	Target 18-49 on All 18-49	Y	Y	42	39	37	34	31	25	35
Automotive - Unfused		Y	N	2	2	1	1	3	1	2
Automotive - Unfused		N	Y	41	44	45	47	47	53	46
Automotive - Unfused		N	N	15	15	18	18	18	21	18
Automotive - Unfused	Target 18-49 on All 18+	Y	Y	36	38	32	19	26	16	28
Automotive - Unfused		Y	N	3	2	2	1	3	1	2
Automotive - Unfused		N	Y	43	45	47	45	48	46	46
Automotive - Unfused		N	N	18	16	19	35	23	36	24
CPG - Fused	Target 18+ on All 18+	Y	Y	28	36	24	23	27	10	25
CPG - Fused		Y	N	1	1	3	2	2	8	3
CPG - Fused		N	Y	41	41	45	49	47	34	43
CPG - Fused		N	N	30	22	28	26	25	48	30
CPG - Fused	Target W25-54 on All W25-54	Y	Y	29	33	24	1	17	4	18
CPG - Fused		Y	N	4	3	3	2	4	2	3
CPG - Fused		N	Y	44	45	49	38	51	34	44
CPG - Fused		N	N	23	19	23	59	29	60	36
CPG - Fused	Target W25-54 on All 18+	Y	Y	37	38	33	21	28	17	29
CPG - Fused		Y	N	3	2	2	1	3	1	2
CPG - Fused		N	Y	43	45	46	46	48	48	46
CPG - Fused		N	N	18	15	19	31	22	34	23
Average	Target 18+ on All 18+	Y	Y	38	40	34	34	31	21	33
Average		Y	N	1	1	2	2	1	2	2
Average		N	Y	42	45	46	46	49	50	47
Average		N	N	18	14	18	18	19	26	19
Average	Target Demo on All Demo	Y	Y	38	40	34	22	29	18	30
Average		Y	N	3	2	2	1	3	1	2
Average		N	Y	42	45	46	47	47	50	46
Average		N	N	17	14	18	30	22	31	22
Average	Target Demo on All 18+	Y	Y	37	39	33	21	28	17	29
Average		Y	N	3	2	2	1	3	1	2
Average		N	Y	43	45	46	46	48	48	46
Average		N	N	18	15	19	31	22	34	23
Total	All Options	Y	Y	38	40	34	28	30	20	31
Total		Y	N	2	2	2	2	2	2	2
Total		N	Y	42	45	46	47	48	50	46
Total		N	N	18	14	18	24	20	29	20

Appendix B: Results of Alternative Planning Approaches

This Appendix presents GRP's achieved in a period based on planning from a previous period, across six different intervals, ranging from 1 month to 6 months. Alternative planning approaches tested were standard demo indexing, brand target by demo indexed on demo, and brand target indexed on Persons 18+.

	Planning using A18-49			Planning Using Brand Target x Demo indexed on Demo			Indices		
	A1849	Financial - Fused x A18-49	Index	A1849	Financial - Fused x A18-49	Index	A1849	Financial - Fused x A18-49	Index
Sep-Oct	35	25	73	33	33	101	93	129	138
Oct-Dec	36	27	75	38	38	101	106	142	135
Sep-Dec	34	25	73	30	29	96	88	116	131
Dec-Mar	33	23	69	34	32	93	104	139	134
Oct-Mar	32	23	70	24	23	94	75	101	134
Sep-Mar	31	21	69	29	28	96	95	132	140
Average	34	24	72	31	31	97	94	127	135

	Planning using A18-49			Planning Using Brand Target x Demo indexed on 18+			Indices		
	A1849	Financial - Fused x A18-49	Index	A1849	Financial - Fused x A18-49	Index	A1849	Financial - Fused x A18-49	Index
Sep-Oct	35	25	73	35	32	89	101	124	122
Oct-Dec	36	27	75	41	36	88	114	134	118
Sep-Dec	34	25	73	31	27	86	91	107	117
Dec-Mar	33	23	69	38	32	84	113	138	122
Oct-Mar	32	23	70	30	24	80	94	106	114
Sep-Mar	31	21	69	29	28	96	95	132	139
Average	34	24	72	34	30	87	102	124	122

	Planning using A18-49			Planning Using Brand Target x Demo indexed on Demo			Indices		
	A1849	Auto-motive x A1849	Index	A1849	Auto-motive x A1849	Index	A1849	Auto-motive x A1849	Index
Sep-Oct	35	26	73	42	44	104	121	171	142
Oct-Dec	36	26	72	31	33	104	87	126	145
Sep-Dec	34	23	66	36	37	102	106	163	154
Dec-Mar	33	24	72	31	31	99	94	130	138
Oct-Mar	32	23	72	24	24	100	75	104	138
Sep-Mar	31	22	70	29	28	96	95	131	138
Average	34	24	71	32	33	101	96	138	143

	Planning using A18-49			Planning Using Brand Target x Demo indexed on 18+			Indices		
	A1849	Auto-motive x A1849	Index	A1849	Auto-motive x A1849	Index	A1849	Auto-motive x A1849	Index
Sep-Oct	35	26	73	42	40	97	120	157	132
Oct-Dec	36	26	72	40	38	96	111	149	133
Sep-Dec	34	23	66	36	34	94	107	152	141
Dec-Mar	33	24	72	37	34	92	112	144	128
Oct-Mar	32	23	72	35	33	95	107	140	131
Sep-Mar	31	22	70	29	28	96	95	131	138
Average	34	24	71	37	35	95	109	146	134

	Planning using A18-49			Planning Using Brand Target x Demo indexed on Demo			Indices		
	A1849	Fused Auto-motive x A1849	Index	A1849	Fused Auto-motive x A1849	Index	A1849	Fused Auto-motive x A1849	Index
Sep-Oct	35	22	62	48	47	98	138	219	158
Oct-Dec	36	24	66	40	38	95	111	159	143
Sep-Dec	34	21	62	41	38	92	122	182	149
Dec-Mar	33	24	72	37	36	97	111	152	136
Oct-Mar	32	23	72	24	25	102	75	107	143
Sep-Mar	31	22	70	29	29	98	95	134	141
Average	34	22	67	37	35	97	109	158	144

	Planning using A18-49			Planning Using Brand Target x Demo indexed on 18+			Indices		
	A1849	Fused Auto-motive x A1849	Index	A1849	Fused Auto-motive x A1849	Index	A1849	Fused Auto-motive x A1849	Index
Sep-Oct	35	22	62	47	43	91	135	198	147
Oct-Dec	36	24	66	34	32	92	96	133	139
Sep-Dec	34	21	62	42	36	86	124	173	140
Dec-Mar	33	24	72	32	31	97	97	131	135
Oct-Mar	32	23	72	24	23	93	75	97	130
Sep-Mar	31	22	70	29	29	98	95	134	141
Average	34	22	67	35	32	92	104	143	138

	Planning using W25-54			Planning Using Brand Target x Demo indexed on Demo			Indices		
	W25-54	CPG - Fused x W25-54	Index	W25-54	CPG - Fused x W25-54	Index	W25-54	CPG - Fused x W25-54	Index
Sep-Oct	72	108	149	76	141	185	105	131	124
Oct-Dec	67	107	160	74	156	210	111	146	132
Sep-Dec	65	104	160	74	147	199	114	142	124
Dec-Mar	65	78	120	69	102	148	106	131	123
Oct-Mar	61	76	125	28	25	90	46	33	72
Sep-Mar	58	73	126	37	31	84	64	43	67
Average	65	91	141	60	101	168	92	110	120

	Planning using W25-54			Planning Using Brand Target x Demo indexed on 18+			Indices		
	W25-54	CPG - Fused x W25-54	Index	W25-54	CPG - Fused x W25-54	Index	W25-54	CPG - Fused x W25-54	Index
Sep-Oct	72	108	149	87	154	177	120	143	119
Oct-Dec	67	107	160	71	141	200	105	132	126
Sep-Dec	65	104	160	80	150	188	123	145	117
Dec-Mar	65	78	120	69	100	145	106	128	121
Oct-Mar	61	76	125	55	88	160	91	116	127
Sep-Mar	58	73	126	37	31	84	64	43	67
Average	65	91	141	66	111	167	103	122	119

Reference

1. "But What Per Cent of the Cookies Do You Cover?" John Dimling, NAB, ARF 16th Annual Conference, 1970