AN EXAMINATION OF DIFFERENTIAL WEIGHTING OF SUCCESSIVE WAVES OF MAGAZINE READERSHIP ESTIMATES

Stuart Gray, BBDO
Julian Baim & Martin Frankel, MRI

Combining successive waves of magazine readership data, or creating doublebases have been common practices for the purpose of providing more stable data than would be produced by a smaller sample of data. An MRI doublebase consists of two studies of approximately 20,000 adults each, and therefore about 40,000 adults. Each study is based on combining two survey waves of approximately 10,000 adults each.

However, the combining of samples has traditionally been based on *equal* weighting—which has to be considered arbitrary.

The differences observed in two successive report periods of readership data may be partially due to sampling variation, but may be partly due to actual trends in readership for at least some magazines.

Our hypothesis is therefore that combining data that gives more weight to the more recent report period may produce averages that are more predictive of the next set of readership estimates. If so, this would represent a very useful finding for media planners and buyers, since they generally look at readership data for the purpose of estimating a future schedule. If it turned out that 50/50 weighting could not be improved upon, this would serve to support the current practice of doublebase.

In order to test that hypothesis, successive periods of readership estimates were averaged based on various weighting schemes, such as 40/60 and 20/80 (in comparison to the 50/50 of "doublebase") so as to determine the ideal weight to place on the more recent data with the purpose of best predicting the subsequent report.

Historical MRI data (waves 27 through 40, representing March 1992 through March 1999) were analyzed to perform the predictions retrospectively, with the following findings.

Overall, 20/80 weighting proved best, with 5% improvement over 50/50 in predicting the next period. 40/60 weighting produced 3% improvement over doublebase (see Table 1). Since these reflected only very modest improvements, we then hypothesized that magazines with more significant changes from one period to the next would benefit more from giving more weight to the more recent data.

Therefore, for each data set we divided the magazines between those manifesting a z score[1] of 2 or more in their change in average readers and those with a z score less than 2. In the case of the former group, the improvement produced by 20/80 weighting over 50/50 was 12% while the improvement for the latter group was only 3%. 40/60 weighting again proved to be better than 50/50, but not as good as 20/80 for either group of books (see Tables 2 and 3).

We also investigated the consistency of our findings by examining each of the five sets of forecasts separately. In other words, we looked at the results of different weighting for using waves 27/28 and 29/30 to predict 31/32, then 29/30 and 31/32 to predict 33/34, etc. In three of the five data sets, there was clearly improvement for 20/80 over 50/50 in the z = 2+ group, ranging from 19% to 27%! while in the other two periods there was virtually no difference between 20/80 and 50/50 (only 1% difference in each of these remaining two comparisons) (see Table 4).

On average, there were there were 21 magazines that exhibited report to report to changes with z scores of 2 or more, ranging from 15 to 27 over the five periods examined. There was an average of 148 magazines with "change" z scores of less than 2, ranging from 134 to 157 over the five periods examined.

Since the presumption based on the analysis to this point was that it is appropriate to give more weight to more recent data when magazines are trending significantly, an obvious ensuing analysis was to determine whether putting *all* the weight on the most recent data produces even further improvement over 20/80.

[1] The z-score was calculated using the standard formula for difference of proportions between two independent simple random samples. The z-score was then multiplied by 2 to account for the design effect of MRI's stratified area probability sample.

The results of that analysis demonstrated that putting all the weight on the most recent data (0/100), although better than 50/50, was not as good as 20/80 (see Table 5).

We theorize that this results from the utilization of additional sample. In other words, when 0/100 is used, only sample from the most recent waves of data are applied, while 20/80 brings to bear a virtual doubling of sample size while still allowing the more recent data to exert more influence.

These results seem consistent with observations made by Skrapits and Appel in Vancouver, when they concluded that "when one year circulation changes are large enough to produce readership changes that can be reliably measured, the two are correlated with each other [and] as the confidence ratio increases, so does the strength of the correlation."

Finally, we would point out that all of the above analyses were based on total adult readers, with no attempt to examine demographics as variables, of for that matter different types of magazines. It is our feeling that there are probably insufficient observations to draw conclusions about these variables, but we would welcome ideas for research along those lines.

Table 1: Average Relative Error

(All Measured Magazines)

	50/50	40/60	20/80
Waves			
27-32	9.7%	9.5%	9.7%
29-34	11.3%	10.8%	10.2%
31-36	9.0%	8.8%	8.6%
33-38	10.2%	10.1%	10.1%
35-40	11.1%	10.7%	10.2%
Average	**10.3%**	**10.0%**	**9.8%**

Read this table as follows: When readership estimates for waves 27&28 were combined with waves 29&30 on a 50/50 basis (Doublebase) to predict waves 31&32, the average relative error for all measured magazines was 9.7%. The second row reflects analogous data for waves 29&30 and 31&32 predciting 33&34. The bottom line reflects the average of all 5 sets of data. Thus, an average relative error of 9.8% for 20/80 weighting represents a 5% improvement over the 10.3% error for the normal doublebase weighting of 50/50.

Source: MRI, Waves 27 through 40

Table 2: Average Relative Error

(Magazines with z scores of 2+)

	50/50	40/60	20/80
Waves			
27-32	10.1%	9.2%	10.1%
29-34	15.3%	14.0%	12.3%
31-36	13.4%	12.5%	11.3%
33-38	9.9%	9.3%	10.0%
35-40	13.7%	11.8%	9.9%
Average	**12.5%**	**11.4%**	**10.7%**

Read this table as follows: When readership estimates for waves 27&28 were combined with waves 29&30 on a 50/50 basis (Doublebase) to predict waves 31&32, the average relative error (for all magazines whose readership estimate change from waves 27&28 to waves 29&30 reflected a z score of 2 or more) was 10.1%. For these magazines, the average relative error of 10.7% for 20/80 weighting represents a 12% improvement over the 12.5% error for 50/50.

Source: MRI, Waves 27 through 40

Table 3: Average Relative Error

(Magazines with z scores of >2)

Waves	50/50	40/60	20/80
27-32	9.6%	9.6%	9.6%
29-34	10.6%	10.2%	9.7%
31-36	8.5%	8.4%	8.3%
33-38	10.3%	10.2%	10.2%
35-40	10.7%	10.5%	10.3%
Average	**9.9%**	**9.8%**	**9.6%**

Read this table as follows: When readership estimates for waves 27&28 were combined with waves 29&30 on a 50/50 basis (Doublebase) to predict waves 31&32, the average relative error (for all magazines whose readership estimate change from waves 27&28 to waves 29&30 reflected a z score of less than 2) was 9.6%. For these magazines, the average relative error of 9.6% for 20/80 weighting represents a 3% improvement over the 9.9% error for 50/50.

Source: MRI, Waves 27 through 40

Table 4: Year by Year Improvement

(Magazines with z scores of 2+)

Waves	50/50	40/60	20/80	**Improvement**
27-32	10.1%	9.2%	10.1%	1%
29-34	15.3%	14.0%	12.3%	19%
31-36	13.4%	12.5%	11.3%	15%
33-38	9.9%	9.3%	10.0%	-1%
35-40	13.7%	11.8%	9.9%	27%
Average	**12.5%**	**11.4%**	**10.7%**	**12%**

"Improvement" refers to the % reduction in relative error when using 20/80 weighting vs. 50/50. In three of the five periods examined, improvement was substantial, while there was virtually no difference in the other two periods.

Source: MRI, Waves 27 through 40

Table 5: Average Relative Error

(Magazines with z scores of 2+)

Waves	50/50	40/60	20/80	0/100
27-32	10.1%	9.2%	10.1%	12.6%
29-34	15.3%	14.0%	12.3%	12.1%
31-36	13.4%	12.5%	11.3%	11.3%
33-38	9.9%	9.3%	10.0%	11.6%
35-40	13.7%	11.8%	9.9%	8.6%
Average	**12.5%**	**11.4%**	**10.7%**	**11.2%**

Source: MRI, Waves 27 through 40