AN EXAMINATION OF DIFFERENTIAL WEIGHTING OF SUCCESSIVE WAVES OF PRODUCT USAGE ESTIMATES

Stuart Gray, BBDO
Julian Baim & Martin Frankel, MRI-NOP World

In a 1999 paper by the same authors, it was concluded that giving more weight to the more recent data in combining successive waves of magazine readership data led to better predictions of subsequent estimates. Specifically, 20/80 weighting produced better predictions than the traditional 50/50 weighting used in double-bases, especially in cases where there was a particularly significant trend in a magazine's readership.

The purpose of this paper is to report on a similar examination of product usage data. The underlying thought is that if similar findings resulted, then it would make sense to apply 20/80 weighting to both readership *and* product usage data when employing double-bases, allowing media planners to perform their customary cross-tabulations with appropriate weights.

We believe that the differences observed in two successive report periods of product and brand usage data may be partially due to sampling variation, but may be partly due to actual trends in usage for at least some of the product categories and brands.

In order to test that hypothesis, successive periods of product usage 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 31 through 40, representing March 1994 through March 1999) were analyzed to perform the predictions retrospectively, with category usage and brand usage analyzed separately. Following are our findings:

Category Data

Overall, 20/80 weighting proved best, with 10% improvement over 50/50 in predicting the next period. 40/60 weighting produced 4% improvement over doublebase (see Table 1). Since these reflected only very modest improvements, we then hypothesized that categories 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 categories between those manifesting a z score[1] of 2 or more in their change in average users and those with a z score of 3 or more. In the case of the former group, the improvement produced by 20/80 weighting over 50/50 was 20% while the improvement for the latter group was 27%. 40/60 weighting again proved to be better than 50/50, but not as good as 20/80 for either group of categories (see Tables 2 and 3).

Brand Data

Overall, there was virtually no difference among the three weighting schemes (see Table 4), but once again we hypothesized that brands with more significant changes from one period to the next would benefit more from giving more weight to the more recent data.

Therefore, as with the categories, we divided the brands between those manifesting a z score of 2 or more in their change in average users and those with a z score of 3 or more. In the case of the former group, the improvement produced by 20/80 weighting over 50/50 was 3% while the improvement for the latter group was 10%. For brands, 40/60 weighting proved to be as good as 20/80 for the z=2+ group, but not as good as 20/80 for the z=3+ group (see Tables 5 and 6).

[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.

We also examined the consistency of our findings by looking at each of the three sets of forecasts separately. In other words, we looked at the results of different weighting for using waves 31/32 and 33/34 to predict 35/36, 33/34 and 35/36 to predict 37/38, and 35/36 and 37/38 to predict 39/40. In all three category data seta 20/80 weighting produced lower errors that 50/50 or 40/60, regardless of z score group. The comparable brand data analysis was not as clear-cut. On an overall basis, only one of the three data sets produced a lower error for 20/80, and the same finding resulted from the brands with 2+ z scores. With the brands with 3+ z scores, however, 20/80 was better than 50/50 all 3 times, and better than 40/60 2 out of 3 times.

Since the presumption based on the analysis to this point was that it is appropriate to give more weight to more recent data when categories or brands 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.

The results of that analysis demonstrated that while putting all the weight on the most recent data (0/100) was as good or better than 50/50 for categories, in the case of brand data 0/100 was never as good as 20/80 (see Table 7).

We theorize that in the case of category data, the sample sizes are generally large enough to allow the reflection of real trends in the most recent data, whereas in the case of brand data, the utilization of additional sample is needed to produce more stable estimates of future usage. 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 findings support the conclusion that 20/80 is an appropriate weighting scheme for doublebase analyses that utilize both magazine readership and product usage data.

Table 1

Average Relative Error

(44 Product Categories)

	50/50	40/60	20/80
Waves			
31-36	4.1%	3.9%	3.5%
33-38	8.3%	8.0%	7.6%
35-40	3.5%	3.2%	3.1%
Average	**5.3%**	**5.0%**	**4.7%**

Read this table as follows: When usage estimates for waves 31&32 were combined with waves 33&34 on a 50/50 basis (Doublebase) to predict waves 35&36, the average relative error for all 44 categories was 4.1%

The second row reflects analogous data for waves 33&34 and 35&36 predciting 37&38. The bottom line reflects the average of all 3 sets of data. Thus, an average relative error of 4.7% for 20/80 weighting represents a 10% improvement over the 5.3% error for the normal doublebase weighting of 50/50.

Source: MRI, Waves 31 through 40

Table 2

Average Relative Error

(Categories with z scores of 2+)

	50/50	40/60	20/80
Waves			
31-36	4.4%	3.9%	3.0%
33-38	8.6%	8.1%	7.3%
35-40	3.2%	2.9%	2.7%
Average	**5.4%**	**5.0%**	**4.3%**

Read this table as follows: When usage estimates for Waves 31&32 were combined with waves 33&34 on a 50/50 basis (Doublebase) to predict waves 35&36, the average relative error (for categories whose usage estimate change from waves 31&32 to waves 33&34 reflected a z score of 2 or more) was 4.4%

For these product categories the average relative error of 4.3% for 20/80 weighting represents a 20% improvement over the 5.4% error for 50/50.

Source: MRI, Waves 31 through 40

Table 3

Average Relative Error

(Categories with z scores of 3+)

	50/50	40/60	20/80
Waves			
31-36	4.2%	3.6%	2.6%
33-38	8.9%	8.3%	7.3%
35-40	4.2%	3.5%	2.8%
Average	**5.8%**	**5.1%**	**4.2%**

Read this table as follows: When usage estimates for waves 31&32 were combined with waves 33&34 on a 50/50 basis (Doublebase) to predict waves 35&36, the average relative error (for categories whose usage estimate change from waves 31&32 to waves 33&34 reflected a z score of 3 or more) was 4.2%

For these categories the average relative error of 4.2% for 20/80 weighting represents a 27% improvement over the 5.8% error for 50/50.

Source: MRI, Waves 31 through 40

Table 4

Average Relative Error

(60 Brands)

	50/50	40/60	20/80
Waves			
31-36	15.5%	15.1%	14.7%
33-38	12.5%	12.5%	12.8%
35-40	9.4%	9.4%	10.0%
Average	**12.5%**	**12.3%**	**12.5%**

Read this table as follows: When usage estimates for waves 31&32 were combined with waves 33&34 on a 50/50 basis (Doublebase) to predict waves 35&36, the average relative error for all 60 brands was 15.5%.

The second row reflects analogous data for waves 33&34 and 35&36 predciting 37&38. The bottom line reflects the average of all 3 sets of data. Thus, an average relative error of 12.5% for 20/80 weighting represents no improvement over the error for the normal doublebase weighting of 50/50.(40/60 showed a very slight improvement over 50/50).

Source: MRI, Waves 31 through 40

Table 5

Average Relative Error

(Brands with z scores of 2+)

	50/50	40/60	20/80
Waves			
31-36	19.2%	18.1%	16.7%
33-38	11.9%	12.1%	12.9%
35-40	9.4%	9.1%	9.8%
Average	**13.5%**	**13.1%**	**13.1%**

Read this table as follows: When usage estimates for waves 31&32 were combined with waves 33&34 on a 50/50 basis (Doublebase) to predict waves 35&36, the average relative error (for brands whose usage estimate change from waves 31&32 to waves 33&34 reflected a z score of 2 or more) was 19.2%

For these brands the average relative error of 13.1% for 20/80 weighting represents a 3% improvement over the 13.5% error for 50/50.

Source: MRI, Waves 31 through 40

Table 6

Average Relative Error

(Brands with z scores of 3+)

	50/50	40/60	20/80
Waves			
31-36	26.2%	24.8%	23.0%
33-38	14.9%	14.2%	13.0%
35-40	10.2%	9.9%	10.2%
Average	**17.1%**	**16.3%**	**15.4%**

Read this table as follows: When usage estimates for waves 31&32 were combined with waves 33&34 on a 50/50 basis (Doublebase) to predict waves 35&36, the average relative error (for brands whose usage estimate change from waves 31&32 to waves 33&34 reflected a z score of 3 or more) was 26.2%

For these brands the average relative error of 13.1% for 20/80 weighting represents a 3% improvement over the 13.5% error for 50/50.

Source: MRI, Waves 31 through 40

Table 7

	50/50	40/60	20/80	0/100
Categories- all				
Waves				
31-36	4.1%	3.9%	3.5%	3.3%
33-38	8.3%	8.0%	7.6%	7.2%
35-40	3.5%	3.2%	3.1%	3.6%
Average	5.3%	5.0%	4.7%	4.7%
Categories- z=2+				
Waves				
31-36	4.4%	3.9%	3.0%	2.4%
33-38	8.6%	8.1%	7.3%	6.5%
35-40	3.2%	2.9%	2.7%	3.2%
Average	5.4%	5.0%	4.3%	4.0%
Categories- z=3+				
Waves				
31-36	4.2%	3.6%	2.6%	2.1%
33-38	8.9%	8.3%	7.3%	6.5%
35-40	4.2%	3.5%	2.8%	3.1%
Average	5.8%	5.1%	4.2%	3.9%
Brands- all				
Waves				
31-36	15.5%	15.1%	14.7%	14.7%
33-38	12.5%	12.5%	12.8%	13.6%
35-40	9.4%	9.4%	10.0%	11.4%
Average	12.5%	12.3%	12.5%	13.2%
Brands- z=2+				
Waves				
31-36	19.2%	18.1%	16.7%	16.5%
33-38	11.9%	12.1%	12.9%	15.0%
35-40	9.4%	9.1%	9.8%	12.3%
Average	13.5%	13.1%	13.1%	14.6%
Brands- z=3+				
Waves				
31-36	26.2%	24.8%	23.0%	22.7%
33-38	14.9%	14.2%	13.0%	14.1%
35-40	10.2%	9.9%	10.2%	12.3%
Average	17.1%	16.3%	15.4%	16.4%