

EVALUATING TV AND PRINT SALES EFFECTS

Bruce Goerlich, Starcom MediaVest Group

Acknowledgement

Many thanks to Simon Broadbent for his help and insights. I am honored to have him as a colleague.

Introduction

Media researchers are well aware of two trends in media selection. One is the way planning agencies now exploit media fragmentation: schedules are built of more components in the hunt for reach and for effectiveness. These components often combine media different in kind – TV, promotions, print, the net and so on.

The second trend is growing acceptance of market-mix modeling as a way to evaluate the effectiveness achieved by this mixture. Sales results are being used in compliment with measures of ad exposure. SMG, as do other major planning agencies, now provide advertisers with tables of ‘the proportion of sales due to...’ and media investments are evaluated that way.

Given that people have been modeling the effects of various media for years, surprisingly little is known about their findings. The Editor of the *Journal of Advertising Research* wrote¹ recently that the comparisons made between media effectiveness over the last several years have been ‘mostly not too helpful.’

Much of the work has been of little use to hard-headed managers concerned with the allocation of monies across media. They are not much impressed by qualitative assessments, descriptions of media environments and creative potential, improvements in targeting and reach, or laboratory tests. Necessary for creative work and judgement calls, these do not address the question, ‘How was it for you?’

Thus a report² on the relationships consumers have with media may tell us that shoppers for apparel find that it is in magazines they get the most useful information. Fair enough, and the long list of categories covered in these interviews tells us a lot about shoppers’ views of media. But that is not return per dollar.

Sometimes undue expectations are raised by the way research is presented. For example, when we read³ that this is ‘the age of accountability’ when ‘advertisers demand measurable returns’, we are rather let down by the following report about advertising awareness. It is small consolation to read that there is a significant relationship between awareness and the scores on ‘definitely or probably will buy.’ Again, no dollars there.

As Alan Smith has long argued⁴, magazine publishers overspend on traditional media research, while newer methods⁵ are taking too long to become accepted. Erwin Ephron has pointed out⁶ that cost discounting and arguments about opportunities to see cut less ice than evidence of effects.

ESOMAR has taken a full part⁷ in this debate, but papers it has published may not yet have explained fully the problems along the way – nor the size of the potential benefits to the print medium when effectiveness is properly evaluated.

¹ Editorial, May 1999

² *Media Choices 2000*, Erdos & Morgan

³ *The effectiveness of magazines in the marketing mix*, 1998, Millward Brown for the Magazine Publishers Association

⁴ Smith, A. *Money well spent?* In Admap, October 1995

⁵ Smith, A. *How to build planning insight from campaign research* in Admap, November 1999, Consterdine, Guy. *Magazine research synthesised* in Admap, May 1996, Pollard, L. *Measuring the effectiveness of magazine advertising* in Admap, 1998.

⁶ Ephron, E. *Response, not readership, is print’s major problem* in Admap, February 1998.

⁷ Cooke, B. *Advertising in magazines is different*, Cooper, G. and Jones, G. *Women’s magazines – delivering results*, Rolls, W. and Gaines, H. *Diagnostic evaluation: a vital component of print research* – all in 1997

What has to be decided during market-mix modeling?

How should we evaluate the sales effects of two different media used in the same campaign?

I suppose we have historical data about the media costs and achievements, about the sales of the brands and about other factors which may affect those sales (competitors’ advertising, retail prices in the category, store activities, times of product changes and so on).

Most analysts agree that regression is the basic tool to disentangle this complex situation, but then the disagreements start. For example:

- Is our objective variable market share or kilograms?
- What are the units in which to measure media variables? Dollars spent? Target ratings? Reach?
- How should we decay ad effects after exposure, and estimate these rates of decay?
- Do we allow for diminishing returns?
- What about synergy between media? Indirect effects?
- Are price effects from actual or relative prices?
- Should we use the coefficient in a linear or multiplicative regression?

Our current position on these questions is shown first by an example. I then generalize on these findings, based on our Brand Library™ of case histories.

Example

For the brand studied here, we had two years of weekly data. The brand split its media budget 82% to TV and 18% to print. On TV we bought 1,450 target ratings a year, spread rather steadily but with some peaks. The use of print declined in the second year. It was a low-priced brand (relative price 78% of the category average), facing growing competition (its TV competitors outspent it eight to one), and therefore losing share. Chart 1 illustrates the basic facts.

Brand Volume and Media Data

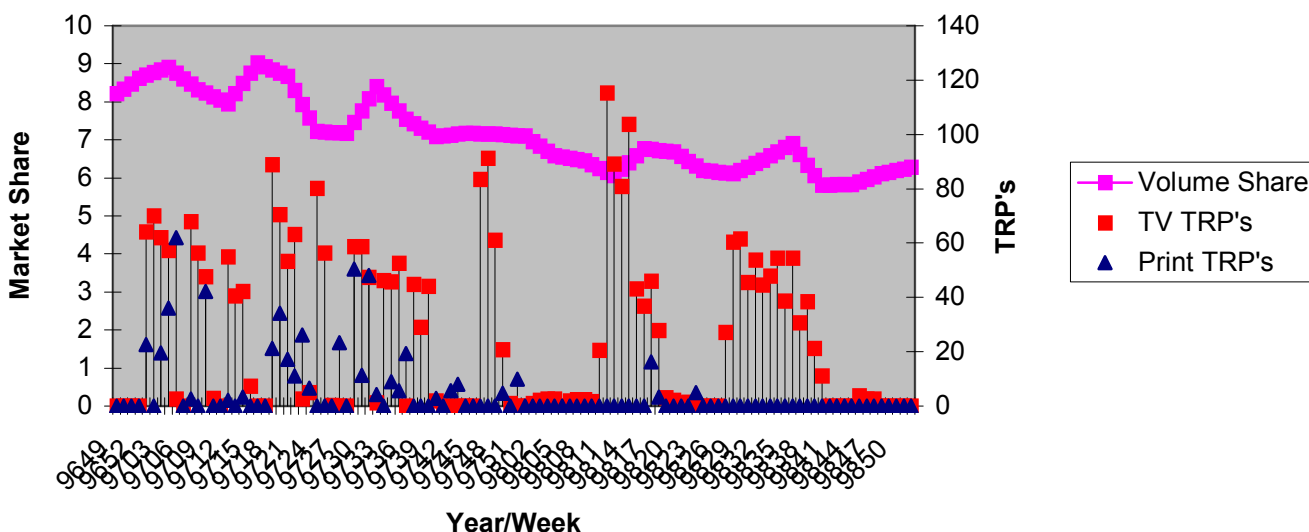


Chart 1

We chose to model volume share – our normal decision but not a fixed one. For TV, our unit is always a 30-second cost equivalent rating, mirroring closely the actual TV impact, despite a variety of spot lengths and costs. For print, our unit was determined by spend, and transformed into the number of equivalent TV ratings we could have bought. Thus both media were in the same units. From these ratings we created adstocks at various half lives (Adstock being the cumulative delayed impact of advertising on sales, and half-life being when half that impact occurs)⁸. We indexed our brand’s retail cost (spend in the local currency divided by kilograms) on the same figure for the category each week, picking up not only our average relative price, but also the effects of our price promotions and competitors’. We had distribution data, but this did not vary enough here to add to the analysis.

⁸ For a thorough review of the Adstock concept see Simon Broadbent’s *When to Advertise*, Admap Publications, Henley-On-Thames, U.K., 1999 .

In regression, we picked the half life for each medium which gave us the best fit, and succeeded in raising r-squared to 0.96. The details were:

	<i>Half life, weeks</i>	<i>Coefficient</i>	<i>t-value</i>
Our relative price		-0.1310	-16.3
Our TV	4	0.0076	4.7
Our print	6	0.0336	6.0
Competitors' advertising	13	-0.0148	-20.2

The unusually high r-squared does not tell us we have an excellent explanation. On the contrary, it warns us that the simultaneous decline in share and rise of competitors' advertising indicates something different. Their advertising is probably a symbol of pressure they are putting on us in other ways, particularly on the shelf and in the trade.

The different half lives make sense. We might expect print to have a longer half life than TV, because of the delay between publication and reading (this can of course be directly estimated for different sorts of newspapers and magazines). However this figure depends more on the purpose and memorability of the advertising. We expect competitors to have a slower effect on us, as this is a rather constant background, with less prominent blips in sales effects.

We have not yet used reach much in our work, and in any case, at the relatively low TV weights here, reach is very highly correlated with ratings.

We found we could improve on this fit, raising r-squared to 0.97, by substituting a diminishing returns expression for the linear assumption above. We define a 'repetition' parameter R, and we replace the adstocks used above with $1-R^{OTS}$, where R is chosen separately for each medium and OTS are the relevant weekly adstocks/100. Thus at the rate of 100 ratings a week, 1-R is the proportion of maximum response reached. At 200 ratings a week, R is the proportion added to the first response. Chart 2 below shows that at a Repetition Factor of .2 diminishing returns are hit fairly soon- e.g. lower levels of frequency provide the best return. With a Repetition Factor of .8, adding frequency continues to exhibit additional impact.

Diminishing Returns

- R(repetition) function ties frequency increases to increases in impact

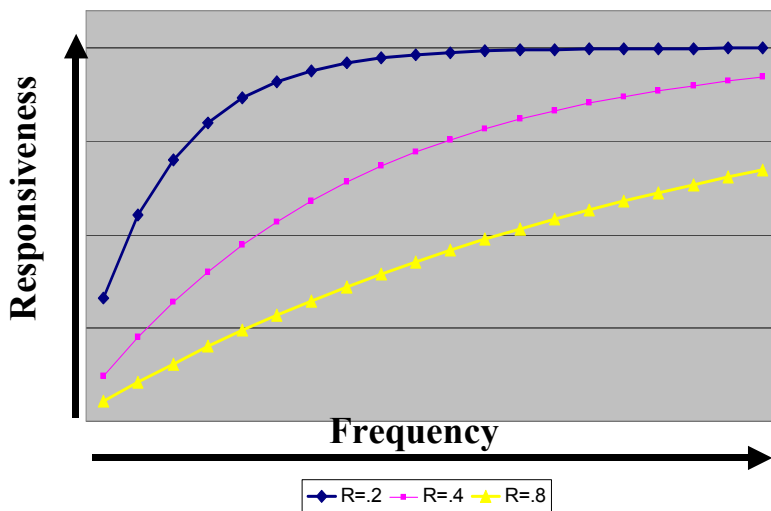


Chart 2

We can make the example more specific by showing the illustrating the impact remaining at different Adstock levels with varying R parameters in chart 3.

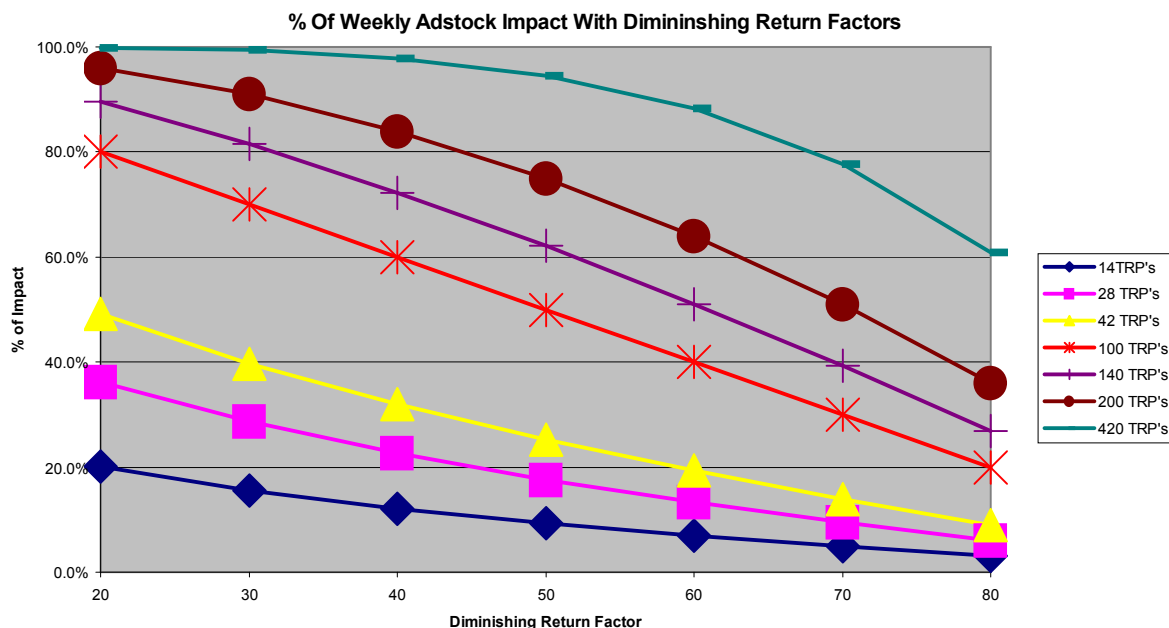


Chart 3

For this brand, we found the best R for both television adstocks is 0.4, and for print .3.

For such low numbers, frequency hits a plateau early: there really is not much return from high repetition. For TV, pressure has reached 60% of its maximum at the rate of 100 ratings a week, and 84% at 200. However, in this case the result is of more theoretical interest than practical: at the low weekly ratings we can afford, the response function is close to linear anyway.

That last statement looks odd, given that I have said the response function is rather convex. But convexity only cuts in at quite high rates. For example, our average is about 28 ratings a week. Consider a range of 14 ratings either way. At 14 ratings, the rate of return per rating is 0.009, hardly more than the rate at 42 ratings, which is 0.008. It is very different for 140 ratings compared with 420, when the returns per rating are 0.005 and 0.002.

Finally, we can collect the information needed to evaluate our two media:

Medium	Coefficient	Average for variable	Coefficient x average	Coeff x var divided by sales
TV	0.0076	27.9	0.21	0.030
Print	0.0336	6.0	0.20	0.028

Interpreting this table has been a problem for many, especially for the last two columns. These may lead some to say, 'TV and print have equal effects.' Since the final column gives the elasticity of the two media, we even hear, 'Each medium contributes about 3% to sales'.

Consider the following:

One of the statistical assumptions made in modeling is that the variables in the equation are independent, which is rarely the case.

Then we must remember the conditions under which the data were collected. We have observed the print effect only when TV was present. Is there synergy between the two? That is, does print work better when the reader has seen a TV commercial? This is very likely – it is assumed in integrated planning. But we cannot measure how much synergy in these circumstances – we have never used print without TV.

Does the direct effect of adstock on sales variation capture all the effects of the medium? This is very unlikely. To accept this would be to say that all the remaining sales are independent of media exposure – not only for shoppers, but also for retailers, representatives and so on. We are talking about an advertised brand here, and it is likely that part of the sales base is due to very long term ad effects.

For all these reasons, we have to exercise extreme caution in interpreting the last two columns of this table.

The first statement quoted above is roughly correct – subject to all the reservations we have made. That is, we have estimated that the short-term sales added by each medium are about the same.

The second statement is definitely misleading, for all the reasons given above, particularly the points about synergy and ‘sales independent of media exposure’. Three percent cannot be guaranteed as the proportion of sales which would be lost if we stopped advertising in one of the media.

It is also incorrect if we interpret it, as ‘another dollar spent on either medium will give about the same return’. The correct number to evaluate another dollar spent is the coefficient⁹. This is because the assumption in regression is that the coefficient (or slope of the line fitted) shows by how much sales increase, as we move along the axis for ratings equivalents (remember we have used the same units here for TV and print). The two coefficients are 0.0076 for TV and 0.0336 for print. By this measure, print is over four times as effective as TV on an equalized spending basis.

A coefficient is not easy to explain to some managers, so we often convert it to the following index, which we call the ‘100 TRP impact’. What if we add annually 100 ratings equivalents to the medium? This is 100/52 or 1.9 ratings a week. Thus the sales increase for TV, if we added 100 ratings equivalents a year, is 0.0146; for print it is 0.0646. As percentages of the current sales average, these are additions of 0.2% and 0.9%. These are the sorts of numbers management understands.

I have pointed out the great difference between ‘print and TV have equal effect’ and ‘print is over four times as effective as TV’. I have shown why the second is more correct for evaluation and planning purposes. There is still the question, Why?

The technical answer is easy. Because TV has over four times as much spent on it, a one percent increase in its budget is over four times as much as it would be for print. Hence the use of elasticity to compare the media is not comparing like with like. In addition, we know that the print effect takes place at a lower position on the x-axis of the response function. Even if the two coefficients and response parameters were the same, the effect would therefore be at a point on the curve with the steeper slope, and we would measure a higher rate of response.

The marketplace answer is harder, and requires more detailed work. It is very likely that the print ads were written for a more direct response, perhaps with coupons, recipes and so on, while TV was more for brand building. Print can also be more precisely targeted than TV, and hence has less wastage.

It was probably a mistake for this brand to have cut print in the second year.

Brand Library™

This SMG resource helps us to learn about how advertising works, through econometric analysis of a database of brand information¹⁰. The brands chosen were those of most interest to Starcom MediaVest, and for which we had access to sufficient data.

They were packaged goods, with an average share of category of 12%. This does not mean they were all large brands; in fact, a third had shares under 5%. The majority were from the US, UK and Australia; the rest were from the European Continent and the Far East. They tended to be heavy advertisers, and 45% of them had retail prices between 1 and 20% above the average for their categories.

In addition to the statistical data, we interviewed the agency and client teams running the campaigns. They completed 29 scales describing the brand situation, the category and communication factors.

From our collection I have picked the 83 cases we have of measured TV effects, and 17 of print, ignoring the radio and outdoor information. In our sample, print advertising tends to be used for larger brands and TV for more expensive brands. Print is definitely the junior partner; getting 15% of the total spend.

⁹ More detail in Simon Broadbent's *If the question is ad effect, the answer is 'Not elasticities'* in the Journal of Advertising Research, March/April 2001, pp 7 – 11.

¹⁰ See for example Simon Broadbent's *What do advertisements really do for brands?* in the International Journal of Advertising, 9, No.2, pp 147 - 165

Not only is print the junior partner overall. Nearly three-quarters of the print examples are at a weight considered very low in TV terms. And print has no examples here of a weight over 50 TV ratings a week, but half of the TV cases are above this level as shown in chart 4.

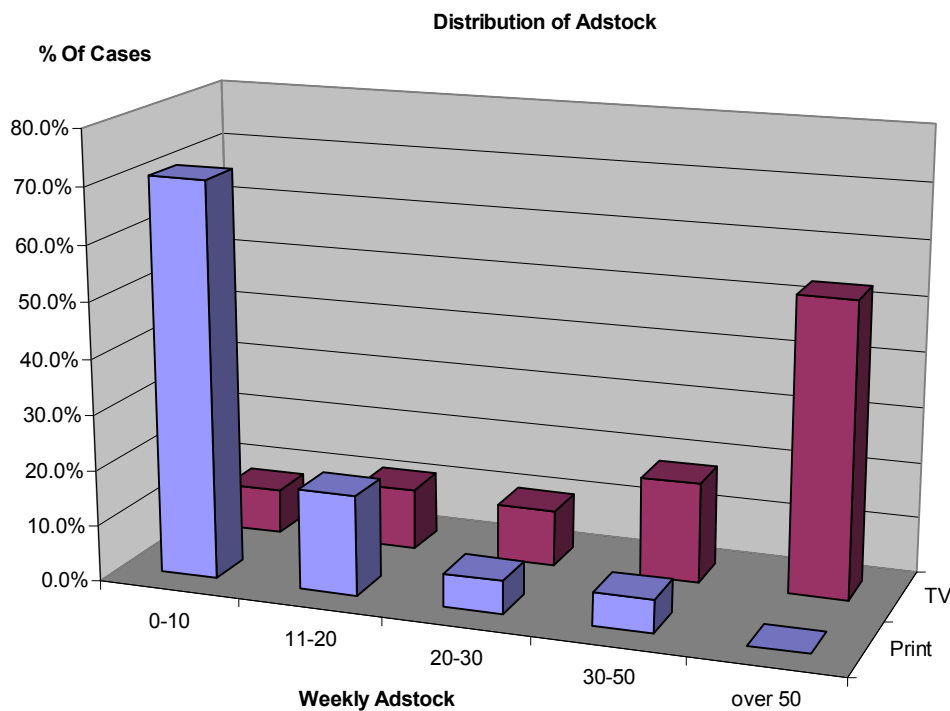


Chart 4

We know that estimates of ad effects are higher for longer half lives, so it is important that these media had about the same averages: 12 weeks for TV and 13 weeks for print as shown in chart 5. Since print effects are delayed by the gap between publication and reading, on average by more than a week, this shows that, once read, the effect of a print ad decays faster¹¹ than of a TV ad. This finding runs somewhat contrary to what we might expect- a TV ad works quickly, and print lingers. While we do know that print exposures accumulate more slowly than TV, what we are finding is that the brand situation is a more important determinate of the half-life, than the medium employed.

¹¹ This confirms a Millward Brown finding: see Farr, A. *How to harness the power of magazine advertising*, in Admap, December 1995

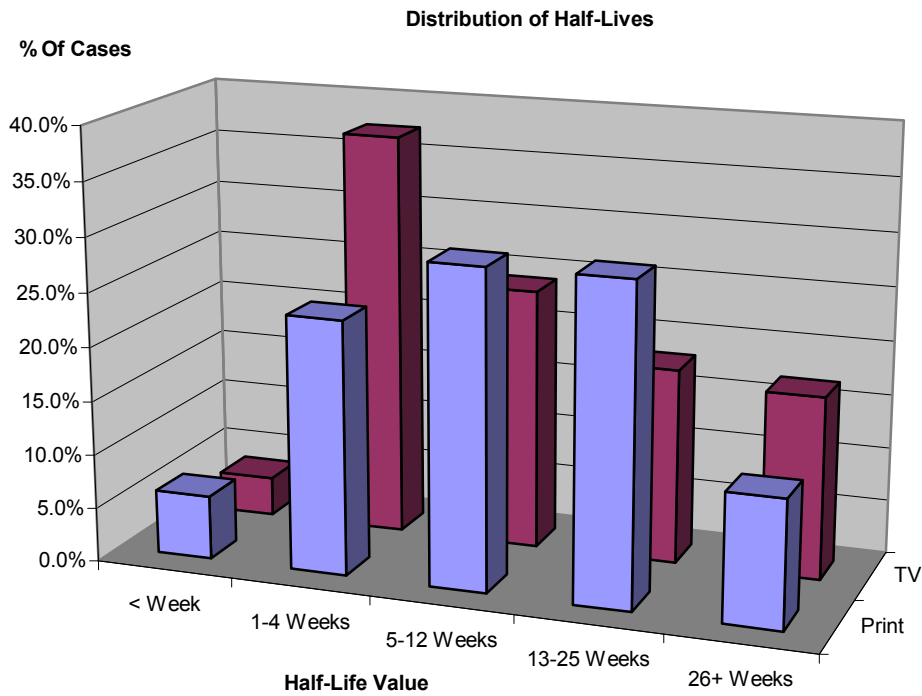


Chart 5

Looking at diminishing returns, TV is likely to be closer to having a linear effect ($R = 0.6$ on average) than print (0.4) as shown in Chart 6. This indicates that TV tends to work best as a reach vehicle, e.g. brands are better off using dispersion and high reach dayparts in their plans. Print on the other hand does not work as well as a reach vehicle. In general, brands would be better off using targeted print that generates multiple exposures. Of course, the individual brand situation is a greater determinate of the Repetition Factor.

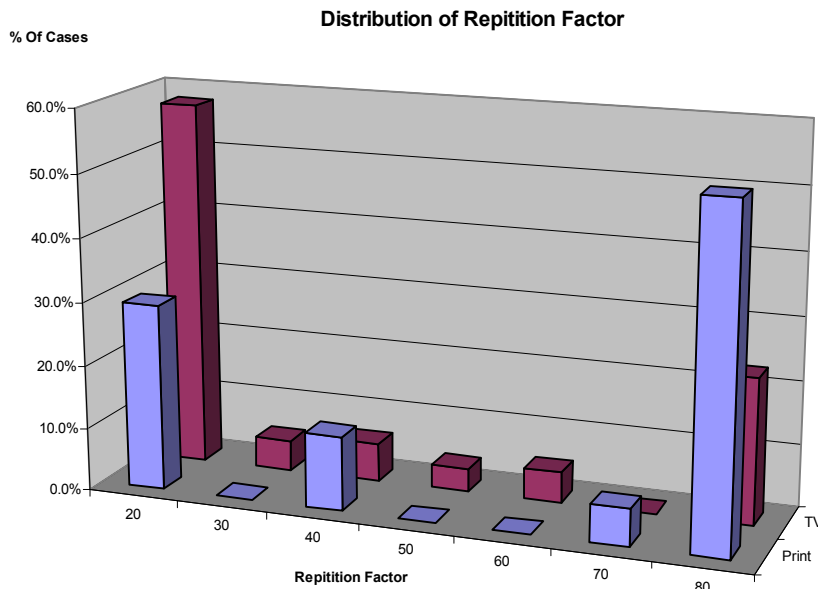


Chart 6

We now get to the key issues – What does our data tell us about how TV and print weights are distributed? About the effects they have? And are these compatible?

The really interesting finding is as follows:

	<i>Elasticity</i>	<i>Percent volume sales increase for 100 more ratings equivalents annually</i>
TV	0.18	0.85
Print	0.09	3.53

Some would conclude from the elasticity numbers that TV is twice as effective as print. This would be wrong. The reason for these very different evaluations is the same as explained in the example – because spend is higher in TV, an elasticity measure makes it look more effective. One percent more spend is a lot more in TV than in print.

When we standardize the additional spend to get a true measure of the return from an extra dollar, we repeat the finding in the example – we get about four times the return from print as from TV. Not only this, the TV average for the percent volume sales increase for 100 more ratings equivalents annually is exceeded by print in 71% of the cases measured as shown in chart 7.

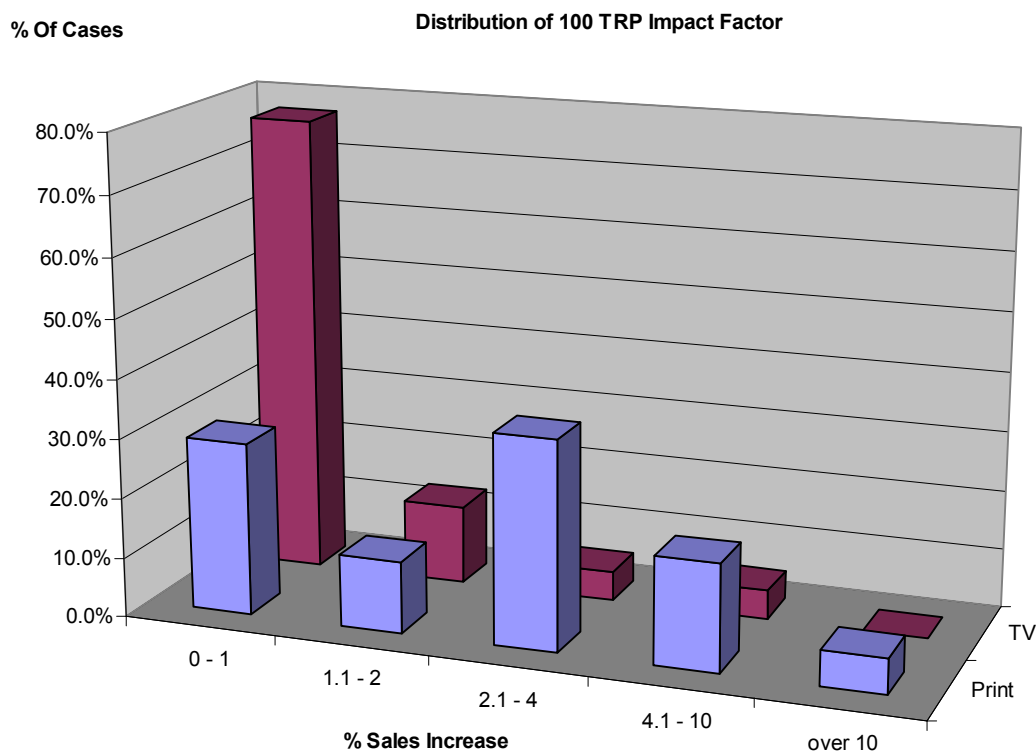


Chart 7

Finally, the teams on each brand seemed to recognize the higher effectiveness of print. They used a scale from 1 to 10 for their estimate of the sales effectiveness of each campaign. TV scored an average of 5.6, but print scored 6.5.

Conclusions

There are many decisions to be made when using market mix modeling to evaluate media. Do not assume that all evaluations are equally reliable.

The choices we normally make are demonstrated in my example.

Generalizing from our database, I suggest the following are important findings:

Per dollar spent, print has usually been more sales-effective than TV.

The method of comparison is vital – if we had incorrectly used elasticity, we would have concluded TV was the more sales-effective.

Print works best as a frequency medium.

Controlling for the delayed exposure patterns for Print, Print and TV have about the same timing impact on sales.

Of course these learnings are not necessarily universal. But they do show that Print does have value- more than perhaps we realize.

