International Journal of Pharmaceutical and Healthcare Marketing

Optimization of the promotion mix in the healthcare industry

Dominique Haughton, Guangying Hua, Danny Jin, John Lin, Qizhi Wei, Changan Zhang,

Optimization of the promotion mix in the healthcare industry

Dominique Haughton
Bentley University, Waltham, Massachusetts, USA and Université Paris I, Paris, France and Université Toulouse I, Toulouse, France
Guangying Hua
Boston Consulting Group, Boston, Massachusetts, USA
Danny Jin, John Lin and Qizhi Wei
Epsilon, Wakefield, Massachusetts, USA, and
Changan Zhang
Bentley University, Waltham, Massachusetts, USA

Received 15 March 2013
Revised 1 March 2014
25 September 2014
23 March 2015
Accepted 14 July 2015

Abstract

Purpose – The purpose of this paper is to propose data mining techniques to model the return on investment from various types of promotional spending to market a drug and then use the model to draw conclusions on how the pharmaceutical industry might go about allocating promotion expenditures in a more efficient manner, potentially reducing costs to the consumer. The main contributions of the paper are two-fold. First, it demonstrates how to undertake a promotion mix optimization process in the pharmaceutical context and carry it through from the beginning to the end. Second, the paper proposes using directed acyclic graphs (DAGs) to help unravel the direct and indirect effects of various promotional media on sales volume.

Design/methodology/approach – A synthetic data set was constructed to prototype proposed data mining techniques and two analyses approaches were investigated.

Findings – The two methods were found to yield insights into the problem of the promotion mix in the context of the healthcare industry. First, a factor analysis followed by a regression analysis and an optimization algorithm applied to the resulting equation were used. Second, DAG was used to unravel direct and indirect effects of promotional expenditures on new prescriptions.

Research limitations/implications – The data are synthetic and do not incorporate any time autocorrelations.

Practical implications – The promotion mix optimization process is demonstrated from the beginning to the end, and the issue of negative coefficient in promotion mix models are addressed. In addition, a method is proposed to identify direct and indirect effects on new prescriptions.

Social implications – A better allocation of promotional expenditures has the potential for reducing the cost of healthcare to consumers.

Originality/value – The contributions of the paper are two-fold: for the first time in the literature (to the best of the authors' knowledge), the authors have undertaken a promotion mix optimization process and have carried it through from the beginning to the end Second, the authors propose the use of DAGs to help unravel the effects of various promotion media on sales volume, notably direct and indirect effects.

Keywords Healthcare industry, Promotion mix

Paper type Research paper

The authors would like to thank the editor and two referees for most useful comments.

International Journal of Pharmaceutical and Healthcare Marketing
Vol. 9 No. 4, 2015
pp. 289-305

1750-6123
DOI 10.1108/IJPHM-03-2013-0008

Introduction; current best practice in the pharmaceutical industry

The matter of marketing expenditures by the pharmaceutical industry continues to attract considerable attention (see editorial by Mukherjee, 2012 and the introduction by Crié and Chebat (2013) to the special issue in the *Journal of Business Research* on health marketing), particularly at a time when healthcare costs are coming under tight scrutiny. In this context, while marketing expenditures in the pharmaceutical industry have continued to decline (Eyeforpharma, 2014), concerns remain about the proportion of such expenditures relative to that of research and development (The Pew Charitable Trust, 2013) and about direct-to-consumer (DTC) marketing programs, which are legal only in the USA and New Zealand (Liang and Mackey, 2011). These concerns flourish in the context of quotes such as "Corporate marketing is the last bastion of unaccountable spending in corporate America", attributed to Eric Schmidt of Google by Stacey (2012).

Marketing efforts in the pharmaceutical industry can be grouped into two categories:

(1) DTC advertising such as TV, print, internet ads as well as customer relationship management programs; and

(2) professional or direct-to-physician (DTP) promotions including samples, details, journal advertising, professional interactions and non-personal promotions.

Most of the marketing mix analyses in the industry focus on measuring the effects of the different DTC advertising channels after controlling for DTP promotions. These analyses are usually performed by creating a control group of physicians in designated market areas (geographic areas defined by the Nielsen Media Research Company as a group of counties that make up a particular television market) or other geographic areas matched to a test group with identical levels of DTP promotions and physician characteristics. The test and control groups have varying levels of DTC activities; this allows for the measurement of the effect of the different DTC efforts.

New product launches typically utilize marketing mix models to simulate potential outcomes (market volume, return on investment (ROI), etc.) under different media spending scenarios. These models incorporate intermediate outcomes such as brand awareness and other consumer behavior that result from consumers asking their physician about the brand and subsequently filling a prescription for this brand.

This paper focuses on the problem of optimizing the DTP promotional mix and also demonstrates how to undertake the optimization process from the beginning to the end while addressing challenges that are common to such processes, such as correlated predictors.

The paper is organized as follows. In the next section, we review pertinent literature and outline the objectives and contributions of this paper. We then describe the two-phase modeling approach adopted here and the data used to demonstrate our methodology. The following two sections describe each of the two modeling phases in more detail. We then move on to a discussion of the directed acyclic graph (DAG) approach used for identifying direct and indirect effects of promotional expenditures on prescription levels and demonstrate why this approach can be useful in promotion mix problems. The last section is the conclusion.

Review of recent literature

While the issue of optimizing marketing mixes – originating in the concept of the four Ps of marketing (product, price, place and promotion) initially introduced by McCarthy (1960), but championed by Kotler (Lilien *et al.* 1991; Mahajan 2013) – has been of interest for many years in the marketing literature, relatively little such published work has been dedicated to the healthcare industry. Of particular relevance to this paper, albeit in another industry (tourism), is the work by Wolfe and Crotts (2011), where a marketing mix methodology is applied to ticket sales for a museum. Of related interest is the paper by Ataman *et al.* (2010) about the long-term effect of marketing strategy on brand sales, which uses dynamic linear transfer function models applied to scanner data for 25 product categories from the four largest retail chains in France. Figure 1 displays the framework typically adopted for marketing ROI assessment and optimization.

In general, the literature in healthcare marketing (refer to the study by Kotler *et al.* 2008 for an extensive discussion of strategic marketing for healthcare organizations) makes the point that decisions on pharmaceutical marketing expenditures are often of a qualitative or policy nature. However, it is in fact important to quantitatively link marketing expenditures and new prescription volume because a more optimal allocation of expenditures would lead to savings that could be expected to decrease the price of drugs for patients. Related work includes the paper by Ruiz-Conde *et al.* (2014), which considers longitudinal and cross-sectional marketing effects, using monthly data on 24 drugs in three therapeutic categories.

Past work has attempted to address the issue of correlated predictors and the problem of a suitable choice of a set of predictors for the number of new prescriptions for a drug or any other yield dependent variable. Lim and Kirikoshi (2008) and Lim *et al.* (2008) propose to use a genetic algorithm (with the predictive power of a model as an objective function) to help select a suitable set of predictors, in combination with a neural net, or a partial least squares regression. The models perform well, but one issue that remains to be addressed is of how to operationalize the model for marketing expenditure allocation, when, for example, some coefficients are negative. In many cases, practitioners faced with inconvenient negative coefficients are forced to ignore them.

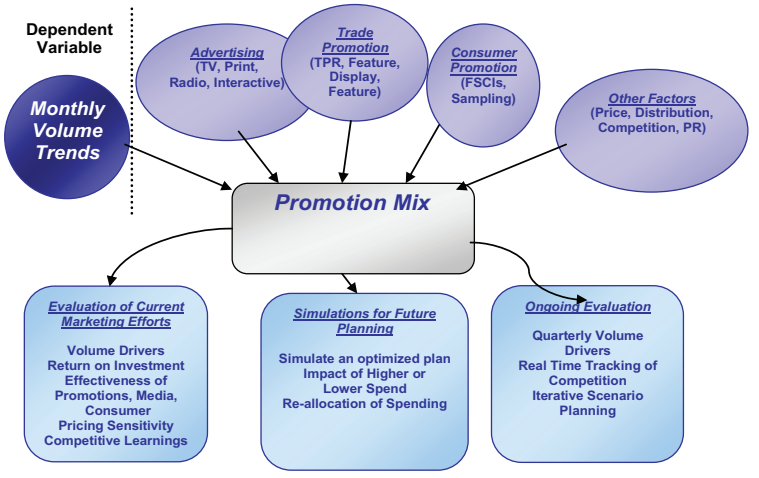


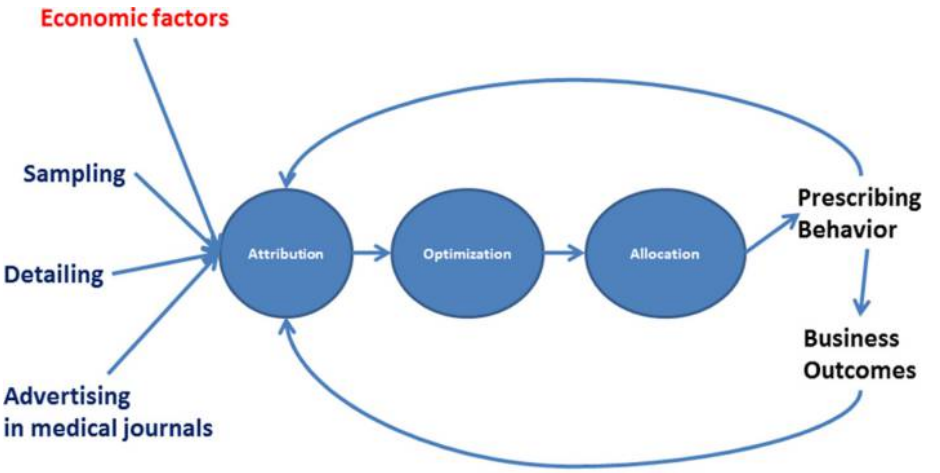
Figure 1. Framework for promotional ROI assessment and optimization

However, negative coefficients can and sometimes do represent reality; we will return to this issue later in the paper.

A recent article by Nichols (2013) proposes an "Advertising Analytics 2.0" framework and creates a case that marketing mix models need to be more sophisticated than traditional linear regression models, notably because the manner in which the various marketing interventions interact is not entirely captured by standard media mix models. An interesting illustrative example suggested by Nichols (2013) is that of a consumer who, while viewing a TV spot for a Toyota Camry, Googles "sedans' on her mobile device. A paid search link for Camry pops up along with car reviews. While looking at reviews, she notices a display ad from a local dealership but does not click on it. One review contains a link to YouTube videos about Camry cars and she watches the Toyota Super Bowl ad from eight months earlier. During her commute to work that week, she sees a Toyota billboard she had not noticed before and she receives a direct mail offer from the company with a time-limited deal. She visits local dealership websites, and eventually heads to a dealer where she test-drives the car and buys it (Nichols, 2013). In Figure 2, we summarize and adapt to the pharmaceutical industry the framework proposed by Nichols (2013).

Manchanda *et al.* (2004) propose a model which takes into account the fact that the level of marketing mix variables is often set with at least partial or a priori knowledge of the likely level of response for each variable and find that this approach results in a more precise estimation of response parameters. They also find that physicians are often not detailed optimally, but that high-volume physicians are detailed to a greater extent than low-volume physicians without regard to responsiveness. The work by Singh (2008) proposes a conceptual model that incorporates several aspects of the network created by physician-salesperson dyads.

More recently, Huber *et al.* (2012) have constructed models applicable to different categories of over-the-counter (OTC) products, which help to identify drivers of product sales in the OTC market. Gönül and Carter (2012) constructed a model of physician



Source: Adapted from Nichols (2013)

Figure 2.
Conceptual framework for pharmaceutical promotional mix optimization

prescribing volume of older versus newer drugs and confirmed the importance of sampling and detailing.

It is useful to note, because markets do differ across countries, the context discussed in this paper is that of pharmaceutical marketing in the USA. In particular, from a research point of view, data availability varies across countries. Obtaining data for research about marketing pharmaceutical expenditures is notoriously difficult in the USA, and data can be obtained almost exclusively in the context of collaborations between academics and practitioners and by of course masking all identifiers of drugs, patients, physicians and even pharmaceutical organizations involved.

In summary, past literature has identified main drivers of prescribing behavior in a number of pharmaceutical industry cases and preliminary promotional mix models have been constructed. However, the issue of correlated predictors and the operationalization of promotional mix models in the pharmaceutical context remains a challenge. In addition, the wider choice of marketing media available given the rapid development of internet platforms and social media has made it all the more important to unravel direct and indirect effects of marketing expenditures on prescription levels. To address these issues, this paper makes the following two-fold contributions:

(1) For the first time in the literature (to the best of our knowledge), we demonstrate how to undertake a promotion mix optimization process and carry it through from the beginning to the end; to achieve this goal, synthetic data are used.

(2) We propose the use of DAGs to help unravel the direct and indirect effects of various promotion media expenditures on sales volume.

We now outline the two main phases that are typically included in a media mix analysis and, after providing details about the synthetic dataset used here, we describe and implement each step of the two phases in a prototype promotional mix process.

Modeling approach

The approach to optimizing the promotion mix for a product typically proceeds in the following two phases.

Phase 1

A model is built for the output variable (typically new prescriptions in a given time period) in terms of a number of relevant independent variables. Phase 1 corresponds to the central and upper sections of Figure 1.

Phase 2

Once a satisfactory model has been constructed, the model is used to evaluate the contribution of each promotional activity to the new prescriptions. This step corresponds to the left lower part of Figure 1.

To illustrate the methods proposed in this paper, we describe below a "Prototype modeling approach for Phases 1 and 2" on the basis of synthetic data. We next introduce the data used to that effect.

Data

Using the Stata statistical software, we generated synthetic data simulated from the correlation matrix and summary statistics for 12 variables used in the study by Lim *et al.* (2008). The variables include, over a period of 71 months, the number of new prescriptions (*nrx*) and 11 variables related to various aspects of different advertising activities (Table I). In the study by Lim *et al.* (2008), the dataset was obtained from a marketing consultant and concerns an antibiotic drug in the USA.

Complete information about the variables can be found in the study by Lim *et al.* (2008), but we reproduced the definitions given in that paper here for the sake of completeness and convenience. As defined by Lim *et al.* (2008):

> "Contacts" (CON) is a product-level report of promotional actions that is provided to physicians and "Calls" (CAL) measures the total number of visits made by pharmaceutical representatives to physicians. A CON can be a full product discussion with a physician, a drug fair set up at the hospital for physicians or a delivery of a product sample. Several products may be discussed during a single call, resulting in the possibility of multiple CONs in a CAL. "Cost of contacts" (COC) includes the costs associated with detailing of representatives that are directed to physicians. "Cost per contact" (CPC) is an estimate of cost per contact whereas "Minutes" (MIN) is the projected sum of time spent with physicians. In a broad sense, we interpret CPC as the quality of CON. This is a rough approximation that accounts for the difficulty and complexity of assessing the physician's overall impression of the representative's visits. "Journal advertising spending" (JAS) captures the expenditure of advertising in medical journals. "Ads" (ADS) measures the number of different layouts of product advertisements in medical journals. If the same ad appears in two journals, it is counted twice. "Ad pages circulated" (ADP) represents the number of total ad pages circulated in journals for a particular product. "Sample" (SAM) shows the projected volume of a product provided as samples to physicians whereas "Extended units samples" (EUS) measures the amount of a product sampled as the number of packages multiplied by the size of the package in tablets, capsules, milliliters, etc. EUS is appropriate for use when the products being compared are similar in terms of dosage form. "Retail value of sample" (RVS) represents the retail value of SAM. "New prescription volume" (NRx) represents the count of new prescriptions dispensed by pharmacists.

We note that in addition to data on the brand of antibiotic (referred to as A) simulated from the statistical summaries in the study by Lim *et al.* (2008), it would be potentially possible to simulate data from the statistical summaries on three other brands B, D and

Table I. List of synthetic variables (following Lim and Kirikoshi, 2008) and summary statistics

Variable	*N*	Minimum	Maximum	Mean	SD
con (contacts)	71	7,657.00	99,188.00	54,113.88	21,606.72
cal (calls)	71	5,361.00	77,657.00	38,964.52	17,162.88
coc (cost of contacts)	71	73,72,97.00	80,417,93.00	43,573,17.52	17,40648.34
cpc (cost per contact)	71	62.00	114.00	87.92	10.67
min (minutes)	71	32,682.00	334,945.00	182,602.33	72,862.46
jas (journal advertising spending)	71	0.00	486,943.00	21,792,6.80	112,177.79
ads (ads)	71	0.00	30.00	14.12	6.57
adp (ad pages circulated)	71	0.00	82.00	39.16	20.22
sam (samples)	71	615.00	26,149,48.00	10,94,355.12	627,396.75
eus (extended unit samples)	71	1,230.00	82,044,39.00	35,566,10.79	17,434,31.60
rvs (retail value of sample)	71	6,660.00	18,502,700.00	85,822,59.60	38,208,27.20
nrx (new prescription volume)	71	84,895.00	24,667,77.00	10,805,44.09	512,578.60

D also covered in the work by Lim *et al.* (2008). We also note that one limitation of our simulated data is that they do not incorporate time auto-correlation – such auto-correlation is however likely to occur in actual data.

Analysis: Phase 1

As is readily visible in Figure 3, a major issue when attempting to build a model for the number of new prescriptions in terms of the 11 predictors is that a high level of correlation exists among the predictors. For example, contacts and cost of contacts are highly correlated as one might expect. Contacts, calls and cost of contacts are highly correlated to the retail value of samples.

The first row of the matrix scatter plot graphs the dependent variable *nrx* in terms of the 11 predictors; a positive linear relationship emerges between *nrx* and each predictor except for *cpc*, where that relationship is negative.

Figure 3.
Matrix scatter plot for 12 synthetic variables

If we attempt to run a linear regression for *nrx* in terms of the 11 predictors, several variables fail to be significant and the negative coefficients for some of the variables might be considered counter-intuitive (Table II). A stepwise regression for *nrx* (Table III) yields a negative coefficient for *jas* and includes a small number of significant predictors.

One option would be to remove *jas* and evaluate the resulting model; this yields an R^2 of 0.910, not a very large loss from the value 0.919, and the coefficients of the remaining three variables are positive and significant. However, this is a matter that would need to be elucidated from domain knowledge of the data; the negative sign could in fact represent a reality. As pointed out by a referee, the fact that *adp* and *ads* might have positive coefficients and *jas* a negative one can occur if the firm substitutes away from advertising in low-cost journals to high-cost journals that are not seen by the target physician as much because of, for example, specialization. We are aware of an interesting case where a negative sign appeared in a similar promotional mix, but this time in the banking industry. It transpired that all coefficients were positive in what looked like a sensible linear regression model, except that of newspaper advertising expenditures. A more careful examination revealed that interest rates were higher for the bank under consideration than for competitors and that those rates were published in the same newspapers.

Rather than dropping a variable with a counter-intuitive sign for its coefficient, a potentially better approach is to extract uncorrelated factors to explain as large as possible an amount of variability in the 11 predictors. A solution with two factors,

Table II.
Linear regression of *nrx* in terms of all 11 predictors (R^2 = 0.931)

Model		Unstandardized coefficients B	Unstandardized coefficients Standard error	Standardized coefficients Beta	*t*	Significance
1	(Constant)	12,967,25.732	522,210.488		2.483	0.016
	con	−31.255	14.707	−1.318	−2.125	0.038
	cal	12.363	16.948	0.414	0.729	0.469
	coc	0.190	0.119	0.646	1.595	0.116
	cpc	−11,301.270	4,710.481	−0.235	−2.399	0.020
	min	1.228	1.648	0.174	0.745	0.459
	jas	−1.556	0.583	−0.341	−2.668	0.010
	ads	33,979.885	14,621.334	0.435	2.324	0.024
	adp	2,713.728	5,320.917	0.107	0.510	0.612
	sam	0.480	0.239	0.588	2.009	0.049
	eus	−0.137	0.077	−0.465	−1.782	0.080
	rvs	0.075	0.025	0.563	3.034	0.004

Table III.
Stepwise linear regression results for *nrx* (R^2 = 0.919)

Model		Unstandardized coefficients B	Unstandardized coefficients Standard error	Standardized coefficients Beta	*t*	Significance
1	(Constant)	22,055.430	52,669.622		0.419	0.677
	jas	−1.214	0.450	−0.266	−2.699	0.009
	ads	33,026.816	9,082.479	0.423	3.636	0.001
	sam	0.228	0.088	0.279	2.602	0.011
	rvs	0.071	0.014	0.527	5.036	0.000

Omponent	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	(%) of variance	Cumulative (%)	Total	(%) of variance	Cumulative (%)	Total	(%) of variance	Cumulative (%)
1	8.487	77.153	77.153	8.487	77.153	77.153	7.938	72.160	72.160
2	1.611	14.649	91.801	1.611	14.649	91.801	2.106	19.148	91.308
3	0.476	4.325	96.126	0.476	4.325	96.126	0.530	4.819	96.126
4	0.220	2.001	98.127						
5	0.097	0.885	99.012						
6	0.051	0.467	99.479						
7	0.022	0.198	99.677						
8	0.021	0.192	99.868						
9	0.009	0.083	99.951						
10	0.003	0.031	99.982						
11	0.002	0.018	100.000						

Table IV.
Factor analysis of all 11 predictors

rotated with a quartimax rotation, explains 91.3 per cent of the total variability in the predictors and is presented in Tables IV and V.

Table V reveals that the first factor seems to represent activities directly targeted to physicians, while the second factor represents more general advertising activities. Table VI displays the coefficients for the Z score of each predictor needed to compute the factors, and Table VII presents the results of a regression of *nrx* on the two factors ($R^2 = 0.885$).

One can then extract a formula for the estimated new prescription volume in terms of the Z scores of the predictors (see the second column of Table VIII), and ultimately, in terms of the predictors themselves. We are now ready to move on to Phase 2 of the process.

Table V.
Loadings of each predictor onto each rotated factor

Variable	Component 1	2
con	0.988	0.046
cal	0.984	0.053
coc	0.964	0.066
cpc	−0.768	0.098
min	0.979	0.031
jas	0.437	0.873
ads	0.594	0.787
adp	0.539	0.828
sam	0.958	0.169
eus	0.921	0.193
rvs	0.951	0.137

Notes: Extraction method: principal component analysis; rotation method: Quartimax with Kaiser normalization; rotation converged in three iterations

Table VI.
Coefficients to compute each factor in terms of the *Z* scores of the predictors

Variable	Component 1	2
con	0.151	−0.109
cal	0.149	−0.105
coc	0.144	−0.094
cpc	−0.135	0.162
min	0.151	−0.117
jas	−0.051	0.446
ads	−0.014	0.375
adp	−0.028	0.406
sam	0.129	−0.034
eus	0.120	−0.015
rvs	0.132	−0.051

Notes: Extraction method: principal component analysis; rotation method: Quartimax with Kaiser normalization; component scores

Model	R	R^2	Adjusted R^2	Standard error of the estimate
Model summary				
1	0.941[a]	0.885	0.882	1.76432E5

		Unstandardized coefficients B	Unstandardized coefficients Standard error	Standardized coefficients Beta	t	Significance
Coefficients[a]						
1	(Constant)	10,805,44.093	20,938.592		51.605	0.000
	Factor 1	470,291.077	21,087.623	0.918	22.302	0.000
	Factor 2	106,415.943	21,087.623	0.208	5.046	0.000

Note: [a]Dependent variable: nrx

Table VII. Linear regression of *nrx* in terms of the two factors

Variable	Coefficient of Z score of variable in objective function	Optimal Z score	Optimal value of variable	Contribution of variable to objective function	Lower limit	Upper limit
con	59,167.95	4	140,540.76	236,671.81	−2	4
cal	58,912.58	4	107,616.04	235,650.32	−2	4
coc	57,713.02	4	113,199,10.88	230,852.08	−2	4
cpc	−46,433.24	−2	66.58	92,866.49	−2	4
min	58,693.49	4	474,052.17	234,773.95	−2	4
jas	23,596.74	−1.94	0.129	−45,841.14	−1.94	4
ads	33,295.52	4	40.40	133,182.09	−2	4
adp	29,855.06	1.94	78.44	57,999.13	−2	4
sam	56,987.67	4	360,3942.12	227,950.67	−2	4
eus	54,746.69	4	105,303,37.19	218,986.76	−2	4
rvs	56,685.92	4	238,655,68.40	226,743.68	−2	4
				Objective function		
	Constraint on the sum of Z scores for all variables	30		184,983,5.83		

Table VIII. Optimization of predictors using linear programming

Analysis: Phase 2

We apply to the objective function obtained in Phase 1 (with weights for each variable Z score given in the second column of Table VIII) an optimization process, here a simple linear programming application, to find the optimal value of the Z scores under some constraints, such as total budget, etc. Figure 4 graphically displays those weights; note the negative weights for the Z score of the variable CPC. For illustrative purposes, we set as constraints the lower limit at −2 and the upper limit at 4 for each Z score (except for *jas* where we set the limit at a Z score of −1.94, which would correspond to a value of *jas* of zero), with an upper limit of 30 for the sum of all Z scores, and obtain the results in Table VIII.

Table VIII reveals that an optimal mix would set the variables *con, cal, coc, min, ads, sam, eus* and *rvs* at their maximum levels; the variables *cpc* and *jas* at their minimum

level; and the variable *adp* at a high, but not at a maximum level (1.94 standard deviations above its mean). Table VIII also gives the optimal values for each variable by transforming the Z scores to actual variable values using the means and standard deviations in Table I. Such an allocation would yield an optimal value of 1,849,835.83 for the new prescription volume.

We note that the constraints we have established in this example are meant to be for demonstration purposes only. In a real-life application, the constraints would apply to actual variables (not Z scores) and the constraint of a sum of at most 30 for Z scores would be replaced by a constraint on the total budget. This would of course take into account the expenditures implied by each variable (for instance the cost of each unit of *jas*, *adp*, etc.).

Another approach to unraveling the effects of correlated predictors: DAGs

Another approach to help unravel the effects of correlated predictors on the new prescription volume is the construction of a DAG followed by the estimation of a structural equation model based on directional links indicated by the DAG. This approach, which relies on the seminal work by Pearl (2009), has been known to work well for the purpose of illuminating direct and indirect effects, the size of which can then be estimated once a hypothetical model is obtained from the DAG (work by Bessler and Loper, 2001, 2002). This is particularly useful in cases, such as here, where theory is scant or non-existent as a guide as to which predictors drive new prescriptions. To give an idea of how this approach works, we use a modified partial correlation (PC) algorithm (PC Pattern) from the Tetrad IV package (The Tetrad Project 2015) to construct a DAG from the synthetic data, with the resulting graph displayed in Figure 5. The values at each node represent the mean of the variable (see Table I), and the values on each directed link represent the impact of an increase of one unit of the origin node variable on the end node variable. Focusing, for example, on the new prescriptions node, labeled *nrx*, we see that only one variable, the *rvs*, impacts new prescription volume directly. An increase of one unit in the total retail value of samples corresponds to an increase in new prescription volume of 0.1261.

A few further interesting features emerge from Figure 5. Samples seem to have an effect on new prescriptions, but indirectly, via their retail value. The cost per contact has a negative effect on the number of contacts, which seems sensible. The variable *con* has a positive effect on *rvs*, but the variable *cpc* has a negative effect on the variable *con*,

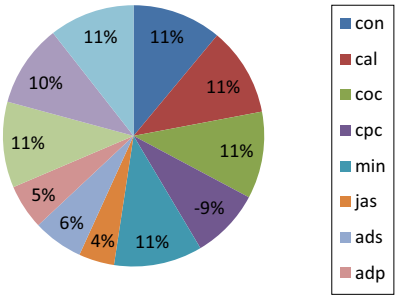


Figure 4.
Weights of Z scores of variables in the objective function

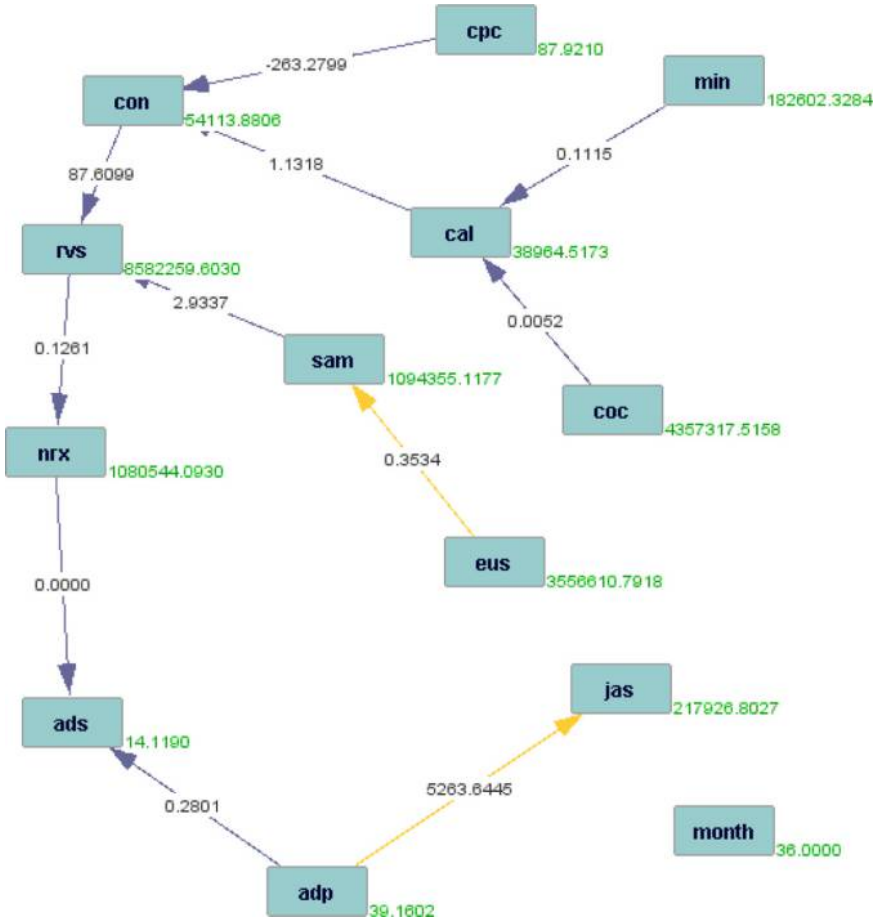


Figure 5.
Directed acyclic graph for 12 synthetic variables

which explains how a negative effect might appear of the variable *cpc* on new prescription volume if one tries a standard regression analysis.

The system of equations which governs the estimated new prescription volume *nrx* is given by:

$$\widehat{nrx} = .1261\ \widehat{rvs}$$

$$\widehat{rvs} = 87.61\ \widehat{con} + 2.934\ \widehat{sam}$$

$$\widehat{sam} = .3534\ eus$$

$$\widehat{con} = 1.1318\ \widehat{cal} - 263.28\ cpc$$

$$\widehat{cal} = .1115\ min + .0052 coc$$

where the hats above the variable names denote estimated values. Combining these equations yields a linear equation for the estimated new prescription volume in terms of the variables *cpc*, *min*, *coc* and *eus*. A linear programming process would then allow for optimizing this objective function under constraints (arising for example from budget limitations) on the variables *cpc*, *min*, *coc* and *eus*, as was done in Table VIII.

We note that the PC algorithm typically yields not one but a class of directed acyclic graphs which are compatible with the data. When a directed arrow appears on the graph, it means that the direction was the same in all graphs generated by the algorithm. When this is not the case, the direction of the link is ambiguous and a link appears without arrows. To estimate the coefficients of each link (estimated here via maximum likelihood), it is necessary to choose one graph in the class of graphs obtained by the PC algorithm and to fix the direction of any remaining ambiguous links. This is the case, for example, for the link between *sam* and *eus* (colored orange in Figure 5). When an ambiguity emerges, a sensible way to lift it is to use domain knowledge and decide which of the orientations is more likely to hold.

For the sake of completeness, we give a brief description of the PC algorithm.

It is important to note that DAGs can infer causality as defined in a study by Pearl (2009), but under very restrictive assumptions, which tend not to be satisfied in many cases. So one would not want to claim any hope of getting definite causality statements out of this analysis; however, what is obtained is a better understanding of which effects might be direct effects and which might be indirect.

Brief description of the PC algorithm

The PC algorithm works essentially as follows. First, a complete undirected graph is created with each variable corresponding to a vertex. Then, the edges are removed in pairs with variables which are independent, either unconditionally or conditionally on a subset of the remaining variables. Independence is tested with standard correlation tests (for continuous data assumed to be multivariate normal) or by a test of independence in contingency tables (for categorical data). Note that Tetrad allows for either continuous data or categorical data, but not a mixture of both types of data, unless the user provides a priori information on which pairs of variables are independent conditionally on other variables. For a discussion of Tetrad and other acyclic graph software packages, refer to the study by Haughton *et al.* 2006.

To orient the surviving links, the PC algorithm proceeds as follows: for each triplet X, Y, Z such that both pairs X, Y and Y, Z are linked but the pair X, Z is not linked, if Y does not appear in any set of variables which when conditioned upon makes X and Z independent, then the triplet X, Y, Z is oriented $X \rightarrow Y \leftarrow Z$, in effect making Y a collider. This makes sense because such an orientation implies that X and Z are dependent given Y. Once all such colliders are identified, the algorithm proceeds like: so if $X \rightarrow Y$, Y and Z are linked and X and Z are not linked, and if there is no arrowhead at Y from Z, then Y, Z is oriented as $Y \rightarrow Z$. Such an orientation implies that X and Z are independent given Y. Appendix B of the Tetrad III user manual (The Tetrad Project 2010) explains how the algorithm unfolds on a particular example. Also refer to Haughton and Haughton (2011), chapter 5, for an introduction to directed acyclic graphs with examples.

Discussion

This paper has proposed a road map which should be helpful to researchers and practitioners involved in promotion or more generally media mix modeling. The synthetic data on which our analyses rely relate to the marketing of a drug, but the methodology can be used in contexts other than healthcare as well.

The issue of negative coefficients in media mix models is one which, while unmentioned in the literature, is quite an important practical problem. In many real-life situations, negative coefficients are either ignored or their sign ignored, when establishing the importance of each component. We have illustrated the fact that the key issue with negative coefficients is whether they arise because of multi-collinearity in the model or because of the genuine negative effect of a particular component. Discovering which of these two options holds is important.

Many promotion or media mix models include expenditures that are quite highly correlated. Because the mechanism which yields new prescriptions or higher sales in general is likely to be quite complex, in this paper, we have suggested that DAGs can be pressed into service to help unravel which expenditures are directly linked to new prescriptions and which expenditures might act via an intermediary variable.

Several issues of interest remain for future work. One is the matter of the presence of competitor promotion expenditures when a few major competitors are present in a particular market. We suggest and investigate elsewhere (Haughton *et al.* 2014) that ignoring competing promotion activity in a media mix model could lead to biased results. Assuming that data on competing sales are available, it is possible to impute with good success the presence or absence of competitor promotional activity when two competitors are in place via the use of hidden Markov models (Haughton *et al.* 2014). For example, more work remains to be done on situations with more than two main competitors.

Two caveats are in order. First, while we have used a factor analysis on the predictors, and then used the factors in a regression analysis for new prescriptions, we could also use a partial least squares regression, which will attempt to extract factors that explain a satisfactory proportion of the variability in the predictors as well as in the new prescriptions. Such an analysis performed on the synthetic data yields similar results to those obtained in this paper, but with a slightly smaller proportion of the variability in the predictors explained (90.5 per cent) and a bit higher proportion of variability in new prescription explained (89.3 per cent). Second, our analysis does not take into account any temporal correlation in the errors of the regression model. The synthetic data were created without such correlation, but it is very likely that temporal error correlation exists in real data and would need to be taken into account, for example, by modeling the errors as autoregressive or even autoregressive-moving average.

It also would be very interesting to apply these techniques to a real-life data set. The synthetic data set presented in this paper is very representative of data used in a promotional mix marketing effort for the marketing of a drug, but, nevertheless, applying the methods to actual data would be very valuable. The difficulty here is the scarcity of such datasets available to researchers.

In particular, a real-life data set is very likely to include time series with auto-correlated variables. An interesting extension of this paper would be to adjust the

methods to take into account the likely auto-correlation in the errors of the media mix regression model.

Last but not the least, the importance of word-of-mouth and viral marketing, notably via social networks, and in particular via social networks that are related to the medical profession, inclusive of co-publication networks, cannot be underestimated. The problem of how to combine social network data (on influential prescribers etc) with more traditional data used in media mix modeling is becoming quite pressing. Media mix models of the future will very probably need to incorporate social network components because ignoring these components risks leading to biased results. A rather key challenge there is to link social network data to existing corporate databases used for targeting physicians and for building marketing programs.

References

Ataman, M.B., Van Heerde, H.J. and Mela, C.F. (2010), "The long-term effect of marketing strategy on brand sales", *Journal of Marketing Research*, Vol. 47 No. 5, pp. 866-882.

Bessler, D.A. (2002), "On world poverty: its causes and effects", available at: http://agecon2.tamu.edu/people/faculty/bessler-david/WebPage/Bessler%20Poverty.pdf (accessed March 2015).

Bessler, D.A. and Loper, N. (2001), "Economic development: evidence from Directed Acyclic Graphs", *The Manchester School*, Vol. 69 No. 4, pp. 457-476.

Crié, D. and Chebat, J.C. (2013), "Health marketing: toward an integrative perspective", *Journal of Business Research*, Vol. 66 No. 1, pp. 123-126.

Eyeforpharma (2014), Decline in Promotional Spending for US Pharma, FC Business Intelligence, available at: http://social.eyeforpharma.com/sales-marketing/decline-promotional-spending-us-pharma (accessed March 2015).

Gönül, F. and Carter, F. (2012), "Estimation of promotional strategies for newer vs older drugs based on physician prescribing data", *International Journal of Pharmaceutical and Healthcare Marketing*, Vol. 6 No. 1, pp. 39-54.

Haughton, D. and Haughton, J. (2011), *Living Standards Analytics: Development Through the Lens of Household Survey Data*, Springer-Verlag, New York, NY.

Haughton, D., Hua, G., Jin, D., Lin, J., Wei, Q. and Zhang, C. (2014), "Imputing unknown competitor marketing activity with a Hidden Markov Chain", *Journal of Direct, Data and Digital Marketing Practice*, Vol. 15 No. 4, pp. 276-287.

Haughton, D., Kamis, A. and Scholten, P. (2006), "A Review of three directed acyclic graphs software packages: MIM, Tetrad, and WinMine", *The American Statistician*, Vol. 60 No. 3, pp. 272-286.

Huber, M., Dippold, K. and Forsthofer, R. (2012), "Which factors drive product sales in the OTC markets?", *International Journal of Pharmaceutical and Healthcare Marketing*, Vol. 6 No. 4, pp. 291-309.

Kotler, P., Shalowitz, J. and Stevens, R. (2008), *Strategic Marketing for Healthcare Organizations*, Wiley Jossey-Bass, San-Francisco, CA.

Liang, B. and Mackey, T. (2011), "Direct-to-Consumer advertising with interactive internet media", *Journal of the American Medical Association*, Vol. 305 No. 8, pp. 824-825.

Lilien, G., Kotler, P. and Moorthy, K.S. (1991), *Marketing Models*, Prentice Hall, NJ.

Lim, C.W., Kirikoshi, T. and Okano, K. (2008), "Modeling the effects of physician-directed promotion using genetic algorithm-partial least squares", *International Journal of Pharmaceutical and Healthcare Marketing*, Vol. 2 No. 3, pp. 195-215.

Lim, C.W. and Kirikoshi,T. (2008), "Understanding the effects of pharmaceutical promotion: a neural network approach guided by genetic algorithm-partial least squares", *Healthcare Management Science*, Vol. 11 No. 4, pp. 359-372.

McCarthy, J.E. (1960), *Basic Marketing: A Managerial Approach*, Irwin, Homewood, IL.

Mahajan, N. (2013), "The thinker interview with Philip Kotler, the father of modern marketing", CKGSB Knowledge, Cheung Kong Graduate School of Business, available at: http://knowledge.ckgsb.edu.cn/2013/10/08/marketing/philip-kotler-four-ps-model-marketing-still-king/ (accessed March 2015).

Manchanda, P., Rossi, P.E. and Chintagunta, P.K. (2004), "Response modeling with nonrandom marketing-mix variables", *Journal of Marketing Research*, Vol. 41 No. 4, pp. 467-478.

Mukherjee, A. (2012), "Advances in pharmaceutical marketing: theoretical and empirical foundations, Editorial", *International Journal of Pharmaceutical and Healthcare Marketing*, Vol. 6 No. 4.

Nichols, W. (2013), *Advertising Analytics 2.0., Harvard Business Review*, Vol. 91 No. 3, pp. 60-68.

Pearl, J. (2009), *Causality*, 2nd edn., Cambridge University Press, Cambridge, MA.

Ruiz-Conde, E., Wieringa, J.E. and Leeflang, P.S.H. (2014), "Competitive diffusion of new prescription drugs: the role of pharmaceutical marketing investment", *Technological Forecasting and Social Change*, Vol. 88 No. 1, pp. 49-63.

Singh, R. (2008), "Network connectedness of pharmaceutical sales rep (FLE)-physician dyad and physician prescription behaviour: a conceptual model", *Journal of Medical Marketing*, Vol. 8 No. 3, pp. 257-268.

Stacey, E.C. (2015), "Measuring the return on marketing investment", *Center for Measurable Marketing at NYU Stern*, available at: http://pages.stern.nyu.edu/~atakos/studentevents/3-28-12MeasuringROMISlideDeck.pdf

The Pew Charitable Trust (2013), "Persuading the prescribers: pharmaceutical industry marketing and its influence on physicians and patients", available at: www.pewhealth.org/other-resource/persuading-the-prescribers-pharmaceutical-industry-marketing-and-its-influence-on-physicians-and-patients-85899439814 (accessed March 2015).

The Tetrad Project (2014), "Tetrad IV", available at: www.phil.cmu.edu/projects/tetrad/ (accessed March 2015).

Wolfe, M.J. and Crotts, J.C. (2011), "Marketing mix modeling for the tourism industry: a best practices approach", *International Journal of Tourism Sciences*, Vol. 11 No. 1, pp. 1-15.

Corresponding author
Dominique Haughton can be contacted at: dhaughton@bentley.edu