METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
X, 2009, str. 131 - 141

PREDICTIVE MARKETING MIX MODELLING IN FMCG PACKAGED FOOD CATEGORY INCLUDING PRICE AND ADVERTISING IMPACT

Tomasz Kolanowski
Wydział Badań i Analiz, UMWW
e-mail:T.kolanowski@umww.com

Abstract: Case study describing estimation of Marketing Mix econometric sales modelling on packaged food fast moving consumer goods product. Model bases on two step modelling approach, using ordinary least squares method. Adstock data transformation has been used to evaluate advertising impact with time distributed lag. Paper presents non-linear relations between market distribution, consumer price levels and advertising as well as 52 weeks forward sales forecast accuracy evaluation.

Keywords: Marketing Mix Modelling, Advertising, Sales Forecast, Fast Moving Consumer Goods (FMCG).

INTRODUCTION AND THEORETICAL APPROACH

The Following paper illustrates the estimation and evaluation of Marketing Mix econometric sales modelling on FMCG product. This technique is recommended in brand managing (Rószkowska 2002) and particularly in advertising management (Beliczyński 2007) and become very popular in recent years. This kind of modelling evaluates impact of marketing activities (independent variables) on sales of the product (dependent variable). As in every modelling, the key dependent factors have to be included into model. There are many theoretical frameworks identifying the most impactful marketing factors, but in this paper the most popular will be used. The 4P's theory was first formulated in 1960's (McCarthy & Shapiro 1975) and since this time very has been very well described and evaluated (Kotler 1999). This theory identifies key factors as Product, Price, Place and Promotion (4P). Product represents all physical features like size, packaging or quality. Place describes all issues related to point of sales: size of stores, number of stores or supplies. Price includes regular product price but also all discounts and price promotions. Promotion represents all kind of advertising, public relations and merchandising.

One of the most difficult marketing objectives for marketing mix modelling is evaluating the effect of advertising. This part of promotion is usually the highest position in brand marketing budget (Zyman 1999) therefore should be included in proper modelling.

The most common method for advertising modelling is called Adstock which is similar to time distributed lag. The transformation represents fact that effect of advertising impact can be observed long after end of advertising campaign.

$$A_t = T_t + \lambda A_{t-1},\ t = 1,\ldots,n \tag{1}$$

Where A_t is the Adstock at time t, T_t is the value of the advertising variable at time t and λ is the 'retention' or lag weight parameter. Inclusion of the A_{t-1} term, imparts an infinite lag structure to this model, with the effect of the first Adstock term, approaching zero, as t tends to infinity.

Adstock parameter λ can be interpreted in the models as percentage of effectively remembered ad contact from previous week plus contacts from current one (Broadbent, 1979). In literature it is frequently called "Ad Retention", there are also two other popular terms which are describing the same phenomenon: "Decay Factor" simply equals $1-\lambda$ and "Advertising Half-Life" which represents time needed to halve advertising effect. All three coefficients can be simply recalculated to each other.

There are no strict norms for ad retention and there are two popular theoretical approaches. In the first, the predefined lambda is determined for the model, (usually 80% or 90%) probably based on classical Emshoff and Mercer research (1970). The other recommended method is to set this parameter during modelling (Hanssens 2003). Some newer academic analyses suggest the most typical half-life range to be around 7-12 weeks (Leone 1995), approximately equivalent of λ = 90%-95%.

Adstock is only a mathematical transformation; its real-world nature is still a matter of discussion. Generally there are two opinions; one is that Adstock represents forgetting ad contacts with time, the other that it is caused by behavioural habit (Jones, 2002) created by advertising.

Advertising can be divided into two areas: creation and media planning (Czarnecki & Korsak 2001). Media planning focus on finding best campaign timing, most effective media and the optimal investment levels (Sisstors & Bumba, 1996) . The model was designed to help the planning process, so it is focused on constant in time advertising performance, rather than on a particular creative's effectiveness.

CASE DESCRIPTION

The analysed brand is owned by international corporation mainly focused on processed food market. The product belongs to broad salty snacks category. The brand was introduced to the UK market two years before analysis and still is in its early life stage.

THE DATA

Because the analysis was needed urgently the following data was chosen:

Sales in Quantity – delivered by AC Nielsen from their store panel. The data represents sales in Units. Units should be interpreted as number of packages or kg of products. This distinction doesn't change interpretation, but helps shield clients' sensitive data. Sales Data provided on a weekly basis and represents sales in multiple stores.

Sales In Value - delivered by AC Nielsen from their store panel. Defined as the total amount of money (GBP) spent on the product at the cash points. Data provided on a weekly basis, aggregated on a national level, and represents sales in multiple stores.

ACV Percentage Distribution – All commodity value weighted retail distribution from AC Nielsen. The measure estimates the percentage of stores where the product can be purchased. Each store is weighted by its sales potential to reflect the real power of distribution. This is illustrated by an example; suppose that town X has 10 stores. However, one of them is a megastore and represents 50% of all trade in the town. If our brand was present only in this store, its ACV Distribution will be 50%. (Curry 1993). Data provided on a weekly basis, aggregated on national level and represents sales in multiple stores.

Ad Spend - delivered by Nielsen Media Research through AdDynamix. Data provided from an independent monitoring system. Methodology based on monitoring of actually broadcasted/issued advertisements across different media. Each advertisement is attributed to a specific brand, and its cost is estimated. The costs are provided based on the rate-card, which means an official published cost of the placement. The real costs are significantly lower and depend on individually negotiated discounts for clients. For a mid-size client in UK, we estimate a discount of 35%-45% off the rate-card price. The discount is constant year to year. Data provided on a weekly basis and aggregated on a national level.

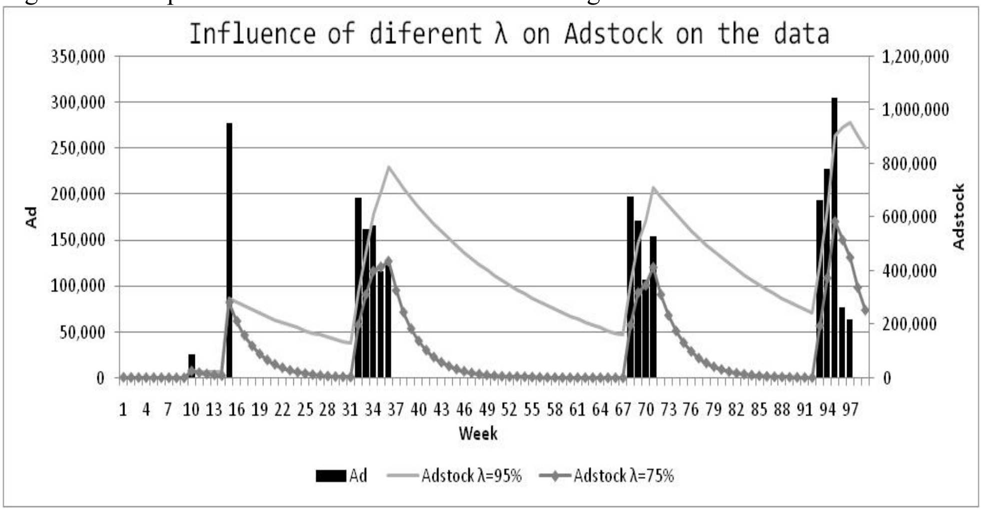
No other data had been used in modelling.

To find the representation of price, the sales (in value), were divided by sales (in volume) in each week. The result was an average purchase price for each week. This number helped to evaluate also effect of in-store promotions (like 2-for-1 offers), when display price is on the same level, but quantity of product increases.

This recalculated indicator also addressed multiple display prices in different stores within the same week. The indicators worked on total aggregated numbers so for example, in a certain week 1000 units were sold for total amount of 800 GBP, the calculated average price was 80p. This indicator will be called "price" later on.

The advertising date ware transformed using ad stock method. As previously mentioned, the most typical λ is in range 0.8 – 0.95. However, it is noted that λ varies strongly on different products and advertisements. In some product categories, it may be lower than 0.5, especially where direct response is a modelled variable. Therefore, in this case the best fitting λ was evaluated from whole theoretical range (0,1). There were five advertising bursts (campaigns) with different pressure during modelling period. According to different λ, their influence on the model varied significantly. (Fig 1)

Figure 1. Example of Adstock transformation according to λ.



Source: own calculations

MODELLING PROCESS

The first model tried to build simple single equation model, and failed. Error had a tendency to rise with time. It was for this reason, building a two-step model was considered.

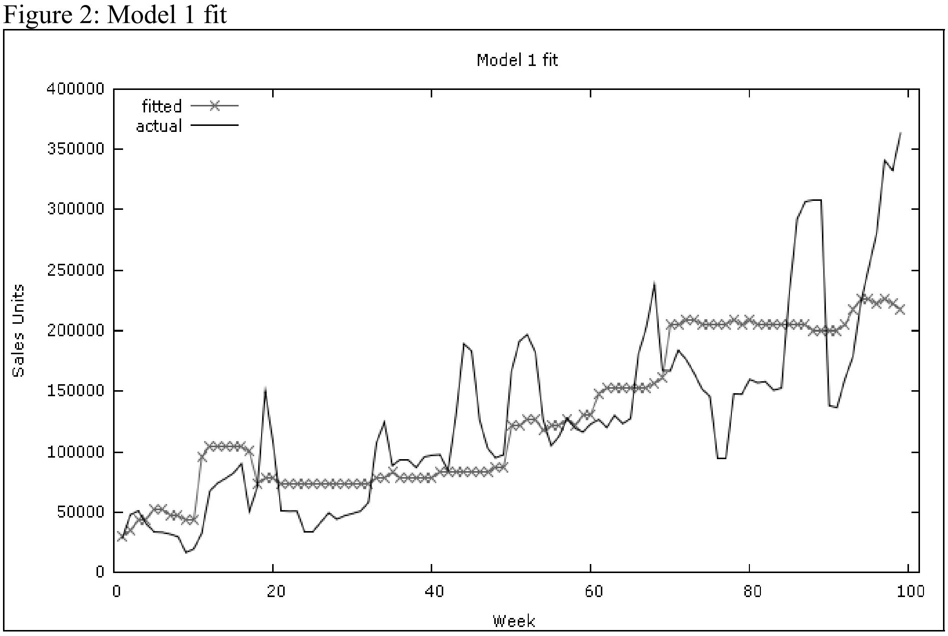
Sales were almost constantly increasing because brand was relatively new and still building its market distribution. ACV distribution was the highest correlated factor. The first model was built to evaluate regression only from ACV on sales.

$$S_d = \beta_1 * ACV \tag{2}$$

where: S_d = Estimated sales in units modelled by distribution, S_a = Actual Sales units

The ordinal least squares method was used. During modelling, the constant was removed because it had a low influence, and was statically not significant. It has also clear economic interpretation; when distribution (ACV) is equal zero, the product is not available in the market, therefore sales is null.

Figure 2: Model 1 fit



Source: own calculations

Residuals from the first stage model were transformed to relative fractional values with this simple formula.

$$S_r = \frac{S_a}{S_d} \tag{3}$$

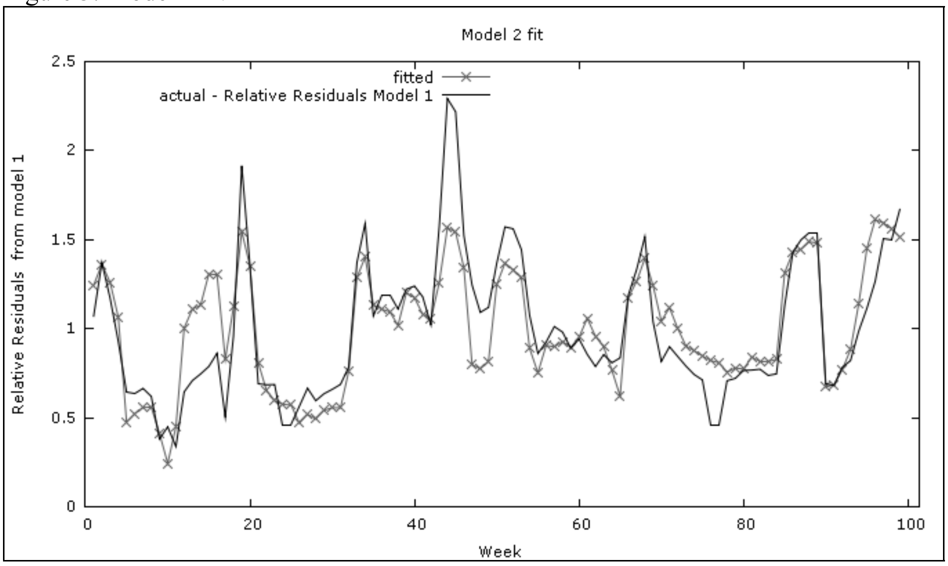
where: S_a = Actual Sales units, S_r = Relative residuals from model 1

The transformed residuals became a new time series for modelling in stage two. Again, the ordinal least squares method was used. The independent variables, Adstock and price, were used. Adstock λ was tested to find the best fit model – process was made by overall model R^2, maximising by changing λ. (Fig. 2) Finally, the λ=96.1% was found as the best fit (equivalent of half-life 17.42 weeks). This level was also the most statistically significant (fig. 3).

$$S_r = \beta_2 \cdot AdSt + \beta_3 \cdot \Pr ice + \beta_4 \tag{4}$$

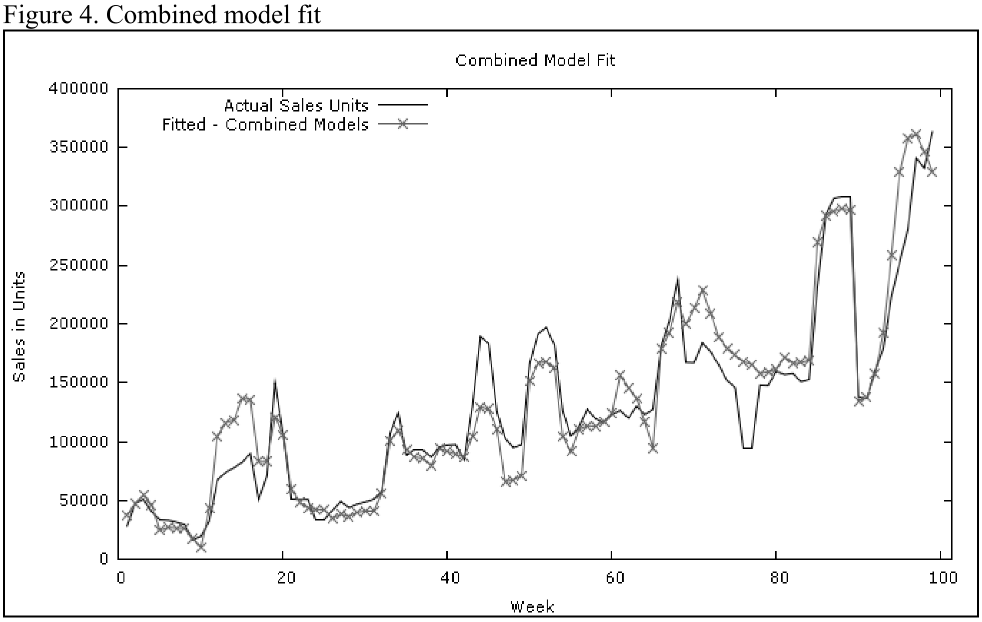
Where S_r = Relative residuals from model 1, S_a = Actual Sales units (measured), *AdSt*= Adstock, *Price* = average product price

Figure 3. Model 2 fit



Source: own calculations

Figure 4. Combined model fit



Source: own calculations

The models were combined to make calculation easier. The formula allows calculation of estimated sales in a single equation. The overall effect of applying two step models estimated sales on satisfying level (Fig 4). There were some problems with accuracy in three periods (around week 15, 50 and 75), but the general shape fitted to all local maximums and minimums.

$$S_Q = \beta_1 \cdot ACV\left(\beta_2 \cdot AdSt + \beta_3 \cdot Price + \beta_4\right) \tag{5}$$

where S_Q=Estimated sales in quantity (units)

It is worth noting that if number of units sold is estimated, and price for average unit is known, the next step would be to multiply equation by sides by price to gain value of sales. This transformation is acceptable (Hanssens 2003) in market mix modelling.

$$S_V = \beta_1 \cdot ACV\left(\beta_2 \cdot AdSt + \beta_3 \cdot Price + \beta_4\right) \cdot Price \tag{6}$$

where S_V= Estimated sales in value (S_Q*Price)

To summarise the modelling process the two ordeal least squares regression were used. The first model estimates sales in units by β_1 parameter its R^2 equals 0.89. The second model evaluates effect of Adstocked ad activity β_2 and price β_3 on residuals from the first one. Its adjusted R^2 equals 0.72 and unadjusted 0.71. All coefficients including constant are highly statistically significant. (Tab 1)

Table 1. Equations parameters

Variable	Distribution	Adstock	Price	Constant
Coefficients	β1	β2	β3	β4
Value	4352.015595	4.60806E-07	-2.458696295	3.005917229
S(β)	155.1746507	7.98862E-08	0.179689217	0.169324788
stat t	28.04591844	5.768273588	13.68304863	17.75237558
p-value 2-sig	1.93928E-60	4.66054E-08	6.88719E-28	3.47167E-38
	***	***	***	***
p-value1-sig	9.6964E-61	2.33027E-08	3.44359E-28	1.73583E-38
	***	***	***	***

Source: own calculations)

MARKETING INTERPRETATION AND DISCUSSION

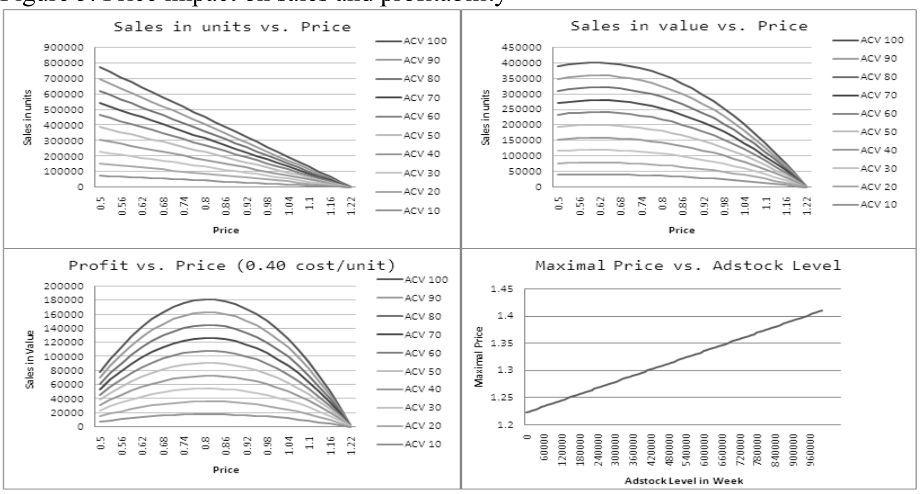
The model has quite an intuitive interpretation. For each percent point of weighted market distribution (ACV), the theoretical constant unit sales were generated. Then this theoretical level was modified by advertising and price factor. The advertising and price worked proportionally to build distribution. Simply speaking, the more people that are able to purchase the product, the more impactful price and advertising are. Price had a negative effect (higher price lower sales) and Adstocked advertising had a positive effect (more advertising, bigger sales). Advertising seemed to work very long-term, even after 17 weeks its effect was only halved. However, it is coherent with Leone (1995) works.

The Price, ACV and Adstock were highly connected and in the equation, and interact with themselves. Charts were quite difficult to read. The most important factor was ACV distribution, therefore wherever it was possible, different levels of ACV were charted.

The unit sales were related to price in a linear way. Lower price generated higher sales. When analysed with zero advertising the level 1.22 was the highest price that generated positive sales. However, for sales expressed in value, effects were non-linear. Maximal sales were generated when the price was set to 0.61 (no advertising).

Another interesting question is ‘What is the most profitable price level’? We did not have data regarding company cost levels per unit (manufacturing, packaging, transport etc) but it was assumed at 0.40. This was a little lower than the lowest observed price in analysed period. This assumption was done to show how the model can be used by clients. In this case, the optimal price level was 0.81 (without advertising) (Fig 5).

Figure 5. Price impact on sales and profitability



Source: own calculations

The effect of increasing price elasticity (because of advertising), was also observed. With null ad pressure the maximal price was 1.22, but with strong Adstock presence it rose above 1.4. The same shift occurred on optimal price for profitability, it rose above 0.91 from the effect of strong ad pressure. The effect of increasing customer’s price elasticity as a result of advertising was identified as the most profitable on UK market (Binet & Field 2007) in recent IPA case study base.

Advertising works in long time, its effect despairs with ratio 96.1% per week. Product familiarity had a key effect on product sales and advertisements were well remembered. Each invested amount of money paid itself back 88% of its

Adstock in the first year, and the rest was present in the following years. The amount of generated sales was proportional to market distribution level ACV. There was also the effect of shifting consumers' price elasticity because of advertising. People seem to accept higher prices when the effect of advertising is present.

FORECASTING AND MODEL EVALUATION

Unfortunately because of external factors the model had never been used in planning. But when the following year's data became available, it was great opportunity to check model accuracy on real values. Again all data was taken from the same sources but further modelling was not made. The proposed approach belongs to ex-post methods and MAPE parameter has been used to evaluate model performance (Witkowska 2006).

This exercise checked how accurate the findings from the first two years modelling were, for product sales dynamics in the third year. In the 52 weeks, new promotions and price levels had been created, and also a new ad campaign was broadcast. There was also a significant increase in market penetration (ACV) from 50% on beginning to 62% by the end of forecast period. This meant that all significant factors were not stable (which makes a good testing period).

The model seemed to be useful in long-term predictions. The 52 forecasted weeks are the usual annual marketing planning period. The model predicted the whole period sales with 3.7% error and its mean absolute percentage weekly error (MAPE) was equal 12.4%. The model was also stable in time and had no tendency to accumulate error in time (fig. 6). Forecasts from the end period were not worse (and sometimes were better) than those from early data closer to the modelling period. (Table 2)

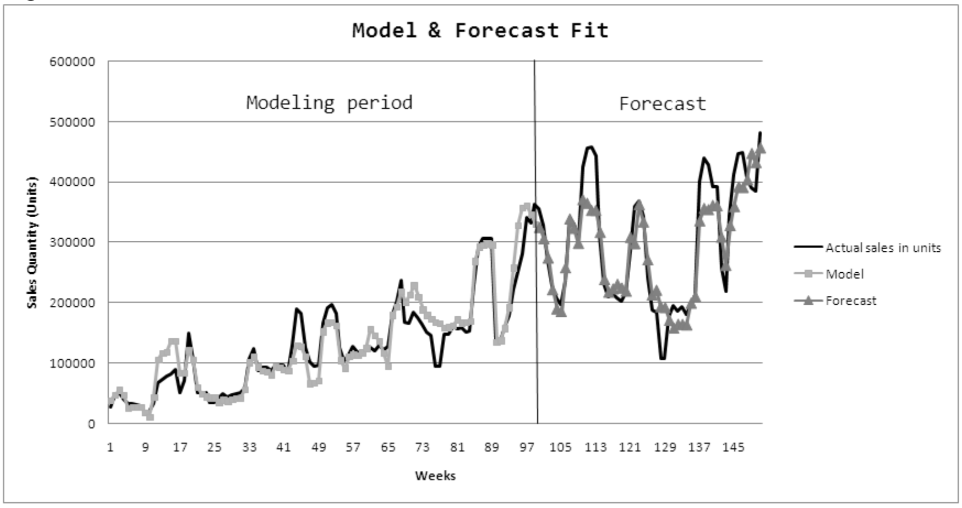
Table 2. Forecast Evaluation

13 week quarters	Q1	Q2	Q3	Q4
Absolute % Error - Total Period	8.1%	1.4%	1.3%	3.1%
MAPE weekly	8.2%	7.2%	22.2%	11.5%

Source: own calculations

It should be noted that all local minimums and maximums were properly indicated. Only the strength of sales response were underestimated in the lowest price (0.7) periods, around week 111 and 138. The prices at this level were almost not present in modelling period. This may indicate a nonlinear price response curve application for the following year's modelling.

Figure 6: Model evaluation



Source: own calculations

CONCLUSIONS

The case showed that market mix modelling can be successfully applied to brands, even when they are in early life stages or a dynamic growth period. The publicly available market monitoring data like AC NIELSEN can be a base for modelling and in situations where a company does not gather its own marketing data, this is sufficient enough to understand market dynamics. The most important conclusion is that properly built marketing mix models are able to forecast brand sales performance. This forecast can be accurate enough to set advertising strategies and marketing activities over a one-year period. This time scale allows brand managers to build their annual plans more precisely and improve their return on investment.

BIBLIOGRAPHY:

Beliczyński, Jan. (2007) Metody planowania i kontroli w zarządzaniu reklamą, Wydawnictwo Akademii Ekonomicznej w Krakowie,, Wydawnictwo Akademii Ekonomicznej w Krakowie.

Binet, L and Field, P (2007) Marketing in the Era of Accountability: Identifying the Marketing Practices and Metrics That Truly Increase Profitability, World Advertising Research Center 2007.

Broadbent, S. (1979) One Way TV Advertisements Work, Journal of the Market Research Society Vol. 23 no.3.

Curry, David J. (1993) New Marketing Research Systems: How to Use Strategic Database Information For Better Marketing Decisions, Published by Wiley, John & Sons, Incorporated.
Czarnecki, A, Korsak, R (2001) Planowanie mediów w kampaniach reklamowych., Published by PWE.
Emshoff, J. R. and Mercer A. (1970) Aggregated Models of Consumer Purchases, Journal of the Royal Statistical Society, Series A, 133:1, 14-32.
Hanssens, D. M. Parsons, L. J. Schultz R. L. (2003) Market Response Models: Econometric and Time Series Analysis, Published by Springer.
Jones, John Philip (2001) The Ultimate Secrets of Advertising, Edition: 3, Published by Sage Publications.
Kotler (1999) Marketing. Analiza, planowanie, wdrażanie i kontrola. Published by Warszawa.
Leone, R.P. (1995) Generalizing what is known about temporal aggregation and advertising carry-over, Marketing Science, 14, G141-G150.
McCarthy, E. J and Shapiro S. J. (1975) Basic Marketing, published by Irwin-Dorsey.
Rószkiewicz, Małgorzata (2002): Narzędzia statystyczne w analizach marketingowych, published by C.H.Beck.
Sissors, J. Z. Bumba, L. (1996) Advertising Media Planning, Published by McGraw-Hill.
Witkowska, Dorota (2006): Podstawy ekonometrii i teorii prognozowania, Published by Oficyna Ekonomiczna.
Zyman, Sergio (1999) The End of Marketing as We Know it Published by Harper-Business.

Predyktywne modelowanie marketing mix na przykładzie marki z kategorii dóbr szybkozbywalnych opakowanej żywnosci z uwzględnieniem wpływu reklamy i ceny

Streszczenie: Artykuł ilustrujący estymacje modelu ekonometrycznego sprzedaży na przykładzie produktu z kategorii dóbr szybkozbywalnych opakowanej żywności. Model opiera się na podejściu dwukrokowym z zastosowaniem metody najmniejszych kwadratów. Do modelowania wpływu reklamy z rozłożonym opóźnieniem zastosowano transformacje Adstock. Artykuł opisuje nieliniowe zależności pomiędzy dystrybucja produktu na rynku, ceną produktu dla końcowego konsumenta i nakładami reklamowymi jak również ocenę trafności prognozy modelu dokonaną na 52 tygodnie naprzód.

Słowa kluczowe: Modelowanie Marketing Mix, Reklama, Prognoza sprzedaży, Dobra szybkozbywalne (FMCG)