April 02, 2015

SYSTEM AND METHOD FOR SPECIFICATION OF A MARKETING MIX ECONOMETRIC MODEL USING FEEDBACK FROM A DIGITAL ATTRIBUTION SYSTEM

Timothy McDonough

SYSTEM AND METHOD FOR SPECIFICATION OF A MARKETING MIX ECONOMETRIC MODEL USING FEEDBACK FROM A DIGITAL ATTRIBUTION SYSTEM

Accuracy of econometric marketing mix model backcasts and forecasts depends on the ability to include only input factors that have strong correlation with the dependent variable used as the output. For models that use digital channel inputs such as paid search, display ads, or others, the only measurements typically available are highly aggregated and often have weak correlation with output dependent variables.

Deployment of digital marketing treatments can be parsed manually by hand, but it is only theoretical guesswork or intuition that guides the choice of input specifications of models used.

This disclosure provides a method including analyzing outputs from a digital attribution system, determining potential input factor attributes that are highly correlated to output dependent variables, and generating a specification for an econometric marketing mix model, the specification including the input factor attributes thus determined. The result is a marketing mix model specification that is more accurate for backcasting and forecasting than what prior approaches can produce. In some implementations, the method may be implemented using a system comprising a processor and/or one or more computer-readable storage media that stores instructions executable by the processor to perform the operations described herein.

This disclosure can provide many advantages, according to various implementations. For example, a more accurate econometric model enables advertisers to better understand the impact of their advertising dollars spent to influence their customers' purchasing behavior and thus allocation of those expenditures is made more efficient.

Econometric marketing mix models consist of inputs (independent variables) and outputs (dependent variables). FIG. 1 shows a diagrammatic representation of a system for specification of a marketing mix econometric model using feedback from a digital attribution system according to an illustrative implementation. As illustrated in FIG. 1, input variables of a marketing mix model 120 may include measures of marketing treatments such as digital channels (e.g., social network, display, paid search, display, affiliate, email, etc.) and non-digital "traditional" offline advertising channels (e.g., TV, radio, out of home, direct mail, print, events, sponsorships, etc.). Other input variables representing non-marketing factors may include macroeconomic measures such as unemployment, interest rates, geo-political events, etc.

Some previous models include broad specifications for digital variables such as highly aggregated measures of paid search, display or other digital channel impressions or spend on those channels. These broad specifications result in the introduction of errors (differences between actual and modeled output) that mask the true influence of digital and other input factors.

Some previous approaches to this problem involve the use of guessing which attributes of digital channel inputs are more highly correlated with the modeled output. Such attributes include keyword search terms or placements of display ads on certain properties (web sites) or positions on a page. As those skilled in the art can appreciate, even if one may, however unlikely, correctly guess which attributes of digital channel inputs may be more correlated with the modeled output, it can be extremely difficult to ascertain the effectiveness or influence of an offline ad campaign on an online user's behavior.

As an example, a previous marketing mix model might take the form of an additive relationship between an output variable such as SALES and a traditional marketing channel such as TV and a digital advertising channel such as paid search (PS). A specification for this previous marketing mix model may be as follows:

$$SALES = \beta_0 + \beta_1 TV + \beta_2 PS + e$$

where,

β_0 = an intercept term (SALES that would occur without TV or PS deployments)

β_1 = coefficient on TV deployment

β_2 = coefficient on Paid Search deployment

e = error term (difference between actual and predicted SALES)

In this case, both input variables “TV” and “PS” are broadly defined and highly aggregated at the channel level. Thus, it is impossible to determine how and which of the input variables is more correlated to the output variable “SALES” and/or what attribute(s) of an input variable actually contributed to the correlation between the particular input variable and the output variable.

Feedback 115 from a digital attribution system 110 can be used to identify particular digital channel attributes that contribute the most explanatory power to the hypothesized causal relationship between digital channel inputs and key performance indicator outputs. Feedback 115 represents outputs from a high performance digital attribution system. High performance digital attribution systems can correctly and accurately capture the causal relationship among different channels, including online channels and offline channels, and conversions. Thus, it is possible to construct input variables for the econometric marketing mix models with very narrowly defined attributes that scientifically assures the highest possible correlation between input variables and output variables. Such a method of constructing a model specification is neither arbitrary, dependent on economic theory, experience level of the modeler, nor intuition. More specifically, it is now possible to construct input variables for the econometric marketing mix models with sub-channel attributes.

Digital attribution system 110 may provide feedback 115 indicating that within the Paid Search channel, a first set (or Group 1) of search terms and a second set (or Group 2) of search terms may have a higher correlation to SALES (the output variable in this example) than other paid search terms within the Paid Search channel (an input variable in this example). With this knowledge, sub-channel-level variables for the Paid Search channel can be constructed. Below provides an example of a new econometric marketing mix model with two (or more) terms related to the Paid Search channel:

$$SALES = \beta_0 + \beta_1 TV + \beta_2 PS_2 + \beta_3 PS_3 + e$$

where,

β_0 = an intercept term (SALES that would occur without TV or PS deployments)

β_1 = coefficient on TV deployment

β_2 = coefficient on Paid Search deployment for one set of paid search terms

β_3 = coefficient on Paid Search deployment for another set of paid search terms

e = error term (difference between actual and predicted SALES)

The technique can be extended indefinitely to any number of variables until statistical considerations limit the expansion of the model specification (number and definition of variables).

FIG. 2 depicts a flow diagram illustrating a method 200 for specification of a marketing mix econometric model using feedback from a digital attribution system according to an illustrative implementation. Method 200 may include analyzing outputs (e.g., feedback 115 shown in FIG. 1) from a digital attribution system (e.g., digital attribution system 110 of FIG. 1) at step 205, determining potential input factor attributes (e.g., PS2 and PS3, in the econometric marketing mix model described above) that are highly correlated to output dependent variables (e.g., the SALES variable in the econometric marketing mix model described above) at step 210, and generating a specification for an econometric marketing mix model at step 215, the specification including the input factor attributes thus determined. The result is a marketing mix model specification that is more accurate for backcasting and forecasting than what prior approaches can produce.

FIG. 3 depicts a plot diagram illustrating results comparing results provided by marketing mix models with a standard specification and an enhanced specification according to an illustrative implementation. Even though the enhanced specification focused on attributes for the display channel only, it can be seen from the plot diagram that using the outputs from a digital attribution system to identify highly corrected attributes for the display channel has significantly improved the results of the econometric marketing mix model.

FIGURE 4 depicts a diagrammatic representation of an illustrative data processing system. As shown in FIGURE 4, data processing system 400 may include one or more central processing units (CPU) or processors 401 coupled to one or more user input/output (I/O) devices 402 and memory devices 403. Examples of I/O devices 402 may include, but are not limited to, keyboards, displays, monitors, touch screens, printers, electronic pointing devices such as mice, trackballs, styluses, touch pads, or the like. Examples of memory devices 403 may include, but are not limited to, hard drives (HDs), magnetic disk drives, optical disk drives, magnetic cassettes, tape drives, flash memory cards, random access memories (RAMs), read-only memories (ROMs), smart cards, etc. Data processing system 400 can be coupled to display 406, information device 407 and various peripheral devices (not shown), such as printers, plotters, speakers, etc. through I/O devices 402. Data processing system 400 may also be coupled to external computers or other devices through network interface 404, wireless transceiver 405, or other means that is coupled to a network such as a local area network (LAN), wide area network (WAN), or the Internet.

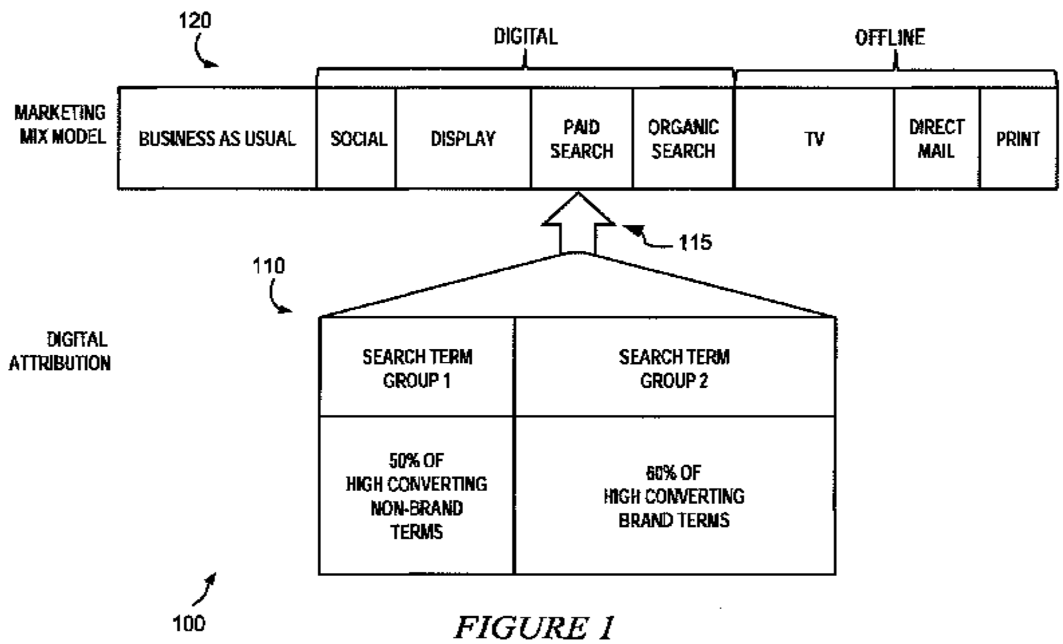

FIGURE 1

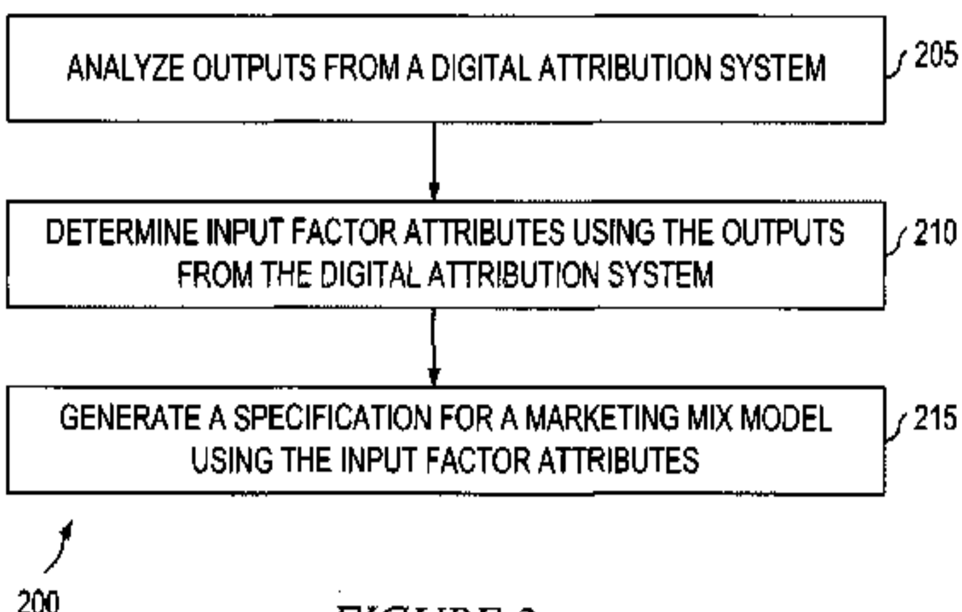

FIGURE 2

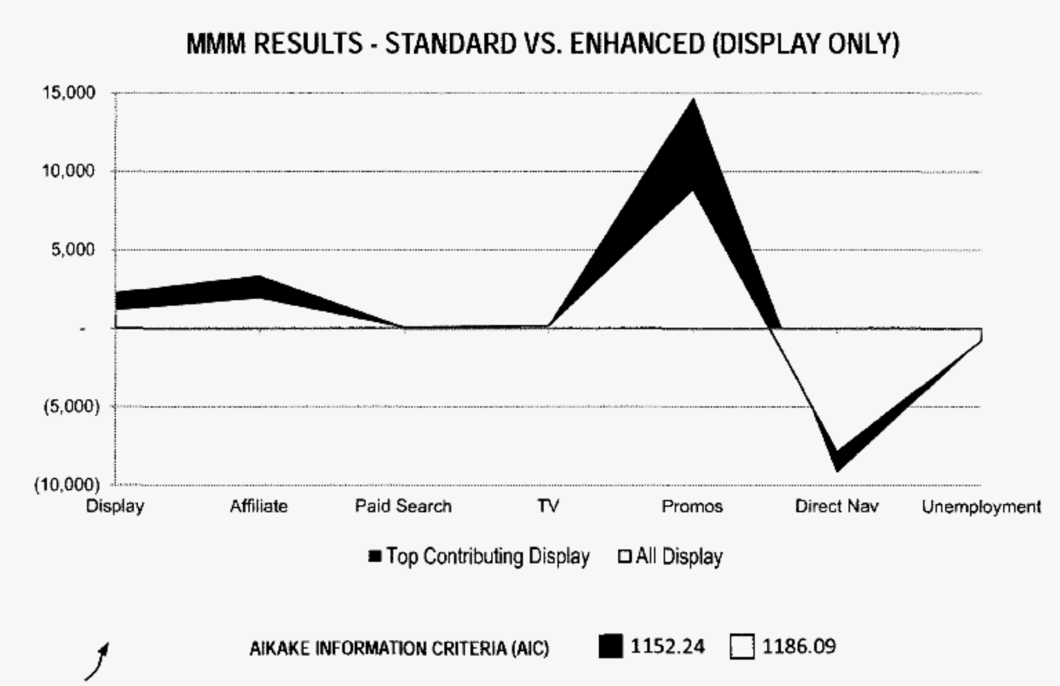

FIGURE 3

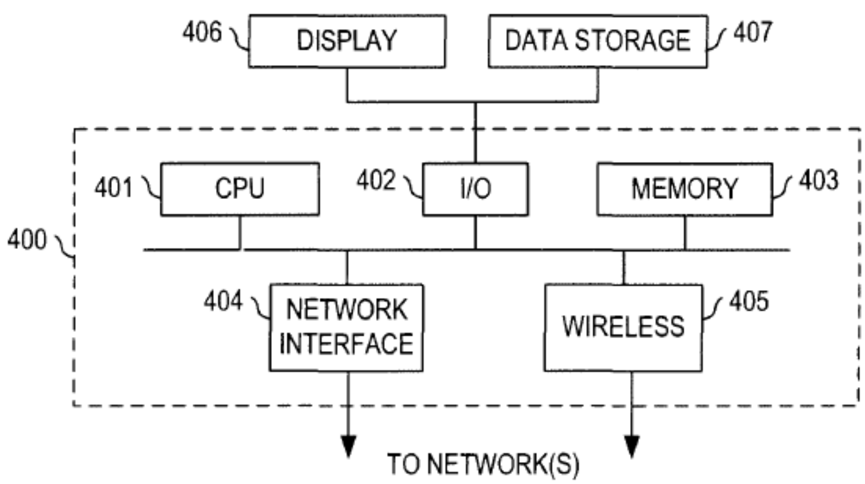

FIGURE 4