International Journal of Research in Marketing 33 (2016) 457–474

Contents lists available at ScienceDirect

IJRM
International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Mapping the customer journey: Lessons learned from graph-based online attribution modeling

Eva Anderl[a,*], Ingo Becker[b], Florian von Wangenheim[b], Jan Hendrik Schumann[a]

[a] Universität Passau, Innstr. 27, 94032 Passau, Germany
[b] ETH Zurich, Weinbergstr. 56/58, 8092 Zurich, Switzerland

ARTICLE INFO

Article history:
First received on April 24, 2014 and was under review for 12½ months
Available online 19 March 2016

Keywords:
Online advertising
Attribution
Marketing models
Markov models
Multichannel

ABSTRACT

Advertisers employ various channels to reach customers over the Internet, who often get in touch with multiple channels along their "customer journey." However, evaluating the degree to which each channel contributes to marketing success and the ways in which channels influence one another remains challenging. Although advanced attribution models have been introduced in academia and practice alike, generalizable insights on channel effectiveness in multichannel settings, and on the interplay of channels, are still lacking. In response, the authors introduce a novel attribution framework reflecting the sequential nature of customer paths as first- and higher-order Markov walks. Applying this framework to four large customer-level data sets from various industries, each entailing at least seven distinct online channels, allows for deriving empirical generalizations and industry-related insights. The results show substantial differences from currently applied heuristics such as last click attribution, confirming and refining previous research on singular data sets. Moreover, the authors identify idiosyncratic channel preferences (carryover) and interaction effects both within and across channel categories (spillover). In this way, the study can help advertisers develop integrated online marketing strategies.

1. Introduction

Online advertising is essential to the promotional mix of many industries (Raman, Mantrala, Sridhar, & Tang, 2012). Today, advertisers employ a variety of online marketing channels[1] to reach potential customers, including paid search and display marketing, as well as e-mail, retargeted displays, affiliates, price comparison, and social media advertising. At the same time, customers visit the advertisers' websites on their own initiative—for instance, by directly typing in the related web address. Using various channels, many customers visit company websites multiple times before concluding a purchase transaction (Li & Kannan, 2014). Previous visits may influence the users' subsequent visits, such that the customer may return to a website through the same channel (carryover effects) or through different channels (spillover effects). Given the proliferation of online channels and the complexity of customer journeys,[2] measuring the degree to which each channel actually contributes to a company's success is demanding.

Despite the widespread and ongoing practice of many advertisers to apply comparatively simple heuristics (e.g., last click attribution), such that the value is attributed solely to the marketing channel directly preceding the conversion (The CMO Club, & Visual IQ,

* Corresponding author at: Universität Passau, Innstr. 27, 94032 Passau, Germany. Tel.: +49 851 509 2421; fax: +49 851 509 2422.
E-mail addresses: eva.anderl@uni-passau.de (E. Anderl), ibecker@student.ethz.ch (I. Becker), fwangenheim@ethz.ch (F. von Wangenheim), jan.schumann@uni-passau.de (J.H. Schumann).

[1] In this study, we use the term "online marketing channels" as an umbrella phrase referring to various online marketing instruments, including search engine advertising, display, or social media advertising.

[2] We define an online customer journey of an individual customer as including all touch points over all online marketing channels preceding a potential purchase decision that lead to a visit of an advertiser's website.

http://dx.doi.org/10.1016/j.ijresmar.2016.03.001

Inc, 2014), this challenge of attributing credit to different channels (Neslin & Shankar, 2009) has recently begun to receive increased attention in academia and practice alike (Berman, 2015). Academics have proposed a variety of substantiated analytical attribution frameworks, including logistic regression models (Shao & Li, 2011), game theory-based approaches (Berman, 2015; Dalessandro, Perlich, Stitelman, & Provost, 2012), Bayesian models (Li & Kannan, 2014), mutually exciting point process models (Xu, Duan, & Whinston, 2014), VAR models (Kireyev, Pauwels, & Gupta, 2016), and hidden Markov models (Abhishek, Fader, & Hosanagar, 2015). Furthermore, several industry players such as Adometry (Google), Convertro (AOL), or VisualIQ have introduced a range of attribution methodologies (Moffett, 2014). Even though sophisticated attribution methods become accessible to a broader audience, and marketing executives call for such performance measures (Econsultancy, 2012), in practice the full browsing history of a user is rarely taken into account when calculating channel effectiveness in multichannel settings (Li & Kannan, 2014).

Based on a survey among marketers using advanced interactive attribution offerings, Osur (2012) reports that the two most widely applied objectives of attribution software are to "measure the value and performance of digital channels" and to "measure how one digital channel affects the performance of another [channel]" (p. 4). In this research, we address both topics, but we extend the issue further by trying to identify generalizable answers to these challenges. For advertisers, it is valuable to know what insights from attribution apply in a company-specific context and what insights may be generalized across companies (industries). Such insights can help to better explain factual channel effectiveness and can also shed light on the interplay of channels in multi-touch environments. Marketing scholars specifically call for further research using customer-level path data across several firms and industries to determine spillover effects in a more generalizable way (Li & Kannan, 2014). Empirical generalizations are important for both theory generation and evaluation and can provide valuable guidance to managers (Kamakura, Kopalle, & Lehmann, 2014). For instance, if advertisers anticipate idiosyncratic channel preferences among some users on their path to purchase, the advertisers could transfer this knowledge into a more adequate channel selection.

To achieve our research objectives, we needed to develop and apply an attribution mechanism with the capability of determining the effectiveness of individual marketing channels and deriving insights on the interplay of channels in a multichannel environment across very different data sets and conditions. In particular, we suggest an attribution framework based on Markovian graph-based data mining techniques, extending an approach originally developed in the context of search engine marketing (Archak, Mirrokni, & Muthukrishnan, 2010). We model individual-level multichannel customer journeys as first- and higher-order Markov graphs, using a property that we call removal effect to determine the contribution of online channels and channel sequences. The graph-based structure of our model reflects the sequential nature of customer journeys, enabling insights into the interplay of channels. Applying this framework to four large, real-world customer-level data sets from three different industries enables us to derive both cross-industry generalizations and industry-specific findings. In doing so, we make the following contributions:

First, we contribute novel insights into online marketing effectiveness of single channels within a multichannel setting. We estimate our graph-based framework on four data sets from different industries and compare the results against two well-known heuristic attribution techniques, namely, first and last click attribution, as well as two logit models. Prior research indicates that heuristic approaches of attributing conversion to the very last (or first) click can produce incorrect conclusions (Abhishek et al., 2015; Li & Kannan, 2014; Xu et al., 2014). A comparison of our results across four data sets enables us to confirm and refine these results and move toward empirical generalizations. We find that firm-initiated channels, where the advertiser initiates the marketing communication (Bowman & Narayandas, 2001; Wiesel, Pauwels, & Arts, 2011), are consistently undervalued by the heuristic attribution approaches. For customer-initiated channels, which are triggered by potential customers, on their own initiative (Wiesel et al., 2011), the contribution of paid search and direct type-ins is consistently overestimated by the last click approach. For other customer-initiated channels, additional factors such as industry and brand characteristics seem to play a role.

Second, higher-order models enable us to shed light on the interplay of channels in a multichannel setting. By comparing the results of our attribution framework across data sets, we generalize findings from prior literature indicating that a majority of channels exhibits idiosyncratic channel carryover (Li & Kannan, 2014). Furthermore, we observe spillover effects both within and between channel categories. Customer-initiated channels show substantial removal effects if they are followed by other customer-initiated channels, whereas spillover effects between firm-initiated channels are, by and large, negligible.

Third, we propose a novel variant that adds to existing advanced attribution modeling techniques (Abhishek et al., 2015; Berman, 2015; De Haan, Wiesel, & Pauwels, 2016; Kireyev et al., 2016; Li & Kannan, 2014; Xu et al., 2014) by representing customer path data as first- and higher-order Markov walks. This graph-based approach, adapted from research on paid search (Archak et al., 2010), represents a useful addition to the emerging attribution literature. Whereas first- and higher-order models offer support in measuring channel contribution in a multichannel setting, higher-order models, in particular, allow investigation of channel sequences and spillovers between channels.

Finally, our framework is beneficial in the approach to several explicit problems that online marketers confront. For instance, the framework may help to calibrate online channel budgets and move toward an improved budget allocation, even though endogeneity issues and different channel elasticities can make budget allocation difficult. Furthermore, using our framework, advertisers can more accurately calculate the conversion probability of a customer, given his or her previous customer journey. This information can be used to support state-of-the-art applications such as real-time bidding decisions in advertising exchanges.

2. Research background

Academic research on attribution and the interplay of online channels has only recently gained momentum. Jordan, Mahdian, Vassilvitskii, and Vee (2011) examine allocation decisions for publishers, using multiple attribution approaches, and derive optimal allocation and pricing rules for publishers who are selling advertising slots. In a study of the economic welfare consequences of the use of attribution technologies, Tucker (2012) finds evidence for more conversions at lower costs, due to the ability to systematically shift budget toward selected campaigns, but does not, however, disclose details on the attribution methodology.

Furthermore, academic studies address the online attribution problem: Shao and Li (2011) introduce two attribution approaches—a bagged logistic regression model and a simple probabilistic model. Building on their work, Dalessandro et al. (2012) propose a more complex, causally motivated attribution methodology based on cooperative game theory. Developing from the basis of simulated campaign data, they find that advertisers tend to assign credit to conversions that are driven by the users' volition to convert rather than on the factual influence of the advertisement. Focusing on the interplay between advertisers and publishers, Berman (2015) evaluates the impact of various incentive schemes and attribution methods on publishers' propensity to show advertisements and on the resulting profits of advertisers. He introduces an analytical model based on the Shapley value and compares it to the last click heuristic. Abhishek et al. (2015) suggest a dynamic hidden Markov model (HMM) that is based on individual consumer behavior and that captures a consumer's deliberation process along typical stages of the conversion funnel. They find that different channels affect consumers in different stages of their decision process. For example, display ads usually impact consumers early in the decision process, moving them from a disengaged state to activity or engagement. Li and Kannan (2014) propose a Bayesian model for measuring online channel consideration, visits, and purchases by using individual conversion path data and validating it in a field experiment. They use the estimated carryover and spillover effects to attribute conversion credit to different channels, and they find that the relative contributions of these channels are significantly different from last click attribution. By means of a mutually exciting point process model, Xu et al. (2014) calculate average conversion probabilities for different online advertising channels, showing that the conversion rate measure underestimates the effect of display ads, as compared to search ads. A multivariate time-series model based on aggregate data by Kireyev et al. (2016) analyzes attribution dynamics for display and for search advertising. They derive spillover effects from display toward search conversion; however, display ads also increase search clicks, thereby increasing costs for search engine advertising. Finally, De Haan et al. (2016) propose a structural vector autoregression (SVAR) model, also based on aggregate data, to determine the effectiveness of various offline and online advertising channels. Given that prior research relies on single data sets, such that findings may be company or industry specific, this study extends the existing literature on multichannel online advertising by applying a novel attribution framework based on Markov graphs on four data sets from three different industries.

3. Data

Our research is based on four clickstream data sets provided by online advertisers, in collaboration with a multichannel tracking provider. Clickstream data record each user's Internet activity and thus trace the navigational path the customer takes (Bucklin & Sismeiro, 2009). For each visit to the advertiser's website during the observation period, the data include detailed information about the source of the click and an exact timestamp. Clicks either represent a direct behavioral response to an advertising exposure or result from the user entering the advertiser's URL directly into the browser, so these sources comprise all online marketing channels, as well as direct type-ins. We also know for each visit whether it was followed by a conversion—in this case a purchase transaction. We use these data points to construct customer journeys that describe the click pattern of individual consumers across all online marketing channels and their purchase behavior. Thus, we not only track successful journeys ending with a conversion, but also journeys that never lead to a conversion, and we do so within a timeframe of thirty days from the last exposure. The data collection occurs at the cookie-level, such that we identify individual consumers—or more accurately, individual devices. The use of cookie data suffers a number of limitations, such as an inability to track multidevice usage or bias due to cookie deletion (Flosi, Fulgoni, & Vollman, 2013), yet cookies remain the industry standard for multichannel tracking (Tucker, 2012). We do not include information on offline marketing channels because measuring individual-level exposure to multiple offline media proves highly difficult in practice (Danaher & Dagger, 2013).

The advertisers that provide the data sets for this study operate in different industries: Data Set 1 was provided by an online travel agency. Data Sets 2, 3, and 4 originate from specialized online retailers selling apparel and luggage, respectively. All the advertisers in our sample address a broad audience and are pure online players, so we can exclude online/offline cross-channel effects. Each data set includes a minimum of 405,000 journeys per advertiser. Their average length is 1.3–2.5 contacts, and between 0.9% and 2.0% of all journeys lead to a successful conversion. In Table 1, we present detailed descriptions of the journeys in our data sets.

Table 1
Descriptions.

Description	Data Set 1	Data Set 2	Data Set 3	Data Set 4
Industry	Travel agency	Fashion retail	Fashion retail	Luggage retail
Number of different channels	8	8	7	7
Number of clicks	1,478,359	1,639,467	1,125,979	615,111
Number of journeys	600,978	1,184,583	862,112	405,339
Thereof with length ≥2	206,519	170,914	142,039	105,031
Thereof with length ≥5	48,344	30,095	12,416	11,475
Journey length	2.46 (8.860)	1.38 (1.916)	1.31 (1.238)	1.52 (4.587)
Number of conversions	9860	10,153	16,200	8115
Journey conversion rate	1.64%	0.86%	1.88%	2.00%

Note. Standard deviations are in parentheses.

All advertisers included in the evaluation distinguish seven or eight different online channels, although across firms, the channels that are used differ. Search engine advertising (SEA) and search engine optimization (SEO) appear in all four data sets. Other channels employed by the advertisers include direct type-in, affiliate marketing, display, price comparison, newsletter, referrer, retargeting, social media advertising, and others. Table 2 provides an overview of the online channels in our data, distinguishing between firm- and customer-initiated channels. As online marketing contacts can be initiated either by customers or by the firm, the origin of the contact is recognized as an important differentiator for online marketing channels (De Haan et al., 2016; Li & Kannan, 2014; Wiesel et al., 2011). In firm-initiated channels such as display advertising, the advertiser determines timing and exposures; in customer-initiated channels, customers actively trigger the communication—for instance, by performing a keyword search.

Table 3 provides information on the distribution of clicks across channels, illustrating the variation within and between data sets. The frequency of channels varies considerably across the four data sets. For example, although affiliate accounts for 46.1% of all clicks in Data Set 2, it is the least frequent channel in Data Set 4 (0.3% of clicks). This variation also alleviates endogeneity concerns. To rule out potential endogeneity, we furthermore conducted an analysis of pairwise correlations between logit model input variables (see Section 5 for variable specifications), which reveals low correlations between channels. The correlation matrices can be found in the Appendix.

4. Model development

We propose a graph-based Markovian framework to analyze customer journeys and derive an attribution model, adapting an approach proposed by Archak et al. (2010) in the context of search engine advertising. Markov chains are probabilistic models that can represent dependencies between sequences of observations of a random variable. They have a long history in marketing (Styan & Smith, 1964) and have frequently been used to model customer relationships (Homburg, Steiner, & Totzek, 2009; Pfeifer & Carraway, 2000). Other applications include advertising frequency decisions (Bronnenberg, 1998) and brand loyalty (Che & Seetharaman, 2009).

Table 2
Online channels.

Online channel	Description	Channel type
Type-in	Visits are classified as (direct) Type-in if users access the advertiser's website directly by entering the URL in their browser's address bar, or by locating a bookmark, favorite, or shortcut.	Customer-initiated
Search (SEA/SEO)	A consumer searching for a keyword in a general search engine (e.g., Google) receives two types of results: organic search results ranked by the search algorithm, and sponsored search results, also known as paid search or search engine advertising (SEA). While organic search or search engine optimization (SEO) results are available for free, SEA clicks are sold via second-price auctions.	Customer-initiated
Price comparison	Price comparison websites are vertical search engines that allow users to compare products by price and features. They aggregate product listings from a multitude of businesses, and direct users toward their websites.	Customer-initiated
Display	Display advertising, respectively banner advertising, entails embedding a graphical object with the advertising message into a website. Timing and exposures of display banners are determined by the firm.	Firm-initiated
Newsletter	Newsletter marketing, also known as e-mail marketing, encompasses sending marketing messages toward potential customers using e-mail.	Firm-initiated
Retargeting	Retargeting is a subclass of display advertising that is personalized toward the user based on his or her browsing history. It aims to re-engage users who have visited an advertiser's website, but did not complete a purchase.	Firm-initiated
Social media	Social media advertising comprises a set of advertising platforms belonging to the field of social media, such as social networks (e.g., Facebook), micromedia (e.g., Twitter), or other (mobile) sharing platforms (e.g., Instagram). In one of our data sets, the advertiser uses targeted Facebook display ads, which we define as social.	Firm-initiated
Affiliate	Affiliate marketing is a form of commission-based marketing in which a business (e.g., retailer) rewards the affiliate (e.g., a product review website) for referring a user toward the business's website. As affiliate in our data sets may include both coupon websites that are customer-initiated and ads provided by affiliate networks that may be more firm-initiated, a clear differentiation between customer- and firm-initiated contacts across data sets is not possible.	Customer-initiated/firm-initiated
Referrer	Referral or referrer traffic covers all traffic that is forwarded by external content websites (with or without remuneration)—for example, by including a text link. As traffic sources vary across data sets, a clear differentiation between customer- and firm-initiated contacts across data sets is not possible.	Customer-initiated/firm-initiated
Other	All forms of advertising that do not clearly fit into one of the categories above, are gathered in a separate category which we call “other”.	Customer-initiated/firm-initiated

Table 3
Channel distribution.

Description	Data Set 1			Data Set 2			Data Set 3			Data Set 4		
	Clicks per journey	Share of clicks	Rank	Clicks per journey	Share of clicks	Rank	Clicks per journey	Share of clicks	Rank	Clicks per journey	Share of clicks	Rank
Type-in	n/a[a]	n/a	n/a	.29(1.05)	20.7%	2	.25(.81)	18.9%	3	.13(4.42)	8.5%	3
SEA	.58(.90)	23.4%	2	.15(.72)	10.5%	4	.18(.55)	14.1%	4	1.13(1.12)	74.3%	1
SEO	.17(.61)	6.8%	3	.17(.795)	12.1%	3	.43(.28)	33.1%	1	.16(.54)	10.8%	2
Price comparison	.09(1.83)	3.7%	4	.00(.053)	0.1%	8	n/a	n/a	n/a	.04(.27)	2.5%	5
Display	1.46(8.75)	59.5%	1	.02(.30)	1.2%	7	n/a	n/a	n/a	n/a	n/a	n/a
Newsletter	.03(.24)	1.2%	7	.07(.72)	4.9%	5	.02(.16)	1.3%	6	n/a	n/a	n/a
Retargeting	.01(.16)	0.4%	8	n/a	n/a	n/a	.00(.01)	n/a	n/a	.02(.20)	1.1%	6
Social media	n/a	n/a	n/a	n/a	n/a	n/a	.28(.82)	21.5%	2	n/a	n/a	n/a
Affiliate	.06(.34)	2.5%	6	.64(1.02)	46.1%	1	n/a	n/a	n/a	.00(.16)	0.3%	7
Referrer	n/a	n/a	n/a	.06(.26)	4.3%	6	.14(.38)	11.0%	5	.04(.21)	2.5%	4
Other	.06(.36)	2.6%	5	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note. Standard deviations are in parentheses.
[a] This advertiser does not track direct type-ins in customer journey data, such that data on direct entries are not available.

In our model, we represent customer journeys as chains in directed Markov graphs.[3] A Markov graph $M = \langle S, W \rangle$ is defined by a set of states

$$S = \{s_1, \dots, s_n\} \tag{1}$$

and a transition matrix W with edge weights

$$w_{ij} = P\left(X_t = s_j | X_{t-1} = s_i\right), 0 \le w_{ij} \le 1, \sum_{j=1}^{N} w_{ij} = 1 \forall i. \tag{2}$$

Using this graph-based approach allows us to represent and analyze customer journeys in an efficient way, as the size of the final graph does not depend upon the number of journeys in the data set, but only on the number of states.

4.1. Base model

Customer journeys contain one or more contacts across a variety of channels. In the base model, each state s_i corresponds to one channel. If an advertiser employs three different channels C1, C2, and C3 in his online marketing mix, the model would include the three states C1, C2, and C3.[4] Additionally, all graphs contain three special states: a START state that represents the starting point of a customer journey, a CONVERSION state representing a successful conversion, and an absorbing NULL state for customer journeys that have not ended in a conversion during the observation period. The full set of states S in our example would then appear as follows: $S = \{\text{START, CONVERSION, NULL, C1, C2, C3}\}$.

The transition probability w_{ij} in the base model corresponds to the probability that a contact in channel i is followed by a contact in channel j. For the first channel in each journey, we add an incoming connection from the START state. If a customer journey ends in a conversion, we connect the state representing the last channel in the journey to the CONVERSION state; otherwise, it leads to the *NULL* state. For modeling reasons, we always add a connection from the CONVERSION state to the NULL state. Cycles in the graph are possible, such as when a sequence of two identical channels appears in a customer journey. Fig. 1 shows an exemplary Markov graph based on four customer journeys. Fig. 2 provides a graphical structure of the simple model for Data Set 1.

4.2. Higher-order models

Markovian models suggest that the present depends only on the first lag and do not incorporate previous observations. However, because prior research suggests that clickstreams should not be regarded as strictly Markovian (Chierichetti, Kumar, Raghavan, & Sarlós, 2012; Montgomery, Li, Srinivasan, & Liechty, 2004), we extend the approach proposed by Archak et al. (2010) by introducing alternative higher-order models in which the present depends on the last k observations. Transition probabilities can thus be defined as follows:

$$\begin{aligned} P(X_t = s_t | X_{t-1} = s_{t-1}, X_{t-2} = s_{t-2}, \dots, X_1 = s_1) \\ = P(X_t = s_t | X_{t-1} = s_{t-1}, X_{t-2} = s_{t-2}, \dots, X_{t-k} = s_{t-k}). \end{aligned} \tag{3}$$

[3] Called adgraphs by Archak et al. (2010).
[4] As we do not make any assumptions on the channels used, we employ dummy channels in our examples. In practice, the set of channels—and thus the set of states—depends on the actual channels used by the advertiser.

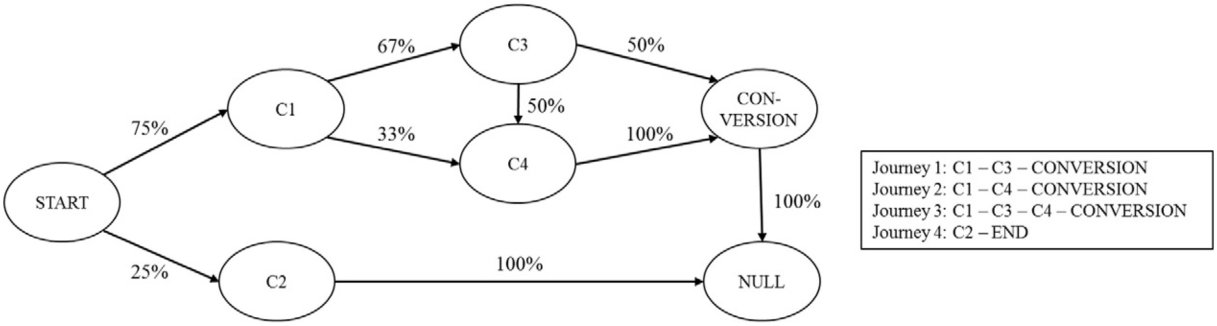

Fig. 1. Exemplary Markov graph.

For our implementation, we utilize the knowledge that a Markov chain of order k, over some alphabet A, is equivalent to a first-order Markov chain over the alphabet A_k of k-tuples. States in higher-order models, therefore, include k-tuples of states in the first-order models, such that we can employ the same algorithms. Unfortunately, the number of independent parameters increases exponentially with the order of the Markov chain and quickly becomes too large to be estimated efficiently with real-world data sets. Considering these implementation issues in relation to algorithmic efficiency, we limit our analyses to Markov chains with a maximum order of four.

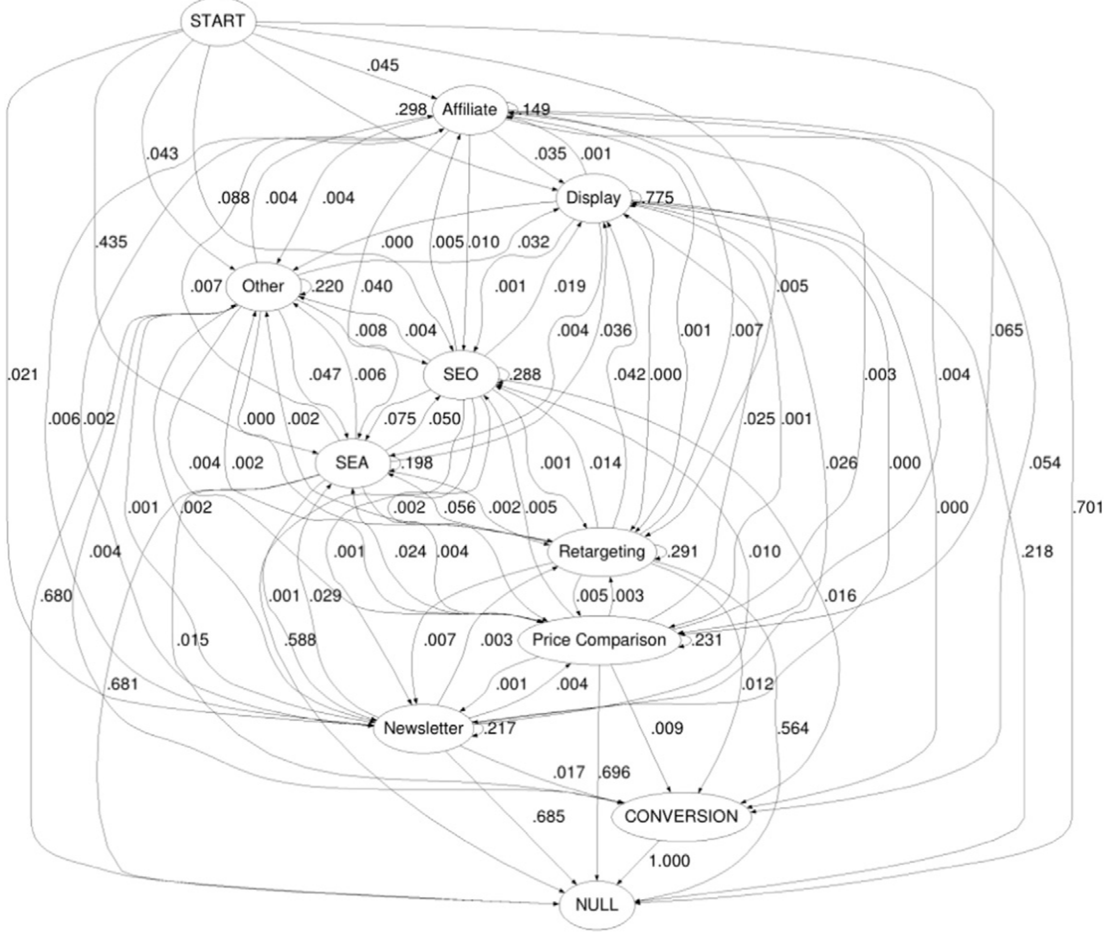

Fig. 2. Markov graph for Data Set 1 (base model).

Table 4
Removal effects for exemplary Markov graph in Fig. 1.

Channel	Visit(s_i)	Eventual Conversion(s_i)	Removal Effect(s_i)	Removal Effect(s_i) in %
C1	.75	1.00	.75	42.86%
C2	.25	.00	.00	.00%
C3	.50	1.00	.50	28.57%
C4	.50	1.00	.50	28.57%

4.3. Removal effect

The representation as Markov graphs allows identifying structural correlations in the customer journey data that can be used to develop an attribution model. Archak et al. (2010) propose a set of ad factors to capture the role of each state, such as Eventual Conversion(s_i)—that is, the probability of reaching conversion from a given state s_i. Visit(s_i) is the probability of passing s_i on a random walk beginning in the START state. For attribution modeling, we propose using the ad factor Removal Effect(s_i), defined as the change in probability for reaching the CONVERSION state from the START state when we remove s_i from the graph. As Removal Effect(s_i) reflects the change in conversion rate if the state s_i was not present, it enables us to simulate counterfactual analysis on historical data and is therefore well suited for measuring the contribution of each channel (or channel sequence). Using the assumption that all incoming edges of the state s_i that we remove are redirected to the absorbing NULL state, Removal Effect(s_i) is equivalent to the multiplication of Visit(s_i) and Eventual Conversion(s_i).[5] The removal effect can thus be efficiently calculated using matrix multiplication or by applying local algorithms provided by Archak et al. (2010). Table 4 illustrates the calculation of removal effects for the exemplary Markov graph presented in Fig. 1. In this graph that is based on four distinct journeys, C1 has a visit probability of .75 and an eventual conversion probability of 1. Multiplying these two values leads to a removal effect of .75 or 42.86% of the total removal effect in percent. Compared to C1, both C3 and C4—again with an eventual conversion probability of 1—have a lower removal effect of .5 (28.57%) due to their lower visit probability of .5. The removal effect of C2, which only appears in non-converting journeys, is zero.

Removal Effect(s_i) can take values between 0 and the total conversion rate. However, as most existing attribution heuristics use percentage values, we report removal effects per state as percentages of the sum of all removal effects (excluding the special states START, CONVERSION, and NULL), when comparing our results to other models. Higher-order models allow us to calculate removal effects for states representing channel sequences. To calculate the removal effect for channel i in higher-order models, we calculate the mean removal effect of all states containing channel i as the last observation or channel.

5. Model fit

In the following, we evaluate the proposed framework with regard to predictive accuracy and robustness and compare it to four benchmark models. Higher-order models outperform both the last click and the first click attribution heuristic models, as well as a simple logit model. In a comparison to a second logit model that includes order effects, they show similar results.

The specification of Logit Model 1 includes the number of clicks for each channel i:

$$\mathrm{logit}(Y_i) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n, \tag{4}$$

where x_i is the number of clicks in channel i in the journey. This model does not reflect the sequence of channels, but only the total number of clicks per channel.

Therefore, we add a second logit model which, in contrast, includes order effects:

$$\mathrm{logit}(Y_{it}) = \alpha + \sum_{i=1}^{n} \beta_{(4i-3)} d_{i1} + \beta_{(4i-2)} d_{i2} + \beta_{(4i-1)} d_{i3} + \beta_{4i} d_{i4}\ , \tag{5}$$

where d_{it} is a dummy variable for a click in channel i at position t, counting from the end of the journey. The dummy variable d_{it} is coded as 1 if channel i is present in position t; otherwise, it is set to 0. To make the number of variables in the logit model tractable, we include the last four contacts of each journey. This definition improves the comparability to fourth-order models, which take the last four contacts into account to determine the next transition.

5.1. Predictive accuracy

Although attribution primarily takes a retrospective view, attribution models should be able to correctly predict conversion events (Shao & Li, 2011). In addition to ensuring scientific rigor, this classification helps to persuade managers of the model's

[5] For a proof, see Archak et al. (2010).

credibility (Lodish, 2001) and lays the foundation for further applications beyond attribution, such as real-time bidding. We use the 10-fold cross-validation, which is superior to leave-one-out validation or bootstrapping since all the data serve as the holdout once (Sood, James, & Tellis, 2009). However, the discriminative power of standard metrics, such as percentage correctly classified or log-likelihood, is limited for measuring classification performance when class distribution is skewed (He & Garcia, 2009). We therefore turn to alternative measures to evaluate predictive accuracy. First, we choose the receiver operating characteristic (ROC) curve that decouples classification performance from class distributions and misclassification costs (Bradley, 1997). To compare our models, we reduce ROC performance to a single scalar value; the area under the ROC curve (AUC). Fig. 3 contains the ROC curves based on a within-sample evaluation of all journeys. As a second measure, we calculate the top-decile lift, which is defined as the proportion of the 10% of journeys predicted to be most likely to convert that actually end in a conversion relative to the baseline conversion rate (Neslin, Gupta, Kamakura, Lu, & Mason, 2006). In Table 5, we report both measures.

Although the overall predictive accuracy varies substantially between data sets, the relative predictive performance of the different model types is comparable. Within and out-of-sample performance for all models is similar, indicating a low risk of overfitting. With the exception of Data Set 2, the base model outperforms the first click approach and leads to results similar to the last click approach. Increasing the memory capacity substantially improves the predictive performance of our graph-based models, such that they clearly outperform the base model as well as the one-click approaches. The second logit model has the best predictive abilities among the four benchmark models, which are at a comparable level with third- and fourth-order models.

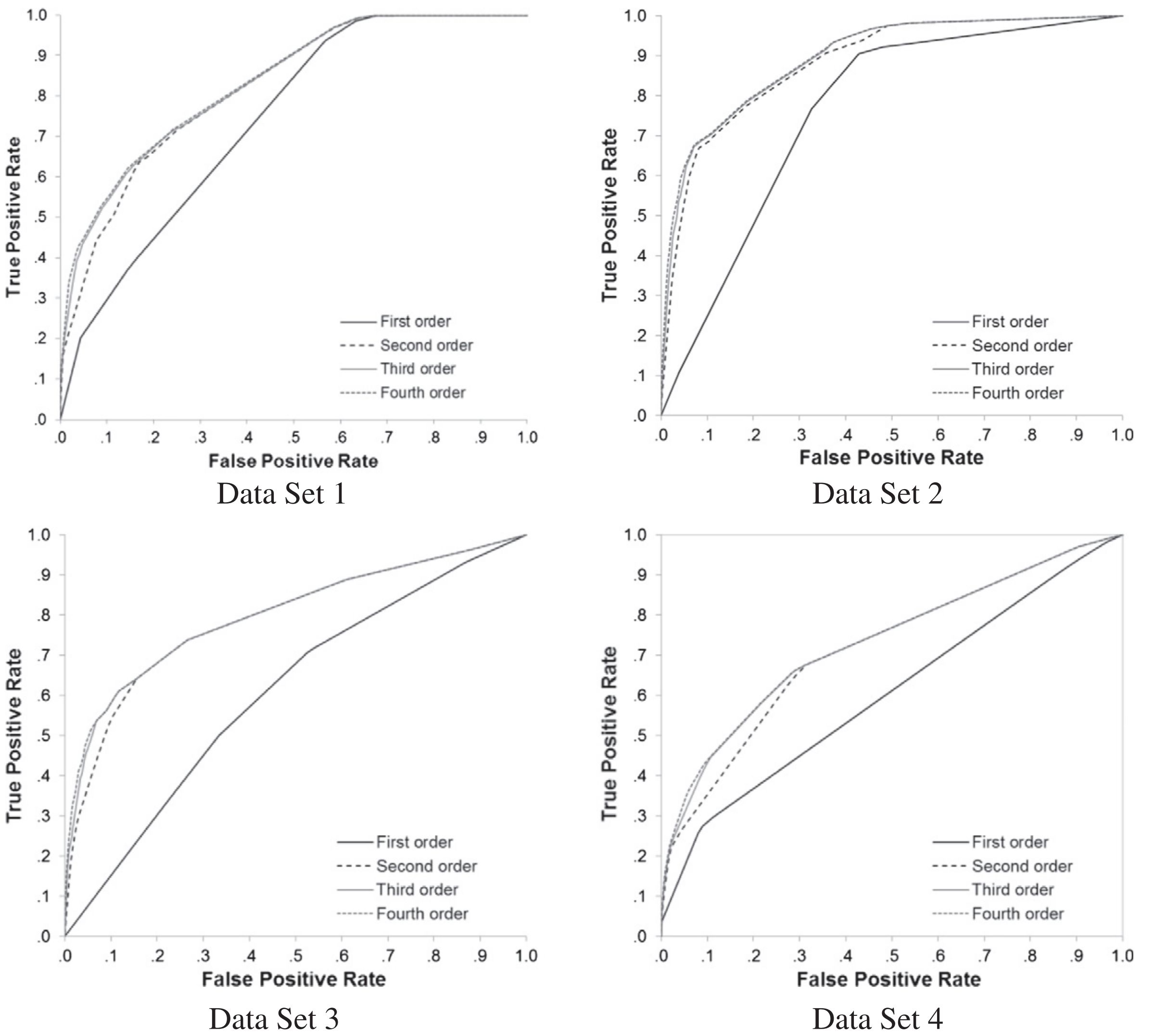

Fig. 3. ROC curves (within sample).

Table 5
Predictive accuracy.

Measure	Sample	Data set	Model							
			Base model	Second order	Third order	Fourth order	Last click	First click	Logit Model 1	Logit Model 2
AUC	Within sample	DS 1	.7398 (.0006)	.8214 (.0007)	.8323 (.0008)	.8357 (.0010)	.7398 (.0006)	.7133 (.0005)	.8185 (.0022)	.8278 (.0010)
		DS 2	.7604 (.0006)	.8834 (.0006)	.8932 (.0005)	.8967 (.0005)	.7609 (.0007)	.7726 (.0008)	.8654 (.0011)	.8888 (.0004)
		DS 3	.6079 (.0009)	.7933 (.0015)	.7996 (.0014)	.8040 (.0018)	.6079 (.0009)	.5980 (.0006)	.8037 (.0011)	.8041 (.0010)
		DS 4	.6029 (.0010)	.7191 (.0029)	.7309 (.0009)	.7333 (.0009)	.6029 (.0010)	.5528 (.0007)	.7281 (.0013)	.7308 (.0011)
	Out-of-sample	DS 1	.7396 (.0061)	.8207 (.0022)	.8305 (.0062)	.8294 (.0063)	.7406 (.0056)	.7136 (.0050)	.8185 (.0074)	.8274 (.0071)
		DS 2	.7607 (.0056)	.8832 (.0054)	.8924 (.0044)	.8933 (.0050)	.7619 (.0056)	.7728 (.0071)	.8653 (.0086)	.8887 (.0037)
		DS 3	.6079 (.0079)	.7930 (.0086)	.7988 (.0088)	.8006 (.0089)	.6093 (.0082)	.5990 (.0064)	.8036 (.0088)	.8040 (.0089)
		DS 4	.6030 (.0093)	.7180 (.0083)	.7280 (.0080)	.7230 (.0072)	.6050 (.0080)	.5550 (.0063)	.7276 (.0084)	.7299 (.0084)
Top-decile lift	Within sample	DS 1	2.9056 (.0211)	4.7181 (.0249)	5.2798 (.0198)	5.3587 (.0203)	2.9056 (.0211)	2.4333 (.0203)	4.9226 (.0261)	5.1854 (.0500)
		DS 2	2.4880 (.0407)	6.8070 (.0156)	6.9306 (.0146)	6.9546 (.0158)	2.4955 (.0312)	2.8692 (.0311)	6.6011 (.0211)	6.8893 (.0146)
		DS 3	1.5065 (.0254)	5.1744 (.0207)	5.6426 (.0221)	5.6450 (.0232)	1.5065 (.0254)	1.4967 (.0188)	5.7731 (.0197)	5.7358 (.0186)
		DS 4	2.8089 (.0137)	3.4708 (.1148)	4.1838 (.0134)	4.2897 (.0161)	2.8089 (.0137)	1.6820 (.0144)	4.0112 (.1042)	4.2194 (.0171)
	Out-of-sample	DS 1	2.9075 (.1385)	4.6937 (.1593)	5.2393 (.1563)	5.2913 (.1704)	2.9175 (.1508)	2.4296 (.0934)	4.9260 (.1487)	5.1856 (.1499)
		DS 2	2.4748 (.0742)	6.7889 (.1442)	6.9277 (.1178)	6.9113 (.1202)	2.5103 (.1078)	2.8709 (.1532)	6.6109 (.1684)	6.8978 (.1023)
		DS 3	1.4896 (.0910)	5.1619 (.1433)	5.6311 (.1993)	5.6107 (.2055)	1.5418 (.1002)	1.5143 (.1368)	5.7714 (.1704)	5.7278 (.1661)
		DS 4	2.8112 (.1074)	3.4556 (.1737)	4.1198 (.1273)	4.1553 (.1464)	2.8256 (.1114)	1.6897 (.1337)	4.0124 (.1332)	4.2090 (.1522)

Note. Standard deviations are in parentheses.

Our framework thus achieves results comparable to widely accepted and tested models, making it well suited for analyzing the structural properties of customer journeys.

5.2. Robustness

Robustness is another important metric to evaluate model fitness (Little, 2004; Shao & Li, 2011), as it conveys the ability of a model to deliver stable and reproducible results if the model is run multiple times and is therefore indispensable for sustainable attribution results. For our models, robustness applies to two measures. First, predictive accuracy should be robust across all cross-validation repetitions. Table 5 lists the standard deviations of the predictive performance measures for each model, as well as for the four models we use as a comparison. The results imply low overall variation without systematic differences between models. Second—and even more important—the variable used for attribution modeling should provide stable attribution results that offer a reliable basis for managerial decisions, such as budget shifts. Therefore, we specifically test the robustness of the Removal Effect(s_i) ad factor. For each model state s_i, we compute the average standard deviation of Removal Effect(s_i) across ten cross-validation repetitions. We report the stability of the removal effects as percentages of the average removal effect across all states, as the number of states per model and, correspondingly, the mean Removal Effect(s_i) varies. We summarize these validation results in Table 6.

Table 6
Removal effect: average standard deviation as % of average removal effect (10-fold cross-validation).

Data set	Model			
	Base model	Second order	Third order	Fourth order
Data Set 1	1.14%	1.92%	3.25%	5.43%
Data Set 2	1.51%	2.10%	3.81%	7.57%
Data Set 3	1.31%	1.72%	2.78%	5.15%
Data Set 4	1.34%	1.80%	2.97%	4.67%

Table 7
Attribution results in comparison to heuristic models (in %).

	Data Set 1			Data Set 2			Data Set 3			Data Set 4		
	Markov graph (third order)	Last click	First click	Markov graph (third order)	Last click	First click	Markov graph (third order)	Last click	First click	Markov graph (third order)	Last click	First click
	%	%	%	%	%	%	%	%	%	%	%	%
Type-in	n/a	n/a	n/a	35.23%	43.91%	40.28%	27.55%	29.77%	25.51%	18.89%	22.02%	13.71%
SEA	46.29%	53.19%	56.36%	19.81%	22.27%	23.60%	18.53%	20.16%	20.70%	59.91%	60.98%	76.26%
SEO	19.54%	16.76%	16.67%	15.79%	13.66%	13.24%	22.03%	21.33%	21.12%	9.24%	7.79%	5.30%
Price comparison	5.29%	4.78%	6.05%	0.21%	0.11%	0.12%	n/a	n/a	n/a	3.56%	2.17%	2.21%
Display	0.93%	0.14%	0.21%	3.70%	1.79%	1.37%	n/a	n/a	n/a	n/a	n/a	n/a
Newsletter	4.47%	2.93%	4.28%	14.54%	8.76%	11.94%	1.21%	1.15%	1.32%	n/a	n/a	n/a
Retargeting	1.25%	0.67%	0.78%	n/a	n/a	n/a	0.07%	0.01%	0.00%	2.09%	1.72%	0.16%
Social media	n/a	n/a	n/a	n/a	n/a	n/a	24.87%	20.73%	23.83%	n/a	n/a	n/a
Affiliate	19.71%	20.17%	13.66%	8.72%	7.83%	6.87%	n/a	n/a	n/a	4.41%	3.67%	0.42%
Referrer	n/a	n/a	n/a	1.99%	1.67%	2.58%	5.72%	6.85%	7.52%	1.88%	1.65%	1.96%
Other	2.52%	1.36%	1.97%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
χ^2		12,911	18,019		40,931	30,076		5948	4822		3357	36,979
df		7	7		7	7		6	6		6	6
P		<.001	<.001		<.001	<.001		<.001	<.001		<.001	<.001

For all data sets in our sample, the average standard deviation as a percentage of the average removal effect increases with model order, leading to a necessary trade-off between predictive accuracy and robustness.

Combining our evaluation results with practical considerations, we recommend third-order models for standard attribution analyses, as they balance predictive performance, robustness, and algorithmic efficiency. Removal effects can be calculated efficiently in $O(|S|^2)$ time (Cormen, Leiserson, Rivest, & Stein, 2009) and hence allow for frequent model updates. However, as the number of states increases exponentially with the order of the Markov chain, we limit our analyses to fourth-order models to allow for updates in near real-time.[6] Using higher-order models also yields additional insights into channel interactions, thereby further increasing managers' understanding of the interplay across channels.

6. Results

6.1. Attribution results

Next, we take a closer look at the attribution results of our proposed framework and compare it with the attribution heuristics most widely used in industry practice, namely, last click and first click attribution(The CMO Club, & Visual IQ, Inc, 2014), and the two logit models, which show comparable predictive abilities. As prior research indicates that heuristic attribution approaches can lead to incorrect conclusions (Abhishek et al., 2015; Li & Kannan, 2014; Xu et al., 2014), a comparison of our results across four data sets enables us to investigate structural differences and move toward empirical generalizations. Our analyses are based on the full data sets and show the contribution of each channel toward final conversions. Given our evaluation results, we use third-order Markov models in our comparison and calculate the mean removal effect of all states containing channel i as the last observation or channel. To enable comparisons to the first click and last click heuristic, we present our model results as percentage values in Table 7.

First, we analyze, across data sets, the attribution results provided by our framework. In sum, the customer-initiated channels (direct type-in, SEA, SEO, and price comparison) accumulate the majority of the contribution (between 68% for Data Set 3 and 92% for Data Set 4), highlighting their general importance for companies' online marketing success (Shankar & Malthouse, 2007). However, the actual contributions and corresponding rankings by channel vary substantially across data sets. SEA, the strongest channel for Data Set 4 (60%) and Data Set 1 (46%), receives less than 20% of the credit for Data Sets 2 and 3. For firm-initiated channels, the results also show considerable variation. For example, a newsletter shows values between 1% (Data Set 3) and 15% (Data Set 2). These differences may reflect the advertisers' individual channel strategies, such that marketers need to individually derive and verify detailed implications.

However, we observe significant differences when comparing the results of the Markov model to those of the heuristic approaches, some of which apply consistently across data sets and allow for generalizations of prior research. In general, the Markov model levels the channel contributions' amplitudes, distributing the contribution more evenly across channels. The heuristic attribution models consistently assign less credit to the firm-initiated channels display, retargeting, social media, and

[6] However, depending on the use case, an advertiser may choose a different trade-off between predictive accuracy, robustness, and algorithmic efficiency.

Table 8
Estimation results for Logit Model 1.

	Data Set 1			Data Set 2			Data Set 3			Data Set 4		
	B	Exp(B)	ME	B	Exp(B)	ME	B	Exp(B)	ME	B	Exp(B)	ME
Type-in				.259***	1.295	.0016***	.707***	2.028	.0098***	.376***	1.457	.0066***
SEA	.443***	1.557	.0000***	.417***	1.517	.0026***	.724***	2.063	.0101***	.243***	1.275	.0043***
SEO	.388***	1.474	.0000***	.196***	1.216	.0012***	.440***	1.552	.0061***	.142***	1.153	.0025***
Price comparison	.002*	1.002	.0000***	.548***	1.730	.0035***				.255***	1.290	.0045***
Display	−3.427***	.032	−.0004***	.188***	1.207	.0012***						
Newsletter	.616***	1.852	.0001***	.277***	1.319	.0017***	.719***	2.052	.0100***			
Retargeting	.310***	1.363	.0000***				.939	2.558	.0131	.195***	1.215	.0034***
Social media							.504***	1.655	.0070***			
Affiliate	1.152***	3.165	.0001***	−.195***	.823	−.0012***				1.848***	6.346	.0327***
Referrer				.108***	1.114	.0007***	.559***	1.748	.0078***	.136**	1.146	.0024**
Other	.032	1.033	.0000									
Intercept	−4.505***	.011		−5.124***	.006		−4.978***	.007		−4.372***	0.013	
Observations	600,978			1,184,583			862,112			405,339		

* $p < .05$.
** $p < .01$.
*** $p < .001$.

newsletter.[7] Display as well as retargeting and social media advertising, which—for the advertisers in our sample—can be regarded as specialized forms of banner advertising, may be underrepresented by one-click approaches, as banners increase brand awareness (Ilfeld & Winer, 2002; Sherman & Deighton, 2001) and thereby reduce the user's cost of subsequent website visits through search channels (Li & Kannan, 2014). Furthermore, retargeting is, by definition, inappropriately reflected by the first click attribution approach, as it highlights specific products or services the user was browsing previously (Lambrecht & Tucker, 2013). For a newsletter, extant research reveals positive but weak short-term effects yet strong long-term effects (Breuer, Brettel, & Engelen, 2011). Newsletters, stored in the inbox, can be repeatedly accessed and may motivate subsequent visits through other channels. Single-click heuristics may not be able to reflect these effects appropriately.

Similar to the Markov model results, customer-initiated channels accumulate the majority of the contribution in the heuristic approaches. However, they leave a more ambiguous picture on a channel level: compared to heuristic approaches, the Markov model consistently assigns less credit to SEA, but more credit to SEO. Direct type-ins, when users directly access the company website, get more credit by the last click approach, whereas price comparison receives less credit. Our findings for SEA and direct type-in are consistent with prior research on attribution: Xu et al. (2014) show that the last click heuristic is biased toward SEA. Li and Kannan (2014) find that the informational stock of other channel visits catalyzes ultimate purchases via direct type-ins, such that type-in is overestimated by the last click approach. However, in contrast to previous findings (Li & Kannan, 2014), our approach assigns more credit to SEO than the heuristic approaches. These results could be explained by differences between product and retailer brands. The hotel chain in the study by Li and Kannan (2014) has a strong "product" brand. By contrast, the advertisers in our sample all act as retailers. For organic search, keywords containing product brands positively predict direct conversions, yet there is no significant effect for retailer brands (Yang & Ghose, 2010). As a consequence, SEO would benefit less from spillovers from other channels. Similar to SEO, price comparison gets more credit than in the last click approach. Breuer et al. (2011) find that price comparison websites not only affect sales in the short term, but compared to other online channels also have a moderate long-term effect, which may not be fully captured by one-click heuristics.

The results for affiliate and referral are inconclusive. This corresponds to the fact that a clear differentiation between customer- and firm-initiated contacts across data sets is not possible. For Data Set 1, the affiliate channel includes coupon websites, which may be frequented by more experienced customers who check for an available discount coupon before finalizing their purchase. Although the last click attribution approach would allow those affiliate websites to "free-ride" on previous contacts (Berman, 2015), our framework assigns less credit to the affiliate channel. For Data Sets 2 and 4, where the affiliate channel includes more firm-initiated contacts, the Markov model assigns more credit to affiliate than the heuristic approaches, which is in line with our finding that heuristic approaches underestimate firm-initiated channels.

In our data sets, referrals mainly relate to aggregator websites that resemble search engines in their functionality and, thus, feature customer-initiated elements. Compared to the first-click model, our framework constantly assigns less credit to referrer—similar to our results for SEA. With regard to the last-click model, the results are mixed: While our model assigns more credit to referrer for Data Sets 2 and 4, it assigns less credit for Data Set 3. Different levels of brand awareness about the advertisers may explain these findings and should be investigated in more detail.

The results of Logit Model 1, which we present in Table 8, reveal a larger variation between data sets. The strongest predictors of conversion events are affiliate for Data Sets 1 and 4, price comparison for Data Set 2, and retargeting for Data Set 3. By contrast,

[7] With the exception of Data Set 3, where the credit assigned to newsletter is higher in the first click model.

Table 9
Attribution results for second-order model (Data Set 1).

	Preceding channel								
Current channel	START	SEA	SEO	Price comparison	Display	Newsletter	Retargeting	Affiliate	Other
SEA	32.52%	14.27%	2.43%	0.34%	0.35%	0.48%	0.15%	0.84%	0.47%
SEO	7.78%	4.40%	4.75%	0.11%	0.07%	0.10%	0.06%	0.42%	0.08%
Price comparison	3.07%	0.22%	0.05%	1.36%	0.03%	0.04%	0.01%	0.04%	0.04%
Display	0.62%	0.09%	0.02%	0.02%	0.42%	0.02%	0.01%	0.03%	0.01%
Newsletter	1.91%	0.27%	0.13%	0.02%	0.04%	0.99%	0.02%	0.07%	0.01%
Retargeting	0.34%	0.20%	0.06%	0.03%	0.02%	0.02%	0.24%	0.04%	0.01%
Affiliate	8.10%	2.61%	1.00%	0.23%	0.15%	0.20%	0.08%	5.31%	0.15%
Other	1.12%	0.27%	0.06%	0.01%	0.02%	0.02%	0.01%	0.03%	0.50%

our graph-based model indicates that all mentioned channels play only a subordinate role. These differences become especially evident when analyzing Data Set 2. According to the Markov model, as well as to the first and last click attribution approach, less than 1% of credit is attributed to price comparison, which is the channel with the lowest number of contacts and is present in only 47 of the 10,153 converting journeys. By contrast, price comparison has the highest marginal effect (ME $= 0.0035$, $p < .001$) in the logit model. These discrepancies derive from different underlying model mechanisms. Logit models do not aim to reflect the overall contribution of a variable (i.e., channel) but rather focus on the predictive ability of a variable with regard to a binary dependent variable (e.g., a conversion event). Thus, the logit models are well suited for calibrating the predictive capabilities of other frameworks—for example, our graph-based approach.

With regard to predictive accuracy, the second logit model performs better than the first logit model. However, due to multicollinearity, the coefficient estimates for the last click are not easy to interpret in an attribution context because coefficients are indeterminate and the standard errors of the estimates are large. As each journey includes at least one click, one of the channel dummy variables for $t = 1$ can be completely explained by a linear combination of the other dummy variables, leading to perfect multicollinearity for $t = 1$.[8] For the other positions, the results show different patterns in different channels across data sets that do not allow for generalizations.

In summary, our findings generalize prior research, showing that heuristics attribution approaches tend to underestimate the contribution of selected firm-initiated channels (Li & Kannan, 2014; Xu et al., 2014). For customer-initiated channels, the contribution of SEA, and especially direct type-ins, is overestimated by the last click approach. For other customer-initiated channels, additional factors such as industry and brand characteristics seem to play a role, such that advertisers need to individually derive and verify detailed implications, ideally on a more granular level. For example, it would be worthwhile to separately analyze the contributions for keywords that contain product brands, retailer brands, or no brand name at all.

6.2. Interplay of channels

In addition to channel-level attribution, higher-order models offer a more detailed view of the interplay of channels, which we illustrate using the second-order model for Data Set 1 in Table 9. The second-order model attribution results for the other data sets can be found in the Appendix.

Across all data sets, the increase in overall purchase probability for most channels is highest right after the START state, near the beginning of the journey[9]—which corresponds to the high share of one-click journeys in our data.

Sequences of identical channels show high removal effects in all data sets. For example, in Data Set 1, affiliate preceded by affiliate has a percentage removal effect of 5.31%, whereas the average removal effect for affiliate preceded by channels other than affiliate is only 0.63%. For sequences of direct type-ins, which show comparably high removal effects, the results may mirror Bowman and Narayandas' (2001) finding on customers' selective firm preferences expressed by direct visits. However, the consistent importance of same-channel sequences across data sets may further indicate idiosyncratic channel preferences for some users, comparable to those found in multichannel relational communication (Godfrey, Seiders, & Voss, 2011). These channel preferences are also in line with related research findings showing that previous visits through specific channels may reduce the cost for current visits through the same channels (Li & Kannan, 2014). Future research should investigate the existence of such preferences, their antecedents and their implications for multichannel online advertising.

Besides idiosyncratic channel preferences, we also find spillover effects between channels both within and between channel groups. First, customer-initiated channels show substantially higher removal effects if they are followed by another customer-initiated channel than if they are followed by firm-initiated channels such as display and newsletter. With increasing experience

8 As an alternative, it would have been possible to select a specific channel as a reference channel. We decided against this solution, as this would have been counterintuitive for $t = 2$ to $t = 4$ where empty positions coded as zero are possible in journeys with fewer than four clicks.

9 A notable exception is retargeting, which explicitly targets customers who have previously visited the advertiser's website (Lambrecht & Tucker, 2013).

in online shopping, users anticipate the benefit and the occurred effort associated with each online channel. In consequence, whenever a purchase need arises, more directed users may sequentially use customer-initiated channels they are familiar with to reduce search cost on their path to purchase (Li & Kannan, 2014). Regarding the interplay of search channels, prior research has argued that click propensities for paid and organic search results are moderated by involvement (Jerath, Ma, & Park, 2014). Our findings suggest that personal channel preferences also play a role, as spillover effects from SEA to SEO are much stronger than from SEO to SEA. Although consumers who have clicked on sponsored search ads react to both paid and organic results, consumers clicking on organic search results seem to generally prefer organic results. These results are consistent with Li and Kannan (2014), but contrary to Yang and Ghose (2010), who find that organic search has stronger spillovers toward paid search. Future research should investigate the role of channel preferences and their interaction with involvement.

The interplay between search and direct type-in shows directional effects across data sets, such that spillovers from both search channels, SEA and SEO, toward direct visits clearly outreach the reversed effects. Users who visit the firm's website via search are more likely to return via direct type-in, potentially because preceding search contacts have shaped their preferences toward a dedicated firm's offering (Bowman & Narayandas, 2001). For price comparison, the results vary considerably across data sets. While for Data Set 1 (travel), price comparison shows stronger spillovers toward both search channels, for Data Set 4 (luggage retail), the corresponding spillovers are in the opposite direction. These variations between data sets indicate industry-specific effects and may also be moderated by the customer's price sensitivity with regard to the respective product category (Mehta, Rajiv, & Srinivasan, 2003).

Furthermore, we also find indications for spillover effects of customer-initiated channels toward firm-initiated channels—and vice versa. These findings again generalize prior research on multichannel customer journeys (Li & Kannan, 2014; Xu et al., 2014). The removal effect for display, retargeting, social media, and newsletter is higher if these channels are preceded by SEO, SEA, or direct type-in instead of other firm-initiated channels. Users originating from customer-initiated channels may be more prone to respond to pushed, firm-initiated media than are users originating from firm-initiated channels. The results further indicate moderate removal effects for firm-initiated channels preceding customer-initiated channels. For instance, prior display visits show spillover effects toward direct type-in and SEA, which is in line with previous research (Ilfeld & Winer, 2002; Li & Kannan, 2014; Sherman & Deighton, 2001). The results for newsletter show similar patterns. Users who respond to newsletter-embedded links are more likely to return via SEA, direct type-in or, above all, a newsletter which, except for direct type-in, corresponds to prior research (Li & Kannan, 2014). According to our results, the spillover effects between firm-initiated channels are, by and large, negligible. Except for some moderate spillovers between display and newsletter, none of the removal effects between firm-initiated channels are noteworthy. These findings correspond to extant research indicating that subsequent firm-initiated contacts reflect stagnation in customer decision making and thus do not positively predict conversion probabilities (Anderl, Schumann, & Kunz, 2015).

On a channel level, it is interesting to analyze which channels induce the strongest spillover effects—and which channels benefit most from spillovers from other channels. To this end, we sum up the removal effects of all two-channel combinations that have channel C1 in the first position (corresponding to a column in Table 9) and compare them to the sum of removal effects of all combinations ending with C1 (corresponding to a row in Table 9). Across all data sets, the summarized removal effects of sequences starting with SEA, price comparison, newsletter, and referrer consistently surpass the removal effects of sequences ending in these channels. The opposite holds true for direct type-in and retargeting. A newsletter exhibits strong long-term effects, yet only weak short-term effects (Breuer et al., 2011). Consequently, newsletter contacts may serve as facilitator for subsequent website visits via other channels, namely, direct type-in, SEA, SEO, and display, rather than as a direct sales channel enabled by other channels (Li & Kannan, 2014). By definition, retargeting focuses on customers who have previously visited the advertiser's website, and thus requires anteceding channels. For SEO, display, and affiliate, the picture is mixed. As the diverging results for customer-initiated channels may not be sufficiently explained by the channels' contact origin alone, a more differentiated view taking brand usage into account becomes appropriate (Anderl et al., 2015). Whereas in some customer-initiated channels, such as direct type-ins, customers actively use the advertiser's brand to initiate the contact, other channels (including price comparison), instead entail generic uses. Compared to generic customer-initiated channels (e.g., price comparison), branded customer-initiated channels (e.g., direct type-in) seem to benefit more from spillovers from other channels. Depending on the keyword used, a search conveys both generic and branded uses; however, the proportion of generic search queries in general outweighs branded search queries regarding both SEA and SEO (Jansen & Spink, 2009), such that—in aggregation—search channels appear closer to generic uses, which supports the above generalization. A more detailed investigation of branded versus generic contacts, as well as the role of brand awareness, would be a valuable extension of our research. Besides increasing predictive performance, the application of higher-order models thus enables a more detailed understanding of the interplay across channels, and allows identifying several promising avenues for future research.

7. Discussion and outlook

In this study, we introduce an attribution framework based on first- and higher-order Markov walks to examine the contribution of individual online channels and to shed light onto carryover and spillover effects between online channels, contributing to marketing theory and practice in the following ways.

First, building on multiple data sets and comparing our estimation results to two prominent attribution heuristics and to two logit models enables us to derive empirical generalizations on individual channel contributions in a multichannel setting. Complementing extant attribution literature (Abhishek et al., 2015; Li & Kannan, 2014; Xu et al., 2014), we find that firm-initiated channels are consistently undervalued by the heuristic attribution approaches across industries. By contrast, the contribution of direct type-ins and SEA is overestimated by these heuristic approaches across industries. For other customer-initiated channels, especially SEO, additional factors such as industry and brand characteristics seem to play a role.

Second, higher-order models allow new insights into the interplay of channels along the customer journey. Using our framework, we generalize prior findings indicating that a majority of channels exhibit idiosyncratic channel carryovers (Li & Kannan, 2014). The consistent relevance of homogeneous channel sequences across data sets may mirror extant channel preferences for some users. Advertisers anticipating these preferences may improve targeting measures for these customer segments. For search channels, paid search touches are often followed by both paid and unpaid search contacts, while their unpaid equivalents are merely followed by unpaid search contacts. These results endorse extant research (Li & Kannan, 2014); however, they contradict industry-specific findings by Yang and Ghose (2010). Moreover, we observe spillover effects within, as well as between, channel categories. Customer-initiated channels exhibit substantial removal effects if they are followed by other customer-initiated channels, whereas—in line with prior research (Anderl et al., 2015)—spillover effects between firm-initiated channels are by and large negligible. Spillovers between customer-initiated and firm-initiated channels are more selective and are carried out at a moderate level in both directions. Our findings are also useful at the channel level, particularly regarding the ability of each channel to induce spillover effects—and to benefit from them. Across all data sets, we observe that the removal effects of sequences that begin with SEA, price comparison, newsletter, and referrer consistently outreach the removal effects of sequences ending in these channels. For direct type-in and retargeting, the opposite holds true. We conjecture that a more differentiated view that includes branded channel usage for customer-initiated channels may help to explain the differences (Anderl et al., 2015).

Third, we develop a novel attribution variant that adds to existing attribution modeling techniques (Abhishek et al., 2015; Berman, 2015; De Haan et al., 2016; Kireyev et al., 2016; Li & Kannan, 2014; Xu et al., 2014) by representing multichannel online customer path data as Markov walks in directed graphs. Higher-order models, in particular, allow investigation of channel sequences and spillovers between channels. Our findings support previous research (Abhishek et al., 2015; Berman, 2015; Li & Kannan, 2014), indicating that one-click attribution measures are neither suited to derive the factual contribution of individual channels nor to determine carryover and spillover effects across channels.

Finally, our framework provides a useful approach to a number of explicit problems that online marketers confront. Because our model captures the contribution of each online channel, it may help to calibrate online channel budgets and facilitate moves toward an improved budget allocation (Raman et al., 2012). If the budget share of a channel is higher than its actual contribution, advertisers should reconsider their budget splits. However, advertisers should proceed carefully, as endogeneity issues and varying channel elasticities can make budget allocation complex. Moreover, using our graph-based framework, advertisers can more accurately calculate the conversion probability Eventual Conversion(s_i) of a customer, given his or her previous customer journey. This information can support state-of-the-art applications such as real-time bidding decisions. In addition, by more fully anticipating the users' channel preferences and the potential cost incurred by each channel, advertisers may be able to reach customers through more efficient channels.

Our research has several limitations that may serve to stimulate research on attribution and online marketing effectiveness. We used four data sets from different industries, but some findings may still be company-specific. In addition, the customer journeys in these data sets were short on average, including a high number of one-click journeys. Yet sophisticated attribution is not required for journeys consisting of just a single click, as in that case one-click heuristics would suffice. Furthermore, our data sets do not include views and exposures to offline channels, although our framework should be well suited to handle such information. Companies may also intend to consider revenues and profits from conversions, and the customer lifetime value of customers acquired. The timing between contacts along the customer journey would be another interesting extension of the research.

The attribution problem is, by definition, endogenic; it measures the relative effectiveness of channels in a given setting (Li & Kannan, 2014), so the results are conditional upon a number of management decisions such as channels used or budget limits per channel. Therefore, optimizing the budget allocation remains an iterative process. Correct attribution, however, is a necessary prerequisite for managers to optimize their budget decisions.

Finally, a strict causal interpretation of customer journeys is difficult, because alternative explanations may exist for correlations between conversions and advertising exposure. Some channels, such as retargeting, explicitly target customers with a higher propensity to purchase (Lambrecht & Tucker, 2013) and observed correlations might be due to selection effects (Lewis, Rao, & Reiley, 2011). Establishing a causal relationship would require large-scale field experiments with randomized exposure. Such experiments are challenging to implement in practice, especially in multichannel settings, but comparing our attribution framework against results of the experiments would be a valuable direction for building on our research.

Acknowledgments

The authors thank intelliAd Media GmbH, especially Franz Graf, Darius Suryadi, Tobias Kiessling, and Mischa Rürup, for providing them with data sets, valuable discussions, and the practical implementation of the framework. They gratefully acknowledge support from the German Federal Ministry of Education and Research (Project “Fre(E)S – Productivity of Free E-Services”, FKZ 01FL10038) and the Wharton Customer Analytics Initiative.

Appendix A

Table 10
Correlation matrix (Data Set 1).

	SEA	SEO	Price Comparison	Display	Newsletter	Retargeting	Affiliate	Other
SEA	1.00							
SEO	.00***	1.00						
Price comparison	−.02***	−.01***	1.00					
Display	−.08***	−.04***	−.01***	1.00				
Newsletter	−.05***	−.01***	.00***	−.02***	1.00			
Retargeting	.01***	.01***	.00	−.01***	.01***	1.00		
Affiliate	−.05***	−.01***	.00**	−.02***	.00**	.00	1.00	
Other	−.06***	−.03***	−.01***	−.02***	−.02***	−.01***	−.02***	1.00

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 11
Correlation matrix (Data Set 2).

	Type-in	SEA	SEO	Price comparison	Display	Newsletter	Affiliate	Referrer
Type-in	1.00							
SEA	.01***	1.00						
SEO	.00***	.10***	1.00					
Price comparison	.00	.00***	.01***	1.00				
Display	.05***	.05***	.08***	.01***	1.00			
Newsletter	.07***	.05***	.04***	.00	.15***	1.00		
Affiliate	−.14***	−.09***	−.10***	−.01***	−.01***	−.05***	1.00	
Referrer	−.04***	−.01***	.00	.00**	.01***	−.01***	−.12***	1.00

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 12
Correlation matrix (Data Set 3).

	Type-in	SEA	SEO	Newsletter	Retargeting	Social media	Referrer
Type-in	1.00						
SEA	−.06***	1.00					
SEO	−.11***	−.07***	1.00				
Newsletter	−.03***	−.03***	−.05***	1.00			
Retargeting	.00	.00	.00	.00	1.00		
Social media	−.06***	−.08***	−.15***	−.03***	.00	1.00	
Referrer	−.09***	−.10***	−.17***	−.04***	.00	−.11***	1.00

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 13
Correlation matrix (Data Set 4).

	Type-in	SEA	SEO	Price comparison	Retargeting	Affiliate	Referrer
Type-in	1.00						
SEA	−.01***	1.00					
SEO	.00*	−.07***	1.00				
Price comparison	.00*	−.05***	.02***	1.00			
Retargeting	.00	.06***	.00	.01***	1.00		
Affiliate	.06***	.02***	.02***	.01***	.00	1.00	
Referrer	.00	−.14***	−.04***	.00**	.00	.01***	1.00

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 14
Estimation results for Logit Model 2.

		Data Set 1			Data Set 2			Data Set 3			Data Set 4		
		B	Exp(B)	ME	B	Exp(B)	ME	B	Exp(B)	ME	B	Exp(B)	ME
	Type-in				3.27	26.32	0.01	5.64	282.46	0.72	9.30	10,895.44	0.98**
	SEA	10.27	28.827.82	0.97	4.38	80.19	0.16	5.56	259.87	0.56	6.70	808.79	0.90
	SEO	4.54	93.85	0.04	5.01	149.17	0.23	7.14	1257.64	0.92	6.63	757.92	0.89
	Price comparison	8.86	7073.34	0.91	4.54	94.00	0.18				6.62	750.40	0.90
	Display	8.66	5782.05	0.85	4.30	73.72	0.12						
$t = 1$	Newsletter	8.42	4525.44	0.88	5.56	260.55	0.27	6.19	490.26	0.68			
	Retargeting	9.21	9955.45	0.29				5.41	224.66	0.29	6.77	873.87	0.07
	Social media							5.45	232.40	0.46			
	Affiliate	9.15	9402.58	0.86	4.91	135.52	0.17				6.46	637.29	0.81
	Referrer				5.61	273.54	0.18	6.11	450.80	0.58	7.82	2498.47	0.94
	Other	7.67	2143.53	0.72									
	Type-in				1.55***	4.71	0.01***	1.87***	6.52	0.05***	1.70***	5.49	0.06***
	SEA	2.02***	7.51	0.01***	1.74***	5.71	0.01***	1.69***	5.43	0.04***	1.10***	3.01	0.03***
	SEO	−1.79***	0.17	0.00***	1.87***	6.48	0.01***	−6.89	0.00	−0.01***	1.05***	2.85	0.02***
	Price comparison	1.68***	5.37	0.01***	1.97***	7.16	0.01**				1.06***	2.88	0.03***
	Display	0.86***	2.36	0.00***	1.55***	4.69	0.01***						
$t = 2$	Newsletter	1.24***	3.47	0.00**	1.61***	5.01	0.01***	1.35***	3.85	0.03***			
	Retargeting	1.11***	3.04	0.00***				1.21***	3.36	0.02***	1.05***	2.86	0.02***
	Social media							1.53***	4.64	0.03***			
	Affiliate	0.68***	1.98	0.00***	1.45***	4.26	0.01***				0.96***	2.61	0.02***
	Referrer				1.76***	5.84	0.01***	1.97***	7.20	0.06***	1.56***	4.76	0.05***
	Other	0.96***	2.62	0.00***									
	Type-in				0.62***	1.86	0.00***	0.56**	1.76	0.01*	0.99***	2.69	0.02**
	SEA	0.73***	2.08	0.00***	0.63***	1.87	0.00***	0.96***	2.62	0.02***	0.55***	1.73	0.01**
	SEO	−1.67***	0.19	0.00***	0.78***	2.18	0.00***	−0.25	0.78	0.00	0.58**	1.78	0.01*
	Price comparison	0.87***	2.39	0.00***	−0.13	0.88	0.00				0.40*	1.49	0.01
	Display	0.62***	1.87	0.00***	0.58***	1.79	0.00***						
$t = 3$	Newsletter	0.40	1.49	0.00	0.74***	2.10	0.00***	0.69***	1.99	0.01***			
	Retargeting	0.47***	1.60	0.00***				0.74***	2.10	0.01***	0.48***	1.62	0.01***
	Social media							0.92***	2.52	0.02***			
	Affiliate	0.63***	1.88	0.00***	0.64***	1.89	0.00***				0.42***	1.52	0.01***
	Referrer				0.67***	1.95	0.00***	1.05***	2.87	0.02***	0.73***	2.07	0.01***
	Other	0.59***	1.81	0.00**									
	Type-in				0.96***	2.61	0.00***	1.73***	5.63	0.05***	1.77***	5.90	0.06**
	SEA	1.06***	2.88	0.00***	1.40***	4.07	0.01***	1.29***	3.61	0.03***	0.67***	1.96	0.01**
	SEO	−2.27***	0.10	0.00***	1.19***	3.29	0.01***	1.10***	3.02	0.02***	0.89***	2.43	0.02**
	Price comparison	1.27***	3.58	0.00***	1.57***	4.79	0.01				0.78***	2.18	0.02*
	Display	1.32***	3.76	0.00***	1.02***	2.77	0.00***						
$t = 4$	Newsletter	1.73***	5.63	0.01**	1.27***	3.54	0.01***	1.21***	3.37	0.02***			
	Retargeting	1.23***	3.41	0.00***							0.54***	1.72	0.01***
	Social media							1.44***	4.20	0.03***			
	Affiliate	1.59***	4.90	0.01***	1.09***	2.98	0.00***				0.56***	1.75	0.01***
	Referrer				1.06***	2.88	0.00***	1.52***	4.56	0.04***	0.59***	1.81	0.01***
	Other	1.05***	2.86	0.00***				0.00	0.00	0.00			
Intercept		−13.57	0.00		−10.57	0.00		−10.57	0.00		−11.43	0.00	
Observations		600,978			1,184,583			862,112			405,339		

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 15
Attribution results for second-order model (Data Set 2).

	Preceding channel								
Current channel	START	Type-in	SEA	SEO	Price comparison	Display	Newsletter	Affiliate	Referrer
Type-in	16.85%	11.33%	2.21%	1.69%	0.04%	0.77%	2.49%	1.08%	0.35%
SEA	10.25%	1.52%	5.25%	1.97%	0.03%	0.35%	0.89%	0.64%	0.24%
SEO	6.25%	1.13%	1.93%	4.70%	0.02%	0.16%	0.43%	0.49%	0.39%
Price comparison	0.07%	0.03%	0.03%	0.02%	0.03%	0.00%	0.01%	0.01%	0.00%
Display	0.62%	0.72%	0.33%	0.29%	0.01%	0.82%	0.23%	0.19%	0.06%
Newsletter	4.61%	1.84%	0.79%	0.50%	0.00%	0.22%	3.86%	0.18%	0.07%
Affiliate	4.24%	0.74%	0.52%	0.40%	0.02%	0.09%	0.12%	2.54%	0.14%
Referrer	1.26%	0.18%	0.16%	0.12%	0.01%	0.04%	0.07%	0.15%	0.19%

Table 16
Attribution results for second-order model (Data Set 3).

	Preceding channel							
Current channel	START	Type-in	SEA	SEO	Newsletter	Retargeting	Social media	Referrer
Type-in	13.26%	8.14%	1.49%	1.72%	0.06%	0.00%	2.36%	0.52%
SEA	11.07%	0.82%	3.74%	1.68%	0.09%	0.00%	1.00%	0.49%
SEO	12.23%	1.10%	1.61%	6.53%	0.05%	0.00%	0.62%	0.35%
Newsletter	0.71%	0.04%	0.04%	0.04%	0.29%	0.00%	0.03%	0.01%
Retargeting	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Social media	12.74%	1.16%	0.63%	0.58%	0.03%	0.00%	8.27%	0.29%
Referrer	4.07%	0.00%	0.28%	0.36%	0.01%	0.00%	0.35%	0.77%

Table 17
Attribution results for second-order model (Data Set 4).

	Preceding channel							
Current channel	START	Type-in	SEA	SEO	Price comparison	Retargeting	Affiliate	Referrer
Type-in	7.43%	4.30%	5.04%	0.83%	0.16%	0.20%	0.29%	0.17%
SEA	39.91%	1.37%	18.19%	1.72%	0.41%	0.28%	0.22%	0.19%
SEO	3.47%	0.38%	3.09%	2.26%	0.13%	0.03%	0.10%	0.06%
Price Comparison	1.20%	0.06%	0.57%	0.18%	0.72%	0.01%	0.02%	0.03%
Retargeting	0.10%	0.21%	1.01%	0.06%	0.03%	0.41%	0.00%	0.02%
Affiliate	0.28%	0.51%	1.21%	0.24%	0.18%	0.03%	0.92%	0.11%
Referrer	1.03%	0.08%	0.26%	0.03%	0.02%	0.00%	0.03%	0.19%

References

Abhishek, V., Fader, P. S., & Hosanagar, K. (2015). Media exposure through the funnel: A model of multi-stage attribution. Retrieved from http://ssrn.com/abstract=2158421

Anderl, E., Schumann, J. H., & Kunz, W. (2015). Helping firms reduce complexity in multichannel online data: A new taxonomy-based approach for customer journeys. *Journal of Retailing*, *92*(2), 185–203.

Archak, N., Mirrokni, V. S., & Muthukrishnan, S. (2010). Mining advertiser-specific user behavior using adfactors. In M. Rappa, & P. Jones (Eds.), *Proceedings of the 19th International Conference on World Wide Web* (pp. 31–40). ACM.

Berman, R. (2015). *Beyond the last touch*: *Attribution in online advertising*. Retrieved from http://ron-berman.com/papers/attribution.pdf

Bowman, D., & Narayandas, D. (2001). Managing customer-initiated contacts with manufacturers: The impact on share of category requirements and word-of-mouth behavior. *Journal of Marketing Research*, *38*(3), 281–297. http://dx.doi.org/10.1509/jmkr.38.3.281.18863.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30*(7), 1145–1159. http://dx.doi.org/10.1016/S0031-3203(96)00142-2.

Breuer, R., Brettel, M., & Engelen, A. (2011). Incorporating long-term effects in determining the effectiveness of different types of online advertising. *Marketing Letters*, *22*(4), 327–340. http://dx.doi.org/10.1007/s11002-011-9136-3.

Bronnenberg, B. J. (1998). Advertising frequency decisions in a discrete Markov process under a budget constraint. *Journal of Marketing Research*, *35*(3), 399–406. http://dx.doi.org/10.2307/3152037.

Bucklin, R. E., & Sismeiro, C. (2009). Click here for Internet insight: Advances in clickstream data analysis in marketing. *Journal of Interactive Marketing*, *23*(1), 35–48. http://dx.doi.org/10.1016/j.intmar.2008.10.004.

Che, H., & Seetharaman, P. B. (2009). "Speed of replacement": Modeling brand loyalty using last-move data. *Journal of Marketing Research*, *46*(4), 494–505. http://dx.doi.org/10.1509/jmkr.46.4.494.

Chierichetti, F., Kumar, R., Raghavan, P., & Sarlós, T. (2012). Are Web users really Markovian? In A. Mille, F. L. Gandon, J. Misselis, M. Rabinovich, & S. Staab (Eds.), *Proceedings of the 21st International Conference on World Wide Web* (pp. 609–618). ACM.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms* (3rd ed.). Cambridge, MA: Massachusetts Institute of Technology.

Dalessandro, B., Perlich, C., Stitelman, O., & Provost, F. J. (2012). Causally motivated attribution for online advertising. *ADKDD '12, Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy* (pp. 7:1–7:9). New York, NY, USA: ACM Retrieved from http://doi.acm.org/10.1145/2351356.2351363.

Danaher, P. J., & Dagger, T. S. (2013). Comparing the relative effectiveness of advertising channels: A case study of a multimedia blitz campaign. *Journal of Marketing Research*, *50*(4), 517–534. http://dx.doi.org/10.1509/jmr.12.0241.

Econsultancy (2012). *Marketing attribution: Valuing the customer journey*. London, UK: Econsultancy (Retrieved from http://econsultancy.com/de/reports/marketing-attribution-valuing-the-customer-journey).

Flosi, S., Fulgoni, G. M., & Vollman, A. (2013). If an advertisement runs online and no one sees it, is it still an ad? Empirical generalizations in digital advertising. *Journal of Advertising Research*, *53*(2), 192–199. http://dx.doi.org/10.2501/JAR-53-2-192-199.

Godfrey, A., Seiders, K., & Voss, G. B. (2011). Enough is enough! The fine line in executing multichannel relational communication. *Journal of Marketing*, *75*(July), 94–109.

De Haan, E., Wiesel, T., & Pauwels, K. (2016). The effectiveness of different forms of online advertising for purchase conversion in amultiple-channel attribution framework. *International Journal of Research in Marketing*, *33*(3), 491–507.

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284. http://dx.doi.org/10.1109/TKDE.2008.239.

Homburg, C., Steiner, V. V., & Totzek, D. (2009). Managing dynamics in a customer portfolio. *Journal of Marketing*, *73*(5), 70–89. http://dx.doi.org/10.1509/jmkg.73.5.70.

Ilfeld, J. S., & Winer, R. S. (2002). Generating website traffic. *Journal of Advertising Research*, *42*(5), 49–61.

Jansen, B. J., & Spink, A. (2009). Investigating customer click through behaviour with integrated sponsored and nonsponsored results. *International Journal of Internet Marketing and Advertising*, *5*(1–2), 74–94.

Jerath, K., Ma, L., & Park, Y. -H. (2014). Consumer click behavior at a search engine: The role of keyword popularity. *Journal of Marketing Research*, *51*(4), 480–486. http://dx.doi.org/10.1509/jmr.13.0099.

Jordan, P., Mahdian, M., Vassilvitskii, S., & Vee, E. (2011). The multiple attribution problem in pay-per-conversion advertising. In W. -Y. Ma, J. -Y. Nie, R. A. Baeza-Yates, T. -S. Chua, & W. B. Croft (Eds.), *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM.

Kamakura, W., Kopalle, P. K., & Lehmann, D. R. (2014). Empirical generalizations in retailing. *Journal of Retailing*, *90*(2), 121–124. http://dx.doi.org/10.1016/j.jretai.2014.05.005.

Kireyev, P., Pauwels, K., & Gupta, S. (2016). Do display ads influence search? Attribution and dynamics in online advertising. *International Journal of Research in Marketing*, *33*(3), 475–490.

Lambrecht, A., & Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research*, *50*(5), 561–576. http://dx.doi.org/10.1509/jmr.11.0503.

Lewis, R. A., Rao, J. M., & Reiley, D. H. (2011). Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. In W. -Y. Ma, J. -Y. Nie, R. A. Baeza-Yates, T. -S. Chua, & W. B. Croft (Eds.), *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM.

Li, H. A., & Kannan, P. K. (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, *51*(1), 40–56. http://dx.doi.org/10.1509/jmr.13.0050.

Little, J. D. C. (2004). Comments on "Models and managers: The concept of a decision calculus". *Management Science*, *50*(12 Supplement), 1854–1860. http://dx.doi.org/10.1287/mnsc.1040.0310.

Lodish, L. M. (2001). Building marketing models that make money. *Interfaces*, *31*(3), S45. http://dx.doi.org/10.1287/inte.31.3s.45.9681.

Mehta, N., Rajiv, S., & Srinivasan, K. (2003). Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Science*, *22*(1), 58–84. http://dx.doi.org/10.1287/mksc.22.1.58.12849.

Moffett, T. (2014). *Cross-channel attribution providers, Q4 2014*.

Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science*, *23*(4), 579–595. http://dx.doi.org/10.1287/mksc.1040.0073.

Neslin, S. A., & Shankar, V. (2009). Key issues in multichannel customer management: Current knowledge and future directions. *Journal of Interactive Marketing*, *23*(1), 70–81. http://dx.doi.org/10.1016/j.intmar.2008.10.005.

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of Marketing Research*, *43*(2), 204–211. http://dx.doi.org/10.1509/jmkr.43.2.204.

Osur, A. (2012). *The Forrester Wave: Interactive attribution vendors, Q2 2012*. (Cambridge, MA. Retrieved from Forrester website: http://www.forrester.com/The+Forrester+Wave+Interactive+Attribution+Vendors+Q2+2012/fulltext/-/E-RES72681).

Pfeifer, P. E., & Carraway, R. L. (2000). Modeling customer relationships as Markov chains. *Journal of Interactive Marketing*, *14*(2), 43–55. http://dx.doi.org/10.1002/(SICI)1520-6653(200021)14:2<43::AID-DIR4>3.0.CO;2-H.

Raman, K., Mantrala, M. K., Sridhar, S., & Tang, Y. (2012). Optimal resource allocation with time-varying marketing effectiveness, margins and costs. *Journal of Interactive Marketing*, *26*(1), 43–52. http://dx.doi.org/10.1016/j.intmar.2011.05.001.

Shankar, V., & Malthouse, E. C. (2007). The growth of interactions and dialogs in interactive marketing. *Journal of Interactive Marketing*, *21*(2), 2–4. http://dx.doi.org/10.1002/dir.20080.

Shao, X., & Li, L. (2011). Data-driven multi-touch attribution models. In C. Apté, J. Ghosh, & P. Smyth (Eds.), *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 258–264). ACM.

Sherman, L., & Deighton, J. (2001). Banner advertising: Measuring effectiveness and optimizing placement. *Journal of Interactive Marketing*, *15*(2), 60–64. http://dx.doi.org/10.1002/dir.1011.

Sood, A., James, G. M., & Tellis, G. J. (2009). Functional regression: A new model for predicting market penetration of new products. *Marketing Science*, *28*(1), 36–51. http://dx.doi.org/10.1287/mksc.1080.0382.

Styan, G. P. H., & Smith, H. (1964). Markov chains applied to marketing. *Journal of Marketing Research*, *1*(1), 50–55. http://dx.doi.org/10.2307/3150320.

The CMO Club & Visual IQ, Inc (2014). *Big data*, *attribution* & *omni-channel*. Retrieved from http://www.visualiq.com/sites/default/files/PDF-download-graphics/study_building_bridges_to_the_promised_land.pdf

Tucker, C. (2012). The implications of improved attribution and measurability for online advertising markets. *Competition in the Online Environment* (Retrieved from https://www.ipria.org/events/conf/Competition_Conference/Measureabiltiy_and_online_ads.pdf).

Wiesel, T., Pauwels, K., & Arts, J. (2011). Marketing's profit impact: Quantifying online and off-line funnel progression. *Marketing Science*, *30*(4), 604–611. http://dx.doi.org/10.1287/mksc.1100.0612.

Xu, L., Duan, J. A., & Whinston, A. B. (2014). Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science*, *60*(6), 1392–1412. http://dx.doi.org/10.1287/mnsc.2014.1952.

Yang, S., & Ghose, A. (2010). Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence? *Marketing Science*, *29*(4), 602–623. http://dx.doi.org/10.1287/mksc.1090.0552.