MASTER

Attribution modeling
using conversion value as an alternative attribution measure to understand the customer journey online

Alblas, R.

Award date:
2018

Link to publication
Attribution Modeling

USING CONVERSION VALUE AS AN ALTERNATIVE ATTRIBUTION MEASURE TO UNDERSTAND THE CUSTOMER JOURNEY ONLINE

by
R. (Robbert) Alblas

BSc Technische Innovatiewetenschappen – TU/e 2015
Student identity number 0791201

In partial fulfilment of the requirements for the degree of

Master of Science
in Innovation Management

Supervisors
Prof.dr. E.J. (Ed) Nijssen, TU/e, ITEM
Dr. S. (Shantanu) Mullick, TU/e, ITEM
S. (Stefan) Lammers, Deloitte
H.H. (Harald) van Doorn, Deloitte
Subject headings: attribution, Markov model, customer journey, clickstream, multichannel, digital marketing, marketing channels, buyer behavior, attribution measure, conversion, conversion value
Abstract

Recent figures indicate an increasing popularity of the digital environment in marketing, advertising and sales. Nowadays, customers increasingly use a vast amount of online marketing channels to gather information and buy products, resulting in more complex customer journeys online. Although this complicates estimation of online channel performance, companies still apply simple approaches to obtain channel attribution estimates. Consequently, scholars have proposed a variety of algorithmic multi-touch attribution models through which significant progress in estimating channel performance has been made. However, scholars do not address the influence of the attribution measure on marketing channel attribution estimates. Some studies take conversion events as attribution measure, others consider value resulting from conversions. Those studies that do take conversion value as attribution measure do not investigate on what aspects conversions of high value differ from conversion of low value. In response, this research applies a graph-based Markovian attribution model to two individual-level data sets from two online-only companies operating in different industries, each entailing ten distinct online channels. The results indicate that attribution measures do influence attribution estimates. Moreover, conversions of high value are shown to differ from conversions of low value in terms of channel usage, customers’ prior buying experience, and in frequency of interactions. In this way, by using conversion value as an alternative attribution measure, the study can help marketers to better understand the customer journey online.
Management Summary

Problem Introduction

Recent figures indicate an increasing popularity of the digital environment in marketing, advertising and sales (Philips, 2015). With more offline customers shifting to digital technologies and "younger, digitally oriented consumers entering the ranks of buyers", the Internet has become a mainstream sales channel (Bughin, 2015). Consequently, online advertising has grown from $9.6 billion in 2004 to a $42.8 billion industry in 2013 (IAB, 2014). As the new digital environment has provided new online marketing channels 1 such as paid search and display marketing, customers increasingly use these channels to gather information and buy products. This proliferation of online marketing channels has caused customers to interact with firms through myriad touchpoints 2 in multiple channels, resulting in more complex customer journeys 3 (Lemon & Verhoef, 2016).

When a customer converts, i.e. makes a purchase, companies want to estimate to what extent their marketing channels have contributed to that decision. Increased complexity of customer journeys complicates such estimations. Still, to obtain their estimates, many companies apply attribution models that are based on simple heuristics such as last-touch attribution, in which the marketing channel that enabled the last touchpoint prior to a purchase decision gets all the credits for it (Li & Kannan, 2014; The CMO Club & Visual IQ, Inc., 2014). However, basic attribution models rely on subjective judgment, ignore timing and sequence of online channels in customer journeys, and disregard paths that did not lead to a purchase decision (Chandler-Peplnjak, 2008; Li & Kannan, 2014). Companies have come to acknowledge these flaws and have expressed the need for more advanced attribution models (Kathibloo, 2010; New York Times, 2012; Szulc, 2012).

In response, scholars have proposed a variety of algorithmic multi-touch attribution models, including logistic regression models (Shao & Li, 2011), game theory approaches (Dalessandro, Perlich, Stitelman, & Provost, 2012), Bayesian models (Li & Kannan, 2014), and Markov models (Anderl, Becker, von Wangenheim, & Schumann, 2016). Instead of subjectively and manually deciding how to credit marketing channels, algorithmic attribution models use advanced statistical modeling to analyze historical data and objectively estimate channel performance (Dalessandro, Perlich, Stitelman, & Provost, 2012). As such, prior research has made significant progress in estimating channel performance and has demonstrated the superiority of algorithmic attribution models.

However, none of the studies in the field of multichannel marketing attribution considers the influence of the attribution measure on marketing channel attribution estimates. The attribution measure, referring to the outcome of interest when estimating marketing channel performance, differs throughout literature. Some studies take conversion events as the attribution measure (e.g. Shao & Li, 2011), others consider revenue or profits resulting from conversion (e.g. Danahar & Dagger, 2013), but none actually argue why they focus on a certain attribution measure nor do they dwell on how the attribution

---

1 The term “online marketing channel” is used as an umbrella phrase referring to online contact point through which the firm and the customer interact (Neslin, et al., 2006).

2 Touchpoints are clearly defined customer interactions with a company, being a combination of a channel, device, and user task, i.e., the customer’s need (Flaherty, 2016).

3 An online customer journey or “path” of an individual customer is defined as including all touch points with a company over its marketing channels prior to a potential purchase decision that lead to a visit of the company’s website (de Haan, Wiesel, & Pauwels, 2013)
measure itself could influence attribution estimates. Marketing channels can influence the probability that a customer converts, but may affect the amount customers spend when they make a purchase differently. Therefore, attribution estimates can alter when including value resulting from conversion in the analysis. Examining such differences may lead to substantive insights into the influence of attribution measures on attribution estimates. This, in return, may help guide managers to make well-informed decisions concerning the attribution measure they use in channel attribution analyses.

Those multichannel attribution studies that take conversion value as attribution measure do not investigate to what extent high-value conversions differ from low-value conversions. In order to ensure current and future revenue growth, companies must focus on those high-value conversions and understand how they are currently performing in driving them (Rawson, Duncan, & Jones, 2013; Hogan, Almquist, & Glynn, 2005). High-value conversions may belong to customers that differ in marketing channel usage, prior buying experience or browsing behavior. To date, no study has focused on such differences between high-value conversions and low-value conversions. This leads to the following problem statement and research questions:

"Current multichannel marketing attribution studies do not provide insights into how attribution measures influence attribution estimates and do not investigate the extent to which high-value conversions differ from low-value conversions, leading to potentially biased inferences concerning marketing channel performance and a lack of understanding about conversions of different value.”

**RQ1:** To what extent are online channel attribution estimates affected by the attribution measure?

**RQ2:** To what extent and on which aspects do high-value conversions differ from low-value conversions?

**Approach**

To achieve the research objectives and provide answers to the research questions, this study applies a graph-based Markovian attribution model. This algorithmic multi-touch attribution model attempts to mine frequent substructures in graph data (Archak, Mirrokni, & Muthukrishnan, 2010). Customer journeys are modeled as chains in first- and higher-order Markov graphs, using a property called removal effect to determine the contribution of online channels and channel sequences. Markov models do not impose a priori constraints on the number of channels and customer paths (Chang & Zhang, 2016), and its graph-based structure reflects the sequential nature of customer journeys (Anderl, Becker, von Wangenheim, & Schumann, 2016). Especially for online retailers operating in the digital environment, having the ability to track responses and performance of their online marketing channels almost instantaneously (Kannan & Li, 2016; Shao & Li, 2011), algorithmic attribution models are of good use. Therefore the Markov model is applied to two large, real-world individual-level data sets from two online-only companies operating in different industries (i.e. B2B and B2C). Analyzing data from companies operating in different industries enables to derive cross-industry generalizations and obtain industry-specific findings.

---

4 Conversion value refers to the amount customers spend when they make a purchase, i.e. transaction value.

5 What is considered to be a "high-value" conversion or “low-value” conversion may vary as it is dependent on product, industry and market. This is to be decided by managers themselves.
Main Findings and Contributions

First, this research contributes novel insights into the impact of the dependent variable, i.e., the influence of the attribution measure, on attribution estimates. By comparing channel attribution estimates from using conversion with channel attribution estimates from using conversion value as attribution measure, this study finds that attribution estimates of firm-initiated channels become greater when using conversion value as attribution measure, whereas attribution estimates of customer-initiated channel become smaller. Although these results are consistent across both data sets, findings do indicate there is a significant influence of industry context on attribution estimates.

Second, including conversion value enables to divide conversions into low-value conversions and high-value conversions. This research contributes to a better understanding of customer journeys by revealing that high-value conversions differ from low-value conversions in terms of channel usage, customers’ prior buying experience, and in frequency of interactions. Price comparison sites and e-mail are consistently found to be of greater importance in driving high-value conversions than in driving low-value conversions, which implies that companies should invest more heavily in those specific channels to increase value of conversion. Furthermore, the share of returning customers is slightly greater in high-value conversions than in low-value conversions. Hence, returning customers are of greater value to companies than new customers and should therefore require more attention. Finally, frequency of interactions is found to be significantly greater for high-value conversions than for low-value conversions. This finding appears to be stronger for customers in a B2C setting, such that especially in that environment, a greater frequency of interactions should be pursued.

Finally, channel attribution estimates from using the Markov model are compared to those of three commonly applied basic attribution models, namely first-touch, last-touch and linear attribution. Prior research indicates that basic attribution models can produce incorrect conclusions regarding channel performance (Li & Kannan, 2014; Xu, Duan, & Whinston, 2014; Anderl, Becker, von Wangenheim, & Schumann, 2016). In this research, a comparison of channel attribution estimates across two data sets provides further indications that basic attribution models are not capable of accurately estimating online marketing channel performance.

Limitations and Further Research

The data sets in this research do not include any information on offline marketing channels. Prior research has shown that online channels do interact with offline channels and create synergies (Kannan & Li, 2016; Joo, Wilbur, Cowgill, & Zhu, 2014). This research does not deny that these interactions and synergies exist, but is simply limited by the unavailability of data about exposures and responses to offline marketing channels. Furthermore, sufficient methods to track exposures and responses to offline marketing channels are currently lacking (Kaushik, 2008). Future research may contribute by developing such methods and including information about offline channels in the analysis.

Finally, attribution modeling is endogenic: it measures the relative effectiveness of channels in a given setting (Li & Kannan, 2014), so results from this study may not apply for other companies in a different setting. However, there are strong indications that industry context affects channel attribution estimates. This is likely to be caused by different customer buyer behavior in B2C markets compared to customer buyer behavior in B2B markets. Hence, future studies might more prominently anticipate and address such differences.
Preface

Writing this final piece of text, I realize that my master thesis project has come to an end. It has been a very turbulent six months. I swapped houses twice, moving from Eindhoven to Amsterdam, and back again to my hometown, Breda. I met a Finnish girl I’ve fallen in love with, making trips to Helsinki on the regular, and throughout the last year I’ve met tons of new people at Deloitte. Throughout the whole period, one thing remained consistent, that is, investing a great amount of time in this document.

I can state with great certainty that writing my master thesis has required the same time investment as studying for all my exams combined. During your time as a student, most things are scripted. You participate in mandatory courses, choose some free electives and occasionally hand-in an assignment that is part of a course. Once you finish one course, the next one has already been set up for you. This is completely different for the master thesis project. From scratch, you will have to build it yourself. This gives you loads of freedom, but can also drown you in information and overwhelm you. Without guidance, soon you’ll become lost in the world of science.

Thankfully, I have had two company supervisors who have guided me throughout the whole process. Harald, thanks for your support. You have opened doors for me I couldn’t have opened myself. Without you, I do not know whether I could have found the right entry to companies and take care of those dreadful NDAs. Special thanks go out to Stefan. You’ve really made me feel welcome at Deloitte, invested a large amount of time in helping me out, and most of all, have shown to be real good support in times I felt a bit down or unmotivated. I sincerely couldn’t have wished for better company supervisors.

Furthermore, I would like to thank my two university supervisors, Ed Nijssen and Shantanu Mullick. Both of you have invested a great amount of time and effort in providing lots of valuable feedback. Your criticism has definitely helped to take my master thesis to greater heights. In hindsight, I do realize that your comments have played a big role in keeping my thesis on track.

Most importantly, I want to thank my friends, family and girlfriend. Mom, dad, thanks for all the emotional support (and financial aid) you have provided me. Not only during the period of writing this thesis, but throughout my whole life. I am really happy to be able to rely on your unconditional support. Finally, a big ‘thank you’ to Janina, who stood by me at all times. I feel blessed I’ve met you and hope we can both enjoy life after completing this thesis even more than we already do now.
# Table of Contents

ABSTRACT ................................................................................................................................. I

MANAGEMENT SUMMARY ........................................................................................................ II

PREFACE ...................................................................................................................................... V

1. INTRODUCTION ..................................................................................................................... 1
   1.1 PROBLEM INTRODUCTION ............................................................................................. 1
   1.2 RESEARCH QUESTIONS AND METHODOLOGY ............................................................. 2
   1.3 ACADEMIC AND MANAGERIAL RELEVANCE ............................................................... 3
   1.4 REPORT STRUCTURE ....................................................................................................... 3

2. RESEARCH BACKGROUND ...................................................................................................... 4
   2.1 UNDERSTANDING MULTICHANNEL ATTRIBUTION ....................................................... 4
       2.1.1 Digital Marketing ..................................................................................................... 4
       2.1.2 The Online Customer Journey .............................................................................. 4
       2.1.3 Online Marketing Channels .................................................................................. 5
       2.1.4 Online Multichannel Marketing Attribution ........................................................ 5
   2.2 LITERATURE REVIEW OF ATTRIBUTION STUDIES ................................................... 7
       2.2.1 Online Multichannel Marketing Attribution Studies .............................................. 7
       2.2.2 Attribution Measures throughout Literature .......................................................... 9
       2.2.3 The Influence of Industry Context ......................................................................... 10
       2.2.4 Literature Gaps ...................................................................................................... 10
   2.3 PROPOSITIONS ................................................................................................................ 11
       2.3.1 The influence of Attribution Measures on Attribution Estimates .......................... 11
       2.3.2 High-Value Conversions versus Low-Value Conversions ...................................... 12

3. DATA ....................................................................................................................................... 14
   3.1 COMPANY CONTEXT ...................................................................................................... 14
   3.2 DATA DESCRIPTION ....................................................................................................... 14
       3.2.1 Sampling Procedure ............................................................................................... 15
       3.2.2 Online Marketing Channels .................................................................................. 15
       3.2.3 Variables of Interest .............................................................................................. 16

4. METHODOLOGY ................................................................................................................... 17
   4.1 MARKOVIAN ATTRIBUTION MODEL .......................................................................... 17
       4.1.1 Markov Graphs ..................................................................................................... 17
       4.1.2 First-order Markov Model ..................................................................................... 18
       4.1.3 Higher-order Markov Model .................................................................................. 19
       4.1.4 The Removal Effect .............................................................................................. 19

5. RESULTS .................................................................................................................................. 20
   5.1 MODEL FIT ....................................................................................................................... 20
       5.1.1 Predictive Accuracy ............................................................................................... 20
       5.1.2 Robustness ........................................................................................................... 22
       5.1.3 Determining Markov Order ................................................................................... 22
   5.2 COMPARING THE MARKOV MODEL .............................................................................. 23

VI
CHAPTER 1
Introduction

1.1 Problem Introduction

Recent figures indicate an increasing popularity of the digital environment in marketing, advertising and sales (Philips, 2015). With offline customers shifting to digital technologies and “younger, digitally oriented consumers entering the ranks of buyers”, the Internet has become a mainstream sales channel (Bughin, 2015). Consequently, online advertising has grown from $9.6 billion in 2004 to a $42.8 billion industry in 2013 (IAB, 2014). As the new digital environment has provided new online marketing channels\(^6\) such as paid search and display marketing, customers increasingly use these channels to gather information and buy products. The proliferation of online marketing channels has caused customers to interact with firms through myriad touchpoints\(^7\) in multiple channels, resulting in more complex customer journeys\(^8\) (Lemon & Verhoef, 2016). This complexity is only likely to increase with the number of touchpoints growing by over 20% annually (Bughin, 2015).

When a customer converts, i.e. makes a purchase, companies want to estimate to what extent their marketing channels have contributed to that decision. Increased complexity of customer journeys complicates such estimations. This is referred to as the attribution problem. Still, to obtain their estimates, many companies apply attribution models that are based on simple heuristics such as last-touch attribution, in which the marketing channel that enabled the last touchpoint prior to a purchase decision gets all the credits for it (Li & Kannan, 2014; The CMO Club & Visual IQ, Inc., 2014). However, basic attribution models rely on subjective judgment, ignore timing and sequence of online channels in customer journeys, and disregard paths that did not lead to a purchase decision (Chandler-Peplnjak, 2008; Li & Kannan, 2014). Companies have come to acknowledge these flaws and have expressed the need for more advanced attribution models (Kathibloo, 2010; New York Times, 2012; Szulc, 2012).

In response, scholars have proposed a variety of algorithmic multi-touch attribution models, including logistic regression models (Shao & Li, 2011), game theory approaches (Dalessandro, Perlich, Stitelman, & Provost, 2012), Bayesian models (Li & Kannan, 2014), and Markov models (Anderl, Becker, von Wangenheim, & Schumann, 2016). Instead of subjectively and manually deciding how to credit marketing channels, algorithmic attribution models use advanced statistical modeling to analyze historical data and objectively estimate channel performance (Dalessandro, Perlich, Stitelman, & Provost, 2012). As such, prior research has made significant progress in estimating channel performance and has demonstrated the superiority of algorithmic attribution models.

However, none of the studies in the field of multichannel marketing attribution considers the influence of the attribution measure on marketing channel attribution estimates. The

\(^6\) The term “online marketing channel” is used as an umbrella phrase referring to online contact point through which the firm and the customer interact (Neslin, et al., 2006).

\(^7\) Touchpoints are clearly defined customer interactions with a company, being a combination of a channel, device, and user task, i.e., the customer's need (Flaherty, 2016).

\(^8\) An online customer journey or “path” of an individual customer is defined as including all touch points with a company over its marketing channels prior to a potential purchase decision that lead to a visit of the company’s website (de Haan, Wiesel, & Pauwels, 2013)
attribution measure, referring to the outcome of interest when estimating marketing channel performance, differs throughout literature. Some studies take conversion events as the attribution measure (e.g. Shao & Li, 2011), others consider revenues or profits resulting from conversion (e.g. Danahar & Dagger, 2013), but none actually argue why they focus on a certain attribution measure nor do they dwell on how the attribution measure itself could influence attribution estimates. Marketing channels can influence the probability that a customer converts, but may affect the amount customers spend when they make a purchase differently. Therefore, attribution estimates can alter when including value resulting from conversion in the analysis. Examining such differences may lead to substantive insights into the influence of attribution measures on attribution estimates. This, in return, may help guide managers to make well-informed decisions concerning the attribution measure they use in channel attribution analyses.

Those multichannel attribution studies that take conversion value as attribution measure do not investigate to what extent high-value conversions differ from low-value conversions. In order to ensure current and future revenue growth, companies must focus on those high-value conversions and understand how they are currently performing in driving them (Rawson, Duncan, & Jones, 2013; Hogan, Almquist, & Glynn, 2005). High-value conversions may belong to customers that differ in marketing channel usage, prior buying experience or browsing behavior. To date, no study has focused on such differences between high-value conversions and low-value conversions. This leads to the following problem statement:

"Current multichannel marketing attribution studies do not provide insights into how attribution measures influence attribution estimates and do not investigate the extent to which high-value conversions differ from low-value conversions, leading to potentially biased inferences concerning marketing channel performance and a lack of understanding about the properties of conversions of different value."

1.2 Research Questions and Methodology

This study aims to fill this gap by examining the difference in online channel attribution estimates resulting from taking a conversion event or the resulting conversion value, i.e. the amount people spend when they make a purchase, as the attribution measure in multichannel attribution analysis. Furthermore, this research investigates the extent to which high-value conversions differ from low-value conversions. This is done in terms of channel importance, customers’ prior buying experience and frequency of interactions. Thus, this research investigates two key research questions:

**RQ1:** To what extent are online channel attribution estimates affected by the attribution measure?

**RQ2:** To what extent and on which aspects do high-value conversions differ from low-value conversions?

To achieve the research objectives and provide answers to the research questions, this study applies a graph-based Markovian attribution model. This algorithmic multi-touch attribution model attempts to mine frequent substructures in graph data (Archak, Mirrokni, 2012, Mirrokni & Archak, 2011).
Customer journeys are modeled as chains in first- and higher-order Markov graphs, using a property called removal effect to determine the contribution of online channels and channel sequences. Markov models do not impose a priori constraints on the number of channels and customer paths (Chang & Zhang, 2016), and its graph-based structure reflects the sequential nature of customer journeys (Anderl, Becker, von Wangenheim, & Schumann, 2016). Especially for online retailers operating in the digital environment, having the ability to track responses and performance of their online marketing channels almost instantaneously (Kannan & Li, 2016; Shao & Li, 2011), algorithmic attribution models are of good use. Therefore the Markov model is applied to two large, real-world individual-level data sets from two online-only companies operating in different industries. Analyzing data from companies operating in different industries enables to derive cross-industry generalizations and obtain industry-specific findings.

1.3 Academic and Managerial Relevance

First, this research contributes to existing literature by providing novel insights into the impact of the dependent variable, i.e. the influence of the attribution measure, on attribution estimates in online multichannel marketing settings. “Good” attribution measurement may rely by and large on the subjective eyes of the beholder (Dalessandro, Perlich, Stitelman, & Provost, 2012). Thus, as for-profit organizations are driven by revenue, it would make more sense to measure channels on their ability to drive conversion value instead of conversion events. Still, most companies focus on the latter. Potential differences in attribution estimates resulting from using different attribution measures may indicate that using conversion events as the attribution measure leads to suboptimal budget allocations across marketing channels.

Second, existing research has not focused on differences between high- and low-value conversions. This study fills this gap and contributes to a better understanding of customer journeys by revealing aspects on which high-value conversions differ from low-value conversions. Successful companies consciously resist investing on every touchpoint, marketing channel and customer journey. Instead, they identify and spend aggressively only on those interactions they know will have the most impact on revenue growth and profitability (Hogan, Almquist, & Glynn, 2005). The insights resulting from this study can guide companies in identifying and driving those conversions that are most likely to be of high value.

Finally, this study provides the participating companies with a better estimation of the performance of their online marketing channels. Both companies currently apply basic attribution models relying on simple heuristics. Basic attribution models have shown to impair attribution measurement, leading to biased inferences concerning channel performance and suboptimal allocation of marketing budgets (Li & Kannan, 2014). In order to obtain more accurate channel attribution estimates this may encourage the participating companies to start using algorithmic multi-touch attribution models instead.

1.4 Report Structure

In the next chapter, concepts and constructs that are relevant in the field of multichannel attribution will be clarified. In addition, chapter two provides an overview of multichannel marketing attribution literature. It concludes with several propositions. Chapters three and four, respectively, describe the data and methodology. Chapter five validates the Markov model and presents the results. Finally, key findings, managerial implications, limitations, and further research directions are discussed in chapters six and seven.
CHAPTER 2
Research Background

2.1 Understanding Multichannel Attribution

To provide definitional clarity, this section contains a brief overview of concepts and constructs that are relevant for this study.

2.1.1 Digital Marketing

Marketing can be seen as the connection between people and products, customers and companies (Burnett, 2008). It involves attracting customer, persuading them to maximize their purchases, and building customer satisfaction and loyalty (Jones & Baron, 1991).

Nowadays, marketing and sales is increasingly moving more towards the digital world. This is a logic result from more offline customers shifting to digital technologies and “younger, digitally oriented consumers entering the ranks of buyers” (Bughin, 2015). Recent figures indicate an overall increasing popularity of the digital environment in marketing, advertising and sales (Philips, 2015; IAB, 2014). “Digital marketing” has evolved over time from a specific term describing the marketing of products and services using digital channels – to an umbrella describing the process of using digital technologies to acquire customers and build customer preferences, promote brands, retain customers and increase sales (Kannan & Li, 2016).

2.1.2 The Online Customer Journey

The new digital environment enables companies to track online customer responses and marketing performance almost instantaneously (Kannan & Li, 2016; Shao & Li, 2011). By recording the customer’s Internet activity companies can trace the navigational path the customer takes (Bucklin & Sismeiro, 2009). For each visit to the advertiser’s website, data is gathered regarding the marketing channel that was used by the customer, the timestamp, and whether the customer made a purchase and the size of that purchase. These data points are then used to construct online customer journeys that describe the pattern of touchpoints of individual customers across marketing channels prior to a potential purchase decision (Anderl, Becker, von Wangenheim, & Schumann, 2016). In a customer journey analysis, firms focus on how customers interact with these touchpoints (Lemon & Verhoef, 2016). In this analysis, the goal is to understand the many possibilities and paths a customer may take to complete his or her “job”.

Despite considerable dissent regarding exact definitions (Hauser & Wernerfelt, 1990), marketing academics generally agree that customer journeys are multistage processes (Roberts & Lattin, 1997). In its broadest sense, the customer journey consists of three stages (Howard & Sheth, 1969; Neslin, et al., 2006). The first stage – pre-purchase – encompasses behaviors such as need recognition, search, and consideration. The second stage – purchase – covers all touchpoints during the purchase event itself. It is characterized by behaviors such as choice, ordering, and payment. The third stage – post-purchase – encompasses touchpoints following the actual purchase. It includes behaviors such as usage and consumption, engagement, and service request (Lemon & Verhoef, 2016). Figure 1 displays the full customer journey.
Figure 1. The full customer journey

The first stages of the online customer journey, as customers move from need recognition, search, consideration and choice to the actual purchase, are of great importance for digital marketers and are therefore focused on in this research. Firms should attempt to identify specific touchpoints that lead customers to continue or discontinue this path (Lemon & Verhoef, 2016).

2.1.3 Online Marketing Channels

A marketing channel is defined as a contact point through which the firm and the customer interact (Neslin, et al., 2006), whereas touchpoints refer to clearly defined customer interactions with a company, being a combination of a channel, device, and user task, i.e., the customer’s need (Flaherty, 2016). Examples of traditional marketing channels are television, radio, magazines and newspapers. Now, the digital environment has provided new online marketing channels. Online retailers can choose from a vast array of online channels such as retargeted displays, search engine advertising (SEA) and email, with European marketers reporting to use an average of seven channels in parallel (Teradata Corporation, 2013). Table 13 in Appendix A provides an overview and description of some commonly deployed online marketing channels.

Scholars have proposed various ways to classify online marketing channels, including brand usage (Anderl, Schumann, & Kunz, 2015), browsing goal (Klapdor, 2013), content-integration (de Haan, Wiesel, & Pauwels, 2016), and degree of personalization (de Haan, Wiesel, & Pauwels, 2013). An overview of classifications of online marketing channels is given in Table 14 in Appendix B. The only classification of marketing channels academics do agree upon is along the origin of the contact (de Haan, Wiesel, & Pauwels, 2013; Li & Kannan, 2014). Marketing activities traditionally have been initiated by the firm, but online marketing channels are often initiated by customers (Shankar & Malthouse, 2007). Although a display advertisement is still initiated by the firm, customers can also visit the advertisers’ websites on their own initiative – for instance, by directly typing in the related Web address. Whereas the first is a firm-initiated channel (FIC), where the firm initiates the marketing communication, the latter depicts a customer-initiated channel (CIC), which is triggered by potential customers, on their own initiative (Li & Kannan, 2014).

Considering the proliferation of online marketing channels, classifying them into meaningful categories is used to reduce complexity in multichannel attribution analyses.

2.1.4 Online Multichannel Marketing Attribution

Understanding the online customer journey is a major consideration when studying customer experience (de Haan, Wiesel, & Pauwels, 2013; Xu, Duan, & Whinston, 2014). Improving this experience tops the list of business priorities companies have (Accenture, 2015). One of the key questions in understanding this process is how each individual online channel is evaluated in terms of its contribution to purchase decisions and sales. Typically, touchpoints over multiple online marketing channels have delivered impressions to a
customer. When the customer then makes a purchase decision, the company wants to
determine which touchpoints through which channels have contributed to that decision.
This step is critical in completing the feedback loop so that one can analyze, report and
optimize the purchase decision process.

However, given the proliferation of online channels and complexity of online customer
journeys, measuring the degree to which channels contribute to purchase decisions and
sales has become a demanding task (Anderl, Becker, von Wangenheim, & Schumann,
2016). This problem of interpreting the influence of marketing channels to the customer’s
decision process and assigning credit to them for driving the customer to desirable actions
such as making a purchase is called the attribution problem (Shao & Li, 2011).

The goal of attribution modeling is to provide an answer to the attribution problem by
pin-pointing the credit assignment of a purchase decision to touchpoints, as illustrated in
Figure 2. The resulting assignment can be aggregated along different dimensions,
including online marketing channel, to derive overall insights (Shao & Li, 2011).

![Figure 2. Multi-touch attribution](image)

To determine which online marketing channel is to be credited for a purchase decision,
initially a simple rule was developed: the last touchpoint, and corresponding marketing
channel, prior to a purchase decision receives 100% of the credits (Li & Kannan, 2014;
The CMO Club & Visual IQ, Inc., 2014). This last-touch attribution principle is simple, but
completely ignores the influences of all touchpoints - and online channels - except the last
one (Chandler-Peplnjak, 2008). In practice, the multiple touchpoints prior to conversion
are still rarely taken into account when estimating online channel performance. Instead,
heuristics like last-touch and first-touch attribution are used. Sometimes the uniform,
weighted, or exponential attribution model, which considers all touchpoints leading up to
a conversion and allocates the credit for the conversion (value) accordingly, is used (Li &
Kannan, 2014; The CMO Club & Visual IQ, Inc., 2014). Some commonly applied basic
attribution models are listed in Figure 3.

In short, basic attribution models rely on subjective judgment, ignore timing and
sequence of online channels in customer journeys, and disregard paths that do not lead
to a purchase decision (Chandler-Peplnjak, 2008; Li & Kannan, 2014). However, use of
online marketing channels may not result in immediate purchases, but they do stimulate
subsequent usage of marketing channels which then leads to purchases (Xu, Duan, &
Whinston, 2014). Also, when optimizing marketing operations, it is of special interest why
certain customers have not converted, instead of only being able to optimize the journeys
of converting customers. Hence, the inability to track non-conversions results in a large
amount of valuable information being lost. These shortcomings of basic attribution models
impair the measurement of the impact of online marketing channels on purchase decisions
and resulting value from purchases (Google Analytics Help, 2017).
Figure 3. (a) first-touch, (b) last-touch, (c) linear, and (d) time decay attribution

To provide guidance in developing more “appropriate” attribution models, Dalessandro et al. (2012) lists properties attribution models should have. First, an attribution model should be fair, i.e., it should reward an individual channel in accordance with its ability to affect likelihood of conversion and value. Secondly, attribution models should be easy to interpret. Finally, as the customer’s purchase decision process is largely dependent on the company, the product it offers, and how it’s offered, a desirable attribution model should be driven by a solid statistical analysis of customer response data (Shao & Li, 2011).

In line with these criteria and as a better alternative to basic attribution models, academics have proposed algorithmic multi-touch attribution models. Algorithmic attribution models are fair and often easy to interpret, and, instead of applying heuristic rules, they use advanced statistical modeling to extensively analyze historical data and objectively estimate online marketing channel performance (Järvinen & Karjaluoto, 2015). Examples of such algorithmic multi-touch attribution models are Bayesian models, hazard models, VAR models, and the Markov model used in this research. How these models have been applied in prior research will be covered in the next section of this report.

2.2 Literature Review of Attribution Studies

This section of the report embeds the focal research in the field of online multichannel marketing attribution studies. It outlines state of the art knowledge and insights from prior research related to this study. Furthermore, it identifies and lists those gaps in existing literature that are being pursued in this research.

2.2.1 Online Multichannel Marketing Attribution Studies

In the early stages of online marketing channel attribution research, studies only investigated the effectiveness of single channels (e.g. display or search) in isolation. An overview of these studies is provided in Table 15 in Appendix C. Although these studies examined ways to improve the effectiveness of individual marketing channels or investigated the relationship between two particular channels, they were of limited use for strategic decision making with respect to budget allocation across multiple marketing channels (de Haan, Wiesel, & Pauwels, 2016).

Only recently, likely due to increasing data availability (Anderl, Schumann, & Kunz, 2015), academic research has focused on marketing channel performance in online multichannel environments. An overview of those studies that have covered an extensive range of channels can be found in Table 1.
<table>
<thead>
<tr>
<th>Study</th>
<th>Online / Offline</th>
<th>Channels</th>
<th>Attribution Measure</th>
<th>Data Type</th>
<th>Method</th>
<th>Market</th>
<th>Insights</th>
<th>Segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shao &amp; Li (2011)</td>
<td>Online</td>
<td>Display, search, social, email, video</td>
<td>Conversion</td>
<td>Field data</td>
<td>Bagged Logistic Regression + Prob. model</td>
<td>Consumer software and services</td>
<td>Last-touch attribution models are inaccurate</td>
<td>No</td>
</tr>
<tr>
<td>Danaher &amp; Dagger (2013)</td>
<td>Online + Offline</td>
<td>TV, radio, mail, newspaper, magazine, display, search, social media, catalog</td>
<td>Revenue, Profits</td>
<td>Field data + Survey</td>
<td>Type II Tobit</td>
<td>Department store with large range of products</td>
<td>7 out of 10 channels significantly influence purchase outcomes</td>
<td>No</td>
</tr>
<tr>
<td>Li &amp; Kannan (2014)</td>
<td>Online</td>
<td>Display, SEA, SEO, email, referral, direct type-in</td>
<td>Traffic, Conversion</td>
<td>Field data + Field experiment</td>
<td>Hierarchical Bayes Model</td>
<td>Hospitality</td>
<td>Carryover and spillover effects exist between marketing channels</td>
<td>No</td>
</tr>
<tr>
<td>Anderl et al. (2015)</td>
<td>Online</td>
<td>Affiliate, display, email, price comparison, retargeting, SEA, SEO, direct type-in, partners</td>
<td>Conversions</td>
<td>Field data</td>
<td>Proportional hazard model</td>
<td>Fashion</td>
<td>Channel taxonomy reveals interaction effects over time</td>
<td>No</td>
</tr>
<tr>
<td>De Haan et al. (2016)</td>
<td>Online + Offline</td>
<td>TV, radio, email, SEA, retargeting, referral, portals, price comparison</td>
<td>Traffic, Conversion, Revenue</td>
<td>Field data</td>
<td>SVAR</td>
<td>Fashion, beauty, electronics, home, leisure</td>
<td>Content-integrated channels are most revenue effective</td>
<td>No</td>
</tr>
<tr>
<td>Anderl et al. (2016)</td>
<td>Online</td>
<td>Affiliate, display, email, price comparison, retargeting, SEA, SEO, referral, social media</td>
<td>Conversion</td>
<td>Field data</td>
<td>Markov model</td>
<td>Travel agency, fashion, luggage</td>
<td>Attribution model captures contribution of online channels</td>
<td>-</td>
</tr>
<tr>
<td>This Study</td>
<td>Online</td>
<td>Direct type-in, SEO, SEA, price comparison, referral, affiliate, email, retargeting, social media, App, shopping, other</td>
<td>Conversion, Conversion Value</td>
<td>Field data</td>
<td>Markov model</td>
<td>Travel, technical supplies</td>
<td>High-value and low-value conversions</td>
<td>-</td>
</tr>
</tbody>
</table>
Being among the first to cover multiple online channels in marketing channel attribution research, Shao & Li (2011) apply two algorithmic multi-touch attribution approaches – a bagged logistic regression model and a simple probabilistic model – to a real-world data set from a multichannel advertising campaign. They provide initial indications that last-touch attribution models are particularly inaccurate for measuring the effectiveness of display advertisements (Shao & Li, 2011). This inaccuracy of last-touch attribution and other basic attribution models in providing estimates of marketing channel performance is also addressed and confirmed in later research (de Haan, Wiesel, & Pauwels, 2016; Anderl, Becker, von Wangenheim, & Schumann, 2016). De Haan et al. (2016) found that last-touch attribution suggests online marketing channel budget allocations that yield 10%-12% less revenue than the status quo, whereas their algorithmic attribution model proposes an allocation that would yield a 21% revenue increase over the status quo.

Together with Danaher & Dagger (2013) and Li & Kannan (2014), de Haan et al. (2016) are also among the few that estimate marketing channel performance using multiple dependent variables, i.e. attribution measures. For five product categories, they investigate how nine different marketing channels generate traffic, affect conversion, and contribute to revenue. They find that content-integrated marketing channels are the most effective in generating revenue, followed by content-separated marketing channels and firm-initiated marketing channels. This finding is congruent with prior research indicating that customer-initiated channels are more effective in driving value than firm-initiated channels by factors of 15 till 25 (Sarner & Herschel, 2008; Wiesel, Pauwels, & Arts, 2011).

Extending an approach originally developed in the context of search engine marketing (Archak, Mirrokni, & Muthukrishnan, 2010), a study done by Anderl et al. (2016) was the first and only to apply an attribution framework based on Markovian graph-based data mining techniques as a means to tackle the attribution problem in an online multichannel environment. A third-order Markov model is compared against last-touch and first-touch attribution and clearly outperforms these basic attribution models in predictive performance. It achieves results comparable to those of widely accepted and tested models, making it well suited for analyzing online marketing channel performance (Anderl, Becker, von Wangenheim, & Schumann, 2016).

2.2.2 Attribution Measures throughout Literature

The attribution measure used in multichannel attribution studies differs throughout literature. The measurement standard most debated is conversion attribution, which is generally defined as the assignment of conversion credit over the multiple channels that have reached the customer (Archak, Mirrokni, & Muthukrishnan, 2010). Shao & Li (2011), Xu et al. (2014), Anderl et al. (2015), and Anderl et al. (2016) have all taken conversion as the sole attribution measure in their research. None of them explain why they have taken conversion events as their outcome of interest, although in the case of Shao & Li (2011) this could be because their model could only handle a binary outcome variable. Anderl et al. (2016) do note that companies may also intend to consider value and profits from conversions and that it would be an interesting extension of their research.

Due to the focus of most multichannel attribution studies on conversion events as the attribution measure, Danaher & Dagger addressed this gap in literature by examining multiple marketing channels using revenue and resulting profits as attribution measures. They matched individual-level exposure to marketing channels with purchase data (Danahar & Dagger, 2013). Taking several attribution measures, De Haan, Wiesel & Pauwels (2016) posit that marketing channels may create value in several ways. First, they can generate traffic to the website. Second, they can attract prospects who are more
likely to go deeper into the website and therefore more likely to convert. Finally, marketing channels can influence the amount shoppers spend at checkout. So they identify whether marketing channels differ in generating additional traffic, increasing the likelihood of conversion, and/or enhancing revenue per conversion (de Haan, Wiesel, & Pauwels, 2016).

2.2.3 The Influence of Industry Context

Studies have addressed various types of markets, covering a large range of product categories. It must be noted that attribution results can differ for each type of market and product. For example, Goldfarb & Tucker (2011a) found that in more private categories, e.g. financial or health products, display advertisements that are both obtrusive and targeted have a lower impact on purchases, although personalization in general was found to substantially enhance display advertisement effectiveness. Also, the amount of different channels being used by a customer is found to vary greatly by category (Bhatnagar & Ghose, 2004). Book purchases predominantly involve a single channel, whereas shopping for consumer electronics engages the customer in several channels (Konus, 2010).

The effect of the number of website visits on purchase likelihood also varies by product category. Focusing on new cars, research by Sismeiro and Bucklin (2004) shows that browsing behavior was not predictive of purchase, whereas Moe and Fader (2004) found that more site visits lead to greater purchase likelihood for books and CDs. Knowledge building and hedonic browsing may be higher for new cars than books, so that many users often return to the site multiple times without ever purchasing (Bucklin & Sismeiro, 2009).

Other research indicates that advertising elasticities are higher for new than for established products (Sethuraman, Tellis, & Briesch, 2011) and that expensive products induce the use of price comparison sites (de Haan, Wiesel, & Pauwels, 2016).

Hence, results from prior studies are conditional upon managerial decisions and context, and, may not hold for other industries or companies, selling different products in a different market. This is because the attribution problem is endogenic, i.e. it measures the relative effectiveness of channels in a given setting (Li & Kannan, 2014). Endogeneity issues make it hard to state what findings are industry-specific and which can be generalized. For this reason, scholars specifically call for further research using individual-level path data across several companies and industries in order to make empirical generalizations (Kamakura, Kopalle, & Lehmann, 2014; Li & Kannan, 2014). This can provide guidance to managers and help firms to more accurately measure the impact of their marketing channels in specific contexts (Kohler, Mantrala, Albers, & Kanuri, 2016).

2.2.4 Literature Gaps

Prior research has made significant progress in estimating online marketing channel performance by applying algorithmic multi-touch attribution models. However, some research and managerial questions still remain and thus require further investigation.

One untapped issue is the influence of the attribution measure that is used on marketing channel attribution estimates. Attribution measurement falls in the domain of descriptive modeling, such as there exists no objective truth or evaluation set with which to measure any type of accuracy or loss function (Dalessandro, Perlich, Stitelman, & Provost, 2012). Aside from developing a method for variance estimation of attribution estimates, what is considered “good” attribution may rely by and large on the subjective eyes of the beholder.

Consequently, some studies take conversion events as the attribution measure, while others consider revenues or profits resulting from conversion. However, channel attribution, more than anything, should align the incentive of the company with the incentives of the marketing channel. A common goal of a company running a marketing campaign is to drive as many conversion events at as low a cost as possible. The goal for
marketing channels, however, is to receive as much payment as possible for these conversion events. Thus, in multichannel marketing, the optimal strategy for a given marketing channel might be guided more by how it drives value and less by how well it can influence conversion (Dalessandro, Perlich, Stitelman, & Provost, 2012), making it more seemly to take conversion value as attribution measure than the conversion event itself. If attribution estimates do not alter when selecting a different attribution measure, there is no need to bother. However, none of the studies in the field of multichannel marketing attribution provides indication for this. Even those studies that address several different attribution measures do not consider a potential influence of the attribution measure on marketing channel attribution estimates. Therefore, comparing marketing channel attribution estimates in terms of conversion with estimates in terms of conversion value may be an interesting contribution.

Second, multichannel attribution studies do not investigate to what extent and on which aspects high-value conversions, i.e. those of high transaction value, differ from low-value conversions, i.e. those of low transaction value. Successful companies consciously resist investing on every touchpoint, marketing channel and customer journey. Instead, they identify and spend aggressively only on those interactions they know will have the most impact on revenue growth and profitability (Hogan, Almquist, & Glynn, 2005). In order to identify those conversions that generate the most value, additional knowledge is needed about differences in channel usage, prior buying experience and browsing behavior for these two categories, i.e. high-value and low-value conversions. Such research might provide fruitful insights and may lead to discovery of new research avenues.

2.3 Propositions

In the last section of this chapter propositions are constructed about the research questions in this research.

2.3.1 The influence of Attribution Measures on Attribution Estimates

When conversion is used as attribution measure, channels are credited for enabling conversion events, but when conversion value is used as attribution measure, channels are credited for driving transaction value resulting from those conversions.

In both enabling conversion events as well as driving value CICs have been found to be more effective than FICs (Anderl, Becker, von Wangenheim, & Schumann, 2016; de Haan, Wiesel, & Pauwels, 2016; Sarner & Herschel, 2008). Whereas FICs often reach customers at the wrong time and with a suboptimal message, CICs appear less intrusive and more relevant (Wiesel, Pauwels, & Arts, 2011). This results in a greater performance of CICs overall. Therefore, for both attribution measures, CICs are expected to obtain attribution estimates that are of greater value than those of FICs. Hence, proposition 1:

**P1** For both attribution measures, attribution estimates of CICs are of greater value than attribution estimates of FICs.

Although CICs have shown to be more effective in enabling conversions as well as driving value, their relative performance is expected to decrease when taking conversion value instead of conversion as attribution measure. Accounting for the transaction value of a conversion provides additional information about the purchase and thus enables to relate to customer buyer behavior. Whereas new customers are more likely to initiate contact with the company on their own initiative and respond less to advertising messages
initiated by the firm, returning customers\textsuperscript{11} are more inclined to respond to advertising messages initiated by the firm as they have built up a relationship with the firm through prior purchases. Therefore, compared to new customers, returning customers are expected to more often use FICs. Prior research has shown that returning customers are of greater value to companies than new customers (Monetate, 2015). This indicates that attribution estimates of FICs, as they are expected to be more often used by (valuable) returning customers, therefore become of relative greater value when including transaction value in the calculation of channel attribution estimates. This translates in proposition 2:

\textbf{P2} \textit{Attribution estimates of FICs become greater and attribution estimates of CICs become smaller when using conversion value as attribution measure instead of conversion.}

\subsection*{2.3.2 High-Value Conversions versus Low-Value Conversions}
Elaborating on the transactional value of a conversion, high-value and low-value conversions may differ on other aspects aside the value they generate. A first differentiator is the importance of online channels for driving these conversions. Some online channels may be of greater importance for driving high-value conversions than they are for driving low-value conversions, e.g. because they reduce perceived risk by offering a safe environment for high value conversions or are simply more relevant for high-value conversions. Following this logic, price comparison sites are expected to be of greater relative importance in high-value conversions as they offer customers a platform to evaluate alternatives. Evaluating alternatives may not be necessary for low-value conversions, as the risk involved with such purchases is low. However, price comparison sites become more relevant when a customer is about to make a valuable purchase and wants to be sure of making the right choice.

The greater perceived risk involved with high value purchases may also induce search channels such as SEO and SEA. To increase trust in a potential purchase decision, customers are expected to search for more information about high value purchases.

Finally, email is expected to be of greater importance for high-value conversions than for low-value conversions. When properly customized to customers, e-mails can effectively encourage high value purchases as they offer a great means for companies to engage with their customers and seduce them in making those purchases. For example, companies may use e-mail to offer personalized discounts on valuable products to customers.

In prior research, high value purchases have indeed shown to induce the use of price comparison sites (de Haan, Wiesel, & Pauwels, 2016). Similar results have been found for e-mail (Danahar & Dagger, 2013; Li & Kannan, 2014) and SEA (Wiesel, Pauwels, & Arts, 2011). Hence, combing practical reasoning with prior research findings leads to the following proposition:

\textbf{P3} \textit{SEA, SEO, price comparison sites, and email are of relative greater importance in driving high-value conversions than they are in driving low-value conversions.}

Another aspect on which high-value conversions may differ from low-value conversions is the prior buying experience of the customer that converts. Customers may have gained experience and trust from purchases they have made before. Consequently, returning customers often have a higher level of trust in the company they are buying from than

\textsuperscript{11} Returning customers refer to those customers that have made a purchase at the company they are buying from before
new customers, which leads to a reduction of uncertainty and perceived risk (de Haan, Wiesel, & Pauwels, 2016; Morgan & Hunt, 1994). Research has shown that this increased trust from experience translates in returning customers to be of higher value (Monetate, 2015). In general, new customers are found to be more risk averse (Moe & Fader, 2004), and therefore reluctant to make high value purchases. This threshold may shrink as customers gain familiarity through repeated purchases, thereby making high-value conversions more likely. This translates in the following proposition:

**P4** The share of returning customers is greater in high-value conversions than in low-value conversions.

Finally, more frequent interactions with a firm should enhance customers’ trust in the firm more quickly. Each interaction in the customer journey has its stochastic impact, and as the effects of these interactions accumulate, the greater the amount of trust becomes (Morgan & Hunt, 1994). Following the same logic as with proposition 4, this increased trust and reduced perceived risk and uncertainty make high-value conversions more likely. Also, as the risk with high value purchases is greater, it may take several interactions until the customer is sure enough to make the purchase. This leads to the final proposition:

**P5** The frequency of interactions within customer paths is higher for high-value conversions than for low-value conversions.
CHAPTER 3

Data

3.1 Company Context
This research examines two data sets from two different companies operating in different industries. Data Set 1 is provided by an online travel company and Data Set 2 originates from an online retailer selling technical supplies. Both companies are pure online players, i.e. they only sell their products online and do not have physical stores, so online/offline cross-channel effects can largely be excluded (Anderl, Becker, von Wangenheim, & Schumann, 2016). The focal companies do not wish to disclose their names. Instead, they will be named Company A and Company B, respectively, hereafter.

Company A is a large B2C travel company, offering travel products (e.g., hotels, flights) through its own network of localized websites and telephone call centers. Currently, in evaluating marketing channel effectiveness, Company A applies a simple heuristic; the last paid touch. This means that a conversion is solely attributed to the last paid touch and corresponding channel used by the customer, prior to making a purchase.

Company B is a B2B e-commerce company with web shops live in several European countries. It offers technical supplies for maintenance and reparations of machinery and installations. With a broad assortment of over 500,000 industrial components the company’s online offer is focusing on a one-stop shop proposition primarily directed at smaller and medium-sized business clients. Company B uses free software from Google Analytics to obtain its attribution estimations (Google Analytics Solutions, 2017). Google Analytics offers Company B basic attribution models such as last-touch, first-touch and linear attribution to estimate online channel performance.

3.2 Data Description
The two data sets contain individual-level clickstream data12. Clickstream data record each customer’s Internet activity and thus trace the navigational path the customer takes online (Bucklin & Sismeiro, 2009). Both data sets contain information about the exact sequence of online channels being used and the corresponding purchase amount. Data Set 2, provided by Company B, only contains conversions, i.e. paths that lead to a purchase. Data Set 1, provided by Company A, also contains non-conversions, i.e. paths that did not lead to a purchase. Additionally, for each path in Data Set 1 it is known whether a purchase was made in the last 12 months by the same customer.

For Company B, data collection occurs at the cookie-level, thereby tracking customer journeys ending in a conversion within a timeframe of thirty days from the last exposure13. Limitations of the cookie mechanism are its inability to track multi-device usage and bias due to cookie deletion (Flosi, Fulgoni, & Vollman, 2013; Boland, 2014). Yet cookies remain the standard for multichannel tracking (Tucker, 2012). Company A also uses cookie data, but overlays it with internal device ID’s and 3rd party cross-device tracking. This enables to track customers across multiple devices. It also means that the limitation of technology,

12 To ensure anonymity and General Data Protection Regulation (GDPR) compliance, the collected data is non-PII (Personally Identifiable Information). No personal customer data was collected.
13 The thirty-day window is based on the limitations of cookies as a cookie is set to expire in thirty days.
i.e. the cookie erase, can be overcome by cross-device ID tracking and that the thirty-day timeframe does not apply to customer paths in Data Set 1 provided by Company A.

3.2.1 Sampling Procedure

Data Set 1 has been sampled by Company A itself, as the volumes of data this company gathers are enormous. Based on uniform sampling, Company A took a sample that retained basic statistical characteristics from the original population and could thus be used to represent the original data set. Furthermore, as is common for online travel, the sample distribution was heavily skewed due to low conversion rates. To achieve a more balanced sample non-conversions were deliberately undersampled.

Company B expanded to more online channels throughout the observed period. As this would influence the estimates resulting from multichannel attribution, analysis of Data Set 2 is limited to those customer paths that occurred after all currently used online marketing channels were deployed. Finally, one-touch journeys have been excluded from both data sets. Sophisticated attribution is not required for journeys consisting of just a single touch, as for those cases single-touch attribution models suffice (Anderl, Becker, von Wangenheim, & Schumann, 2016). After the sampling procedures, the data sets can be described by their properties as presented in Table 2.

<table>
<thead>
<tr>
<th>Description</th>
<th>Data Set 1</th>
<th>Data Set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry</td>
<td>Travel</td>
<td>Technical supplies retail</td>
</tr>
<tr>
<td>Market</td>
<td>B2C</td>
<td>B2B</td>
</tr>
<tr>
<td>Region</td>
<td>EMEA(^{14})</td>
<td>EU</td>
</tr>
<tr>
<td>Observation period</td>
<td>365 days</td>
<td>119 days</td>
</tr>
<tr>
<td>Number of different channels</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Number of touchpoints</td>
<td>1,251.468</td>
<td>35.161</td>
</tr>
<tr>
<td>Number of journeys</td>
<td>214,660</td>
<td>8,876</td>
</tr>
<tr>
<td>Interaction frequency (in paths)</td>
<td>5.83</td>
<td>3.96</td>
</tr>
<tr>
<td>Number of conversions</td>
<td>40,644</td>
<td>8,876</td>
</tr>
<tr>
<td>Number of non-conversions</td>
<td>174,016</td>
<td>0</td>
</tr>
<tr>
<td>Journey conversion rate</td>
<td>18.93(^{%})</td>
<td>100(^{%})</td>
</tr>
</tbody>
</table>

3.2.2 Online Marketing Channels

Both companies deploy an extensive range of online marketing channels, distinguishing between ten different online channels. However, across the focal companies, the channels that are used differ slightly. Direct type-in, search engine optimization (SEO), search engine advertising (SEA), price comparison, referral, affiliate and email appear in both data sets. SEA is differentiated on the basis of the keywords used in a search: if the keyword contains the retailer’s brand, the search is branded. If not, the search is unbranded. In addition to these channels, Company A uses social media and a mobile application, whereas Company B uses retargeting and Google shopping. Table 13 in Appendix A provides an overview and description of the online marketing channels that are present in the data. The classification of these online marketing channels along dimensions of contact origin, brand usage, browsing goal, degree of content integration and degree of personalization can be found in Table 14 in Appendix B. Table 3 provides information on the distribution of clicks across channels for both data sets. This also illustrates the variation in frequency of channels being used between the two data sets.

\(^{14}\) Europe, Middle East & Africa
3.2.3 Variables of Interest

Variables of interest for this research are online marketing channel, conversion, conversion value, prior buying experience and interaction frequency. Operational definitions of these variables of interest are given in Table 4 and an overview of the descriptive statistics for some of these variables is provided in Table 5.

Table 3. Channel distribution

<table>
<thead>
<tr>
<th>Type</th>
<th>Channel</th>
<th>Data Set 1 (Company A)</th>
<th>Data Set 2 (Company B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Clicks per journey</td>
<td>Share of touchpoints</td>
</tr>
<tr>
<td>CIC</td>
<td>Direct type-in</td>
<td>.67</td>
<td>12.8%</td>
</tr>
<tr>
<td></td>
<td>SEO</td>
<td>.68</td>
<td>12.8%</td>
</tr>
<tr>
<td></td>
<td>SEA-branded</td>
<td>.22</td>
<td>4.2%</td>
</tr>
<tr>
<td></td>
<td>SEA-unbranded</td>
<td>.68</td>
<td>12.8%</td>
</tr>
<tr>
<td></td>
<td>Price comparison</td>
<td>1.25</td>
<td>23.7%</td>
</tr>
<tr>
<td></td>
<td>Application</td>
<td>.40</td>
<td>7.5%</td>
</tr>
<tr>
<td></td>
<td>Shopping</td>
<td>n/a15</td>
<td>n/a</td>
</tr>
<tr>
<td>FIC</td>
<td>Referral</td>
<td>.33</td>
<td>6.2%</td>
</tr>
<tr>
<td></td>
<td>Affiliate</td>
<td>.23</td>
<td>4.4%</td>
</tr>
<tr>
<td></td>
<td>Email</td>
<td>.70</td>
<td>13.2%</td>
</tr>
<tr>
<td></td>
<td>Retargeting</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Social media</td>
<td>.13</td>
<td>2.4%</td>
</tr>
</tbody>
</table>

Table 4. Operationalization of variables of interest

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online marketing channel</td>
<td>The online marketing channel that has been used by the customer to enable a touchpoint</td>
</tr>
<tr>
<td>Conversion</td>
<td>The event of a purchase (1: Yes, 0: No)</td>
</tr>
<tr>
<td>Conversion value</td>
<td>The transaction value of the conversion</td>
</tr>
<tr>
<td>Prior buying experience</td>
<td>The customer has made a booking in the last 12 months (1: Yes, 0: No)</td>
</tr>
<tr>
<td>Interaction frequency</td>
<td>The frequency of touchpoints, and corresponding online marketing channels in the customer journey</td>
</tr>
</tbody>
</table>

Table 5. Descriptive statistics

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conversion value16</td>
<td>0</td>
<td>705.357,3017</td>
<td>2.267,05</td>
<td>8.862,73</td>
<td>65.94</td>
<td>7912.72</td>
</tr>
<tr>
<td></td>
<td>Past Purchase Interaction frequency</td>
<td>0</td>
<td>1</td>
<td>.30</td>
<td>.46</td>
<td>.87</td>
<td>-1.23</td>
</tr>
<tr>
<td>Data Set</td>
<td>Conversion value</td>
<td>2</td>
<td>778</td>
<td>5.83</td>
<td>7.09</td>
<td>13.34</td>
<td>775.66</td>
</tr>
<tr>
<td></td>
<td>Interaction frequency</td>
<td>2</td>
<td>133</td>
<td>3.96</td>
<td>8.46</td>
<td>6.60</td>
<td>69.77</td>
</tr>
</tbody>
</table>

15 This company does not use a “shopping” channel. Whenever an online marketing channel is not being used, it is specified as not available (n/a)

16 Values are based only on conversions as non-conversions would affect the mean value of purchases

17 Although this value is high, it is not rare. There are over 40 observations of purchases of a value greater than 100.000
CHAPTER 4
Methodology

4.1 Markovian Attribution Model

The attribution model used in this research depicts a graph-based Markovian attribution model. This model was originally developed by Archak et al. (2010) and later extended by Anderl et al. (2016). This research applies the same model as the one used in the research by Anderl et al. (2016). The Markov attribution model attempts to mine frequent substructures in graph data. The graph-based structure of the model reflects the sequential nature of customer journeys. Customer journeys are modeled as chains in first- and higher-order Markov graphs, using a property called removal effect to determine the contribution of states. The order of a Markov model indicates how many previous steps are included in these states. In this research, steps relate to online marketing channels. States represent single online marketing channels in first-order models and represent channel sequences consisting of multiple online channels in higher-order models (Anderl, Becker, von Wangenheim, & Schumann, 2016).

The merits of employing a Markovian attribution model are its flexibility and parsimony. It does not impose a priori constraints on the number of states and transition paths; the total number of states is instead determined empirically through model selection criteria (Chang & Zhang, 2016). Furthermore, it meets the properties of an attribution standard as proposed by Dalessandro et al. (2012).

![Figure 4. Exemplary first-order Markov graph](image)

4.1.1 Markov Graphs

A Markov graph provides a graphical overview of all the customer journeys that are present in the data by capturing co-occurrences of states adjacent to each other (Archak, Mirrokni, & Muthukrishnan, 2010). In a Markov graph, the “START” node represents the start of the customer journey, the “CONVERSION” node represents a conversion, and the “NULL” node is the end state for customer journeys that have not ended in a conversion during the observation period. For modeling reasons, the “CONVERSION” node is also connected to the “NULL” node. Every other graph node represents a certain state. This state can be an online channel, e.g. C1, or a sequence of channels. The edge weights are the probabilities of observing two states consecutively in a customer journey (Archak, Mirrokni, & Muthukrishnan, 2010). Using this graph-based approach allows to represent and analyze customer journeys in an efficient way, as the size of the final graph does not depend upon the number of paths in the data set, but only on the number of states.
Figure 4 shows an exemplary Markov graph based on four customer journeys. The first-order Markov graphs for Data Set 1 is provided below in Figure 5. The first-order Markov graph for Data Set 2 is provided in Figure 10 in Appendix D. This way, in one graphical overview, all customer journeys present in the data set are combined and visualized, providing information regarding the importance, position and interrelationships among online channels. For example, from Figure 5 we can conclude that customer journeys often begin with a price comparison site visit ($w_{ij} = 0.28$) and end in conversion after a mobile application visit ($w_{ij} = 0.13$).

**Figure 5. First-order Markov graph for Data Set 1 (Company A)**

### 4.1.2 First-order Markov Model

A first-order Markov model looks back zero steps. The customer is currently at a certain state, and the probability of going anywhere is based on being at that state. Hence, a first-order Markov model does not incorporate any previous observations, but only the current state. In the first-order Markov model, each state $s_i$ corresponds to one online channel.
The mathematics behind a first-order Markov model is as follows:

\[ w_{ij} = P(X_t = s_j | X_{t-1} = s_i), 0 \leq w_{ij} \leq 1, \sum_{j=1}^{N} w_{ij} = 1 \forall i \]

With:
- Transition Probability \((w_{ij})\) = The probability that state \(s_i\) is followed by state \(s_j\);
- Transition Probability \((w_{ij})\) is no less than 0 and no greater than 1;
- The sum of all Transition Probabilities equals 1.

### 4.1.3 Higher-order Markov Model

Whereas a first-order Markov model only takes the present state, i.e. online channel, into account, higher-order Markov models can look back one or more steps in the sequence. In the first-order model, each state \(s_i\) corresponded to one channel. In a higher order model, this state \(s_i\) now represents a tuple, i.e. sequence of online channels. The higher the order, the greater this sequence is and the more previous steps are taken into account when calculating the transition probability for the next step (Anderl, Becker, von Wangenheim, & Schumann, 2016). Therefore, higher-order models are more accurate in making attribution estimations. However, the number of independent parameters and the complexity of Markov models rises exponentially as the order increases. At some point, this complexity becomes too great to make attribution estimations with real-world data sets (Anderl, Becker, von Wangenheim, & Schumann, 2016). For higher-order models, in which the present depends on the last \(k\) observations, Transition Probabilities can be defined as follows:

\[ w_{ij} = P(X_t = s_t | X_{t-1} = s_{t-1}, X_{t-2} = s_{t-2}, \ldots, X_{t-k} = s_{t-k}) \]

### 4.1.4 The Removal Effect

The performance of online marketing channels can be represented by the so-called removal effect. The removal effect is the change in probability of converting or total conversion value when a state is completely removed from the graph. The more the probability of converting or conversion value decreases when removing a state, the higher the removal effect of that state. Therefore, a high removal effect suggests high performance, while a low removal effect indicates that the state, and corresponding online channel(s), might not be very important for a customer to reach conversion or drive value.

As mentioned before, in higher-order Markov models, states do not represent single online channels but channel sequences instead. Higher-order models therefore allow to calculate the removal effects of channel sequences. In such models, the removal effect for a single online channel is now reflected by the mean removal effect of all states having that specific online channel as the last channel in their sequence.

In this research, calculations of removal effects are carried out in R, a free software tool for statistical computing and graphics (R Foundation, 2017), using the "Channel Attribution” R package written by Davide Altomare and David Loris (Altomare & Loris, 2016). The removal effects can be calculated according to the attribution measure that is used for doing marketing channel attribution. In Appendix E, removal effects are calculated in terms of conversion and conversion value, based on the exemplary first-order Markov graph in Figure 4. This is done to illustrate how removal effects are calculated in R and affected by the attribution measure.
CHAPTER 5

Results

5.1 Model Fit

Markov models of order one till five\(^{18}\) are evaluated on two important criteria of model fitness, i.e. predictive accuracy and robustness. Testing the model on these criteria ensures scientific rigor and stable results. First of all, although attribution takes a retrospective view, attribution models should be able to correctly predict conversion events and conversion value (Shao & Li, 2011). Second, attribution models should deliver stable and reproducible results when the model is run multiple times. Both qualities are evaluated and compared.

5.1.1 Predictive Accuracy

Two metrics of predictive accuracy are used to evaluate the Markov models on their predictive performance. First, receiver operating characteristics (ROC) graphs are constructed. ROC has become a commonly applied metric in machine learning and data mining research (Fawcett, 2006). ROC graphs visualize, organize and select classifiers based on performance. The curve of the ROC graph is used in evaluating and comparing algorithms. In this case, that is the algorithm behind Markov models of order one till five.

Data is split into a “training” and “test” sample. Markov analysis is done on the training sample, which results in predictions about channels and channels sequences resulting in conversion or not. These predictions are then applied to the test sample and compared against the actual classification for the paths in the test set. The ROC curve decouples classification performance from class distributions and misclassification costs (Bradley, 1997). Basically, it plots the true positive rate (positives correctly classified / total positives) against the false positive rate (negatives incorrectly classified / total negatives) at various threshold settings. The optimal point on the ROC curve is \((FPR, TPR) = (0, 1)\). This means no false positives and all true positives. The closer the curve is to that point, the better the model.

ROC is a powerful metric as the ROC graph is invariant against skewed class distribution of the applied data set. Hence, a data set featuring 60% positive labels, e.g. conversion, will yield the same ROC graph as a data set featuring 45% positive labels (Vogler, 2015). That makes it a particularly useful metric for this research as the division of conversions and non-conversions is not 50/50 in the data sets. However, although ROC can handle a skewed class distribution of conversion, they are intended for binary variables in which the result either occurs (conversion) or does not (non-conversion) (Linden, 2005). This means that the predictive accuracy of Markov models cannot be evaluated for Data Set 2 as it only contains conversions. Furthermore, it does not make sense to test predictive accuracy when the attribution measure that is being predicted is always of value “1”.

Predictive accuracy can also not be evaluated with conversion value as attribution measure as it is a continuous variable and not a binary variable. It is possible to look for

---

\(^{18}\) The number of independent parameters and the complexity of Markov models rises exponentially as the order increases. At some point, this complexity becomes too great to make attribution estimations with real-world data sets. Therefore, evaluation of predictive accuracy and robustness is limited to Markov models of order one till five.
cutoff points using ROC analysis, but you would have to dichotomize the variable into two broad categories (e.g., high and low value). Some investigators dichotomize variables in order to run a ROC curve but this practice is frowned upon. The option of dichotomizing the conversion value variable has been investigated in this research, but this did not lead to sensible ROC curves. Based on Data Set 1, Figure 6 contains the ROC curves of Markov models ranging from order one till five with conversion as the attribution measure.

A second metric that is used to measure and compare the predictive accuracy of Markov models is the AUC value. This AUC value relates to the area under the ROC curve. The AUC value is specifically useful as it reduces ROC performance to a single scalar value. As stated before, the optimal point on the ROC curve is \((FPR, TPR) = (0, 1)\). Hence, the closer the AUC value is to 1, the better is the predictive ability of the model.

To calculate the AUC values for Markov order one till five, a 10-fold cross-validation procedure is used. Cross-validation ensures that results are not dependent on one single split of data. Instead, AUC values are measured across 10 different samples. This is done using a within-sample as well as an out-of-sample sampling procedure. Table 6 lists the mean AUC values with their standard deviations in brackets for Data Set 1.

**Table 6. AUC values for Data Set 1**

<table>
<thead>
<tr>
<th>Sample</th>
<th>First order</th>
<th>Second order</th>
<th>Third order</th>
<th>Fourth order</th>
<th>Fifth order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within sample</td>
<td>.7616 (.0105)</td>
<td>.7869 (.0087)</td>
<td>.8173 (.0087)</td>
<td>.8574 (.0122)</td>
<td>.9043 (.0081)</td>
</tr>
<tr>
<td>Out-of-sample</td>
<td>.7659 (.0068)</td>
<td>.7842 (.0077)</td>
<td>.8009 (.0070)</td>
<td>.8090 (.0092)</td>
<td>.8207 (.0071)</td>
</tr>
</tbody>
</table>

In line with findings from prior research done by Anderl et al. (2016) and as reflected by the ROC graphs in Figure 6 and the AUC values in Table 6, the predictive accuracy of a Markov model increases with rising Markov order. Higher order Markov models get closer to the optimal point on the ROC curve \((0, 1)\) and the AUC value increases towards its optimal value \((1)\). Whereas with a within-sample procedure AUC values increase linearly with rising Markov order, the out-of-sample procedure leads to a different observation; the relative increase in AUC values lessens with rising Markov order.
5.1.2 Robustness

Robustness concerns the ability of a model to deliver stable and reproducible results if the model is run multiple times, and is indispensable for sustainable attribution results (Anderl, Becker, von Wangenheim, & Schumann, 2016). A first indication of robustness is provided by the standard deviations of the AUC values noted in brackets in Table 6 as it tests for robustness across all cross-validation repetitions. Standard deviations are small (<.02) and imply low overall variation without systematic differences between models.

More important however, is that the attribution measure used for attribution modeling provides stable attribution estimates. Only then it can offer a reliable basis for managerial decisions, such as shifts in budget allocations across marketing channels. Therefore, the robustness of the removal effect ($s_j$) is tested. This robustness is captured by computing the average standard deviation of the removal effect ($s_j$) for each online marketing channel across ten cross-validation repetitions. The stability of the removal effects is expressed in the average standard deviation of the removal effects as percentage of the average removal effects across all online marketing channels. Table 7 provides this information. Lower percentages correspond to more robust estimations of removal effects.

Table 7. Robustness of removal effect: average standard deviation as % of average removal effect

<table>
<thead>
<tr>
<th>Data set</th>
<th>Measure</th>
<th>Model Order 1</th>
<th>Model Order 2</th>
<th>Model Order 3</th>
<th>Model Order 4</th>
<th>Model Order 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Set 1</td>
<td>Conversion</td>
<td>.31%</td>
<td>.32%</td>
<td>.35%</td>
<td>.34%</td>
<td>.34%</td>
</tr>
<tr>
<td></td>
<td>Conversion Value</td>
<td>.80%</td>
<td>.79%</td>
<td>1.04%</td>
<td>1.01%</td>
<td>.92%</td>
</tr>
<tr>
<td>Data Set 2</td>
<td>Conversion</td>
<td>2.02%</td>
<td>1.92%</td>
<td>1.95%</td>
<td>1.96%</td>
<td>1.96%</td>
</tr>
<tr>
<td></td>
<td>Conversion Value</td>
<td>2.09%</td>
<td>2.03%</td>
<td>2.07%</td>
<td>2.20%</td>
<td>2.19%</td>
</tr>
</tbody>
</table>

The number of independent parameters and the complexity of Markov models rises exponentially as Markov order increases. As a result, standard deviations of removal effects increase slightly with rising Markov order, indicating that estimates of removal effects become less stable and robust.

It should be noted that in prior research this increase in average standard deviations was found to be much larger. Average standard deviations increased with a factor 5 from Markov order one till four (Anderl, Becker, von Wangenheim, & Schumann, 2016). This difference most likely is the result from using the “Channel Attribution” R package in the calculation of removal effects (Altomare & Loris, 2016). Removal effects for states cannot be provided by this R package. Instead, removal effects for states are aggregated per channel. Hence, the effect of Markov order on the robustness of attribution measure estimates is reflected differently in this research. However, both practices are considered appropriate in determining robustness of the attribution measure estimates.

5.1.3 Determining Markov Order

Based on Figure 6 and Table 6, Markov models of order three till five are recommended for obtaining channel attribution estimations as they show the greatest predictive performance. However, based on Table 7, for both data sets and attribution measures, a second order Markov model would be recommended as it provides the most robust removal effects. These findings suggest a trade-off between predictive accuracy and robustness.

Combining the evaluation results with practical considerations and recommendations from prior research, a fifth order Markov model will be used for the attribution analyses of Data Set 1. This model has the greatest predictive accuracy and was also found to be more
robust than Markov models of order three and four. For attribution analyses of Data Set 2, a third order model will be used. Predictive accuracy of the Markov models cannot be evaluated for Data Set 2 as it only contains conversions. However, findings from prior research suggest that third order models balance predictive performance, robustness, and algorithmic efficiency. Also, from Table 7 can be concluded that a third order Markov model provides more robust removal effects for Data Set 2 than fourth and fifth order models.

The frequency of interactions in customer paths supports the choice for a fifth order Markov model for Data Set 1 and a third order Markov model for Data Set 2. On average, paths in Data Set 1 consist of 5.83 interactions, whereas paths in Data Set 2 consist of 3.96 interactions. Hence, it makes sense to choose for a fifth order model when analyzing customer paths in Data Set 1, as it includes more previous steps in the analysis.

5.2 Comparing the Markov Model

In order to understand how different models result in different channel attribution estimates, a third order Markov Model is compared against some commonly applied basic attribution modes, i.e. first-touch, last-touch, and linear attribution. Figure 7 displays the attribution estimates for Data Set 1 with conversion as attribution measure. The online channels are placed on the x-axis and the number of conversions they generate on the y-axis. Attribution estimates for Data Set 2 with conversion as attribution measure are provided in Figure 11 in Appendix F. Comparisons of attribution models on conversion value are provided in Figure 12 and Figure 13 in Appendix F.

![Figure 7. Attribution estimates in conversion using different attribution models (Data Set 1)
First of all, significant differences can be observed when comparing the results of the Markov model to those of the basic attribution models. In line with Anderl et al. (2016) the Markov model levels the channel contributions’ amplitudes, distributing the contribution more evenly across channels. The results also support prior findings in that last-touch and first-touch attribution underestimate the contribution of FICs such as email, referral and social media (Li & Kannan, 2014). Direct type-ins get more credit with last-touch attribution, whereas price comparison receives less credit. Furthermore, the Markov model consistently assigns more credit to SEO, all in line with prior research (Anderl, Becker, von Wangenheim, & Schumann, 2016; Li & Kannan, 2014).

These findings also relate to linear attribution. However, a linear model distributes the contribution more evenly across channels, making under- and overestimation of contribution less severe. Findings are different for email, a FIC that is overestimated by linear attribution. For emails, stored in the inbox, can be repeatedly accessed and may motivate subsequent visits through other channels. Single-touch models are not able to reflect these effects appropriately, but linear attribution can.

5.3 The influence of Attribution Measures on Attribution Estimates

To investigate the influence of the attribution measure on attribution estimates, for both data sets, channel attribution estimates in terms of conversion are compared to those in terms of conversion value. Channel attribution estimates are represented by their removal effects and are expressed in percentages of the total removal effect for all channels. Results are provided in Table 8. For each data set, the first two columns contain the removal effects for channels using conversion and conversion value, respectively, as attribution measures. The third column notes the difference in attribution estimates. Hence, this column reflects the influence of the attribution measure, i.e. conversion or conversion value, on channel attribution estimates. When a channel is attributed more credit using conversion value instead of conversion, this difference is noted in green. When it is attributed less credit using conversion value, the difference is noted in red. The same has been done on an aggregated level with attribution estimates per channels type, i.e. CICs and FICs. These results are summarized and captured in Table 9.

Table 8. Removal effects (%) taking conversion and conversion value as attribution measure

<table>
<thead>
<tr>
<th>Type</th>
<th>Channel</th>
<th>Data Set 1 (Company A)</th>
<th></th>
<th>Data Set 2 (Company B)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Conversion</td>
<td>C. Value</td>
<td>Dif.</td>
<td>Conversion</td>
</tr>
<tr>
<td>CIC</td>
<td>Direct type-in</td>
<td>19.82%</td>
<td>18.50%</td>
<td>-1.32%</td>
<td>23.19%</td>
</tr>
<tr>
<td></td>
<td>SEO</td>
<td>8.70%</td>
<td>8.69%</td>
<td>-.01%</td>
<td>6.38%</td>
</tr>
<tr>
<td></td>
<td>SEA-branded</td>
<td>7.76%</td>
<td>7.95%</td>
<td>+.19%</td>
<td>8.77%</td>
</tr>
<tr>
<td></td>
<td>SEA-unbranded</td>
<td>7.99%</td>
<td>8.17%</td>
<td>+.18%</td>
<td>18.65%</td>
</tr>
<tr>
<td></td>
<td>Price comparison</td>
<td>9.09%</td>
<td>10.24%</td>
<td>+1.15%</td>
<td>.80%</td>
</tr>
<tr>
<td></td>
<td>Application</td>
<td>11.27%</td>
<td>10.71%</td>
<td>-.56%</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Shopping</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>30.5%</td>
</tr>
<tr>
<td>FIC</td>
<td>Referral</td>
<td>9.63%</td>
<td>9.59%</td>
<td>-.04%</td>
<td>.94%</td>
</tr>
<tr>
<td></td>
<td>Affiliate</td>
<td>8.65%</td>
<td>8.48%</td>
<td>-.17%</td>
<td>3.04%</td>
</tr>
<tr>
<td></td>
<td>Email</td>
<td>14.24%</td>
<td>14.74%</td>
<td>+.50%</td>
<td>2.44%</td>
</tr>
<tr>
<td></td>
<td>Retargeting</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>5.28%</td>
</tr>
<tr>
<td></td>
<td>Social media</td>
<td>2.84%</td>
<td>2.92%</td>
<td>+.08%</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Average difference  .42%  1.09%
Findings confirm that in both enabling conversion events as well as driving conversion value CICs are found to be more effective than FICs. From the aggregated removal effects per channel type in Table 9 we can conclude that, in both data sets and independent of attribution measure, averaged attribution estimates of CICs are of greater value than averaged attribution estimates of FICs. Hence, proposition 1 is supported.

Additionally, from Table 8 we can conclude that attribution measures influence attribution estimates. The average difference in attribution estimates between using conversion value or conversion is .42% for Data Set 1 and 1.09% for Data Set 2. In addition, the aggregated removal effects per channel type in Table 9 show that, in both data sets, attribution estimates of FICs become greater when using conversion value as attribution measure instead of conversion, whereas attribution estimates of CICs become smaller. This is indicted by the negative red values for CICs and positive green values for FICs. Hence, results provide support for proposition 2.

Further, results indicate that there are significant differences across both data sets. Especially in Data Set 2, CICs are of relative great importance. Removal effects of CICs are on average ~5 times greater than removal effects of FICs. In Data Set 1, CICs are on average only ~1.2 times greater than removal effects of FICs. In addition, the attribution measure appears to be of greater influence on attribution estimates for Data Set 2 than on attribution estimates for Data Set 1. Whereas the average difference in attribution estimates between using conversion value or conversion is .42% for Data Set 1, this average difference is 1.09% for Data Set 2. Also on individual channel level, attribution measures affect channels differently for Data Set 1 than for Data Set 2. The most prevailing finding is the opposite direction of the differentia effect for direct type-ins. Whereas for Data Set 1 attribution estimates of direct type-ins decrease in terms of conversion value, estimates of direct type-ins increase in terms of conversion value for Data Set 2. Such findings indicate there is a significant influence of industry context on attribution estimates.

### 5.4 High-Value Conversions versus Low-Value Conversions

To better understand the distribution of conversion value for purchases, a histogram is provided for both data sets. Mean and median values for conversion value are included as well. Figure 8 and Figure 9 contain the histograms for Data Set 1 and Data Set 2, respectively. Both data set face a skewed distribution with a relatively great amount of conversions of low value and few conversions of high value. This is especially the case for Data Set 1. The median is the value that separates the higher half of the data from the lower half. As is common practice in simplifying the presentation of results (DeCoster, Gallucci, & Iselin, 2011), the median is used to categorize conversions into high-value and low-value conversions. For Data Set 1, those conversions below a value of 726.86 are considered low-value conversions. Those conversions above the median are considered to be of high value. For Data Set 2, conversions below a value of 59.03 are considered low-value conversions and those above that value are considered to be high-value conversions.
5.4.1 Marketing Channel Usage

To investigate whether some online channels are of relative greater importance for high-value conversions than they are for low-value conversions the removal effects of online channels for both categories have been estimated. Removal effects are expressed in percentages of the total removal effect for all channels. Table 10 contains the removal effects using conversion as attribution measure, whereas Table 11 contains those removal effects using conversion value as attribution measure.
### Table 10. Differences in removal effects (in %) between high- and low-value conversions (conversion)

<table>
<thead>
<tr>
<th>Type</th>
<th>Channel</th>
<th>Data Set 1 (Company A)</th>
<th>Data Set 2 (Company B)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low-Value</td>
<td>High-Value</td>
<td>Dif.</td>
</tr>
<tr>
<td>CIC</td>
<td>Direct type-in</td>
<td>21.55%</td>
<td>18.57%</td>
<td>-2.98%</td>
</tr>
<tr>
<td>SEO</td>
<td></td>
<td>8.97%</td>
<td>8.57%</td>
<td>-0.40%</td>
</tr>
<tr>
<td>SEA-branded</td>
<td></td>
<td>7.81%</td>
<td>7.76%</td>
<td>-.05%</td>
</tr>
<tr>
<td>SEA-unbranded</td>
<td></td>
<td>8.03%</td>
<td>7.98%</td>
<td>-.05%</td>
</tr>
<tr>
<td>Price comparison</td>
<td></td>
<td>8.31%</td>
<td>9.60%</td>
<td>+1.29%</td>
</tr>
<tr>
<td>Application</td>
<td></td>
<td>11.38%</td>
<td>11.26%</td>
<td>-.12%</td>
</tr>
<tr>
<td>Shopping</td>
<td></td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>FIC</td>
<td>Referral</td>
<td>9.59%</td>
<td>9.62%</td>
<td>+.03%</td>
</tr>
<tr>
<td>Affiliate</td>
<td></td>
<td>8.46%</td>
<td>8.77%</td>
<td>+.31%</td>
</tr>
<tr>
<td>Email</td>
<td></td>
<td>13.57%</td>
<td>14.68%</td>
<td>+1.11%</td>
</tr>
<tr>
<td>Retargeting</td>
<td></td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Social media</td>
<td></td>
<td>2.32%</td>
<td>3.20%</td>
<td>+.88%</td>
</tr>
<tr>
<td>Average difference</td>
<td></td>
<td>.72%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average CICs</td>
<td></td>
<td>-2.32%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average FICs</td>
<td></td>
<td>+2.32%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For each data set, the first two columns contain the removal effects for channels in low-value conversions and high-value conversions, respectively. The third column notes the difference in attribution estimates between the two. Hence, this column reflects the change in importance of online channels when comparing low-value conversions to high-value conversions. Those online channels with positive values given in green are of relative greater importance in driving high-value conversions than they are in driving low-value conversions. The opposite holds for those channels with negative values given in red. These channels are of relative smaller importance in driving high-value conversions than they are in driving low-value conversions.

### Table 11. Differences in removal effects (in %) between high- and low-value conversions (conv. value)

<table>
<thead>
<tr>
<th>Type</th>
<th>Channel</th>
<th>Data Set 1 (Company A)</th>
<th>Data Set 2 (Company B)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low-Value</td>
<td>High-Value</td>
<td>Dif.</td>
</tr>
<tr>
<td>CIC</td>
<td>Direct type-in</td>
<td>21.52%</td>
<td>17.89%</td>
<td>-3.63%</td>
</tr>
<tr>
<td>SEO</td>
<td></td>
<td>8.68%</td>
<td>8.82%</td>
<td>+.14%</td>
</tr>
<tr>
<td>SEA-branded</td>
<td></td>
<td>7.81%</td>
<td>7.98%</td>
<td>+.17%</td>
</tr>
<tr>
<td>SEA-unbranded</td>
<td></td>
<td>7.74%</td>
<td>8.34%</td>
<td>+.60%</td>
</tr>
<tr>
<td>Price comparison</td>
<td></td>
<td>8.25%</td>
<td>10.47%</td>
<td>+2.22%</td>
</tr>
<tr>
<td>Application</td>
<td></td>
<td>11.69%</td>
<td>10.69%</td>
<td>-1.00%</td>
</tr>
<tr>
<td>Shopping</td>
<td></td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>FIC</td>
<td>Referral</td>
<td>9.50%</td>
<td>9.63%</td>
<td>+.13%</td>
</tr>
<tr>
<td>Affiliate</td>
<td></td>
<td>8.73%</td>
<td>8.35%</td>
<td>-0.38%</td>
</tr>
<tr>
<td>Email</td>
<td></td>
<td>13.65%</td>
<td>14.71%</td>
<td>+1.06%</td>
</tr>
<tr>
<td>Retargeting</td>
<td></td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Social media</td>
<td></td>
<td>2.43%</td>
<td>3.10%</td>
<td>+.67%</td>
</tr>
<tr>
<td>Average difference</td>
<td></td>
<td>1.00%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average CICs</td>
<td></td>
<td>-1.49%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average FICs</td>
<td></td>
<td>+1.49%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Depending on the attribution measure, the average difference in importance of online channels is .72% or 1.00% in Data Set 1 and 1.37% or 1.43% in Data Set 2. This shows that there are significant differences in the importance of online channels for low-value conversions versus high-value conversions. Across both data sets and attribution measures, in general, FICs are found to be of greater importance in driving high-value conversions than they are in driving low-value conversions. Consequently, the importance of CICs in driving high-value conversions is found to be smaller compared to their importance in driving low-value conversions.

As expected and in line with proposition 3, price comparison sites and e-mail are consistently found to be of greater importance in driving high-value conversions than in driving low-value conversions. These channels are also the only two out of eight online channels used by both companies that show similar results across both data sets and attribution measures. Hence, again, results indicate that findings are highly dependent on industry context. For example, direct type-ins perform significantly worse in driving high-value conversions than in driving low-value conversions for travel products offered in a B2C market. However, for technical supplies in a B2B market, direct type-ins perform better in driving high-value conversions.

The greater perceived risk involved with buying products of high value was also expected to induce the use of search channels, making them more important in driving high-value conversions than in driving low-value conversions. However, this is not reflected in the results in Table 10 and Table 11. The overall greater importance of SEA-branded and overall smaller importance of SEA-unbranded in high-value conversions does indicate a mediating effect of keyword usage. Apparently, SEA based on branded keywords is more important for driving high-value conversions, whereas SEA based on unbranded keywords is more important for driving low-value conversions. Such differences between SEA-branded and SEA-unbranded appear to be greater for technical supplies offered in the B2B market than for travel products offered in the B2C market.

5.4.2 Prior Buying Experience

Another aspect on which high-value conversions are expected to differ from low-value conversions is the prior buying experience of a customer. Based on risk aversion theory, high-value conversions are more likely to belong to customers that are more experienced, i.e. returning customers, resulting in a greater share of returning customers in high-value conversions when compared to the share of returning customers in low-value conversions. As captured in Table 12, results from this study show that the share of returning customers\(^\text{19}\) is indeed greater for high-value conversions than for low-value conversions. In high-value conversions, 77% of the purchases are made by returning customers, whereas in low-value conversions 74% of the purchases are made by returning customers, thereby indicating a slight difference between low-value and high-value conversions. Hence, proposition 4 is supported. This finding can only be based on Data Set 1 as Data Set 2 provides no information about customer experience.

| Table 12. Returning customers (%) and amount of interactions for low- and high-value conversions |
|-------------------------------------------------|------------------|------------------|------------------|------------------|
| Aspect                                          | Data Set 1 (Company A) | Data Set 2 (Company B) |
|                                                 | High-Value | Low-Value | High-Value | Low-Value |
| Returning Customers (%)                         | 77%        | 74%       | n/a        | n/a        |
| New customers (%)                               | 23%        | 26%       | n/a        | n/a        |
| Frequency of interactions                       | 11.87 (12.22) | 4.55 (3.02) | 8.68 (9.53) | 7.31 (7.04) |

\(^{19}\) In this study, returning customers are those customers that have made a prior purchase within the last 12 months.
5.4.3 Frequency of interactions

A final aspect on which high-value conversions are expected to differ from low-value conversions is the frequency of interactions in the related paths. In customer paths, customers interact with the company via touchpoints that are enabled by online channels. As the risk involved in making high value purchases is higher, it should take more interactions until the customer is sure enough to make such purchases. Therefore, the frequency of interactions is expected to be greater for high-value conversions than for low-value conversions. The results (standard deviations in brackets), as summarized in Table 12, indeed show that the frequency of interactions is significantly greater for high-value conversions than for low-value conversions. This finding is consistent throughout both data sets, although a greater difference in frequency of interactions between high- and low-value conversions can be observed for Data Set 1. Thus, proposition 5 is supported.
CHAPTER 6
Discussion

For this discussion, two research questions are addressed that motivated this study.

6.1 The Influence of Attribution Measures on Attribution Estimates

First, the influence of the attribution measure on online channel attribution estimates is investigated, thereby providing answers to the first research question:

To what extent are online channel attribution estimates affected by the attribution measure?

Results from this study indicate that online channel attribution estimates are affected by the attribution measure. Depending on the company that is being studied, attribution estimates differ between .42% and 1.09% across the attribution measures used in this research, i.e. conversion and conversion value.

As expected, for both companies, attribution estimates for FICs are found to increase when taking conversion value as attribution measure, whereas attribution estimates for CICs are found to decrease. This finding can be explained by returning customers being more inclined to respond to advertising messages initiated by the firm as they have built up a relationship with the firm through previous purchases. In prior research, this relationship has shown to translate in returning customers to be of higher value (Monetate, 2015; Ansari, Mela, & Neslin, 2008). Consequently, attribution estimates of FICs, as they are more often used by (valuable) returning customers, become of relative greater value when including transaction value in the calculation of channel attribution estimates.

Compared to Company A, the attribution measure appears to be of greater influence on attribution estimates for Company B. Such differences are likely due to different industry contexts. Whereas company A is a B2C company, Company B is a B2B company. Actions by consumers are often driven by emotions or aspirations, but businesses act more rational. Consequently, businesses are more likely to proactively identify a need for a product and consumers tend to be influence by advertising and marketing techniques (Scotter, 2018). In general, this translates in the greater importance of CICs for company B. However, this is different for returning business clients. A business client that frequently buys at the focal firm is not likely to use price comparison sites, search for information online or use other CICs. Instead, he reacts on FICs, and most prominently, he directly types in the URL of the company’s website to place an order. This also explains why, for company B, attribution estimates of direct type-ins, as opposed to most CICs, increases in value.

6.2 High-Value Conversions versus Low-Value Conversions

Second, conversions have been split on the basis of transaction value, exploring channel usage, prior buying experience and frequency of interactions for low-value conversions and high-value conversions. In this, the following research question is addressed:

To what extent and on which aspects do high-value conversions differ from low-value conversions?
High-value conversions differ from low-value conversions in terms of channel usage, customers’ prior buying experience, and in frequency of interactions. Hence, from the results in this study can be concluded that accounting for value makes sense as analyses become more sensitive to such differences.

6.2.1 Marketing Channel Usage

Results from this research show that there are significant differences in the importance of online channels for low-value conversions versus high-value conversions. As expected and in line with indications from prior research (de Haan, Wiesel, & Pauwels, 2016; Li & Kannan, 2014; Danahar & Dagger, 2013), price comparison sites and e-mail are consistently found to be of greater importance in driving high-value conversions than in driving low-value conversions.

Price comparison sites offer customers a platform to evaluate alternatives. Such platforms may not be necessary for low-value conversions, as the risk involved with making low value purchases is little. However, price comparison sites become more relevant when a customer is considering making a high value purchase and wants to be sure of making the right choice, therefore resulting in a relative greater importance in driving high-value conversions than in driving low-value conversions.

When properly customized to customers, e-mails can effectively encourage high value purchases as they offer a great means for companies to engage with their customers and seduce them in making those purchases. It must be noted that firms must be cautious while sending as repeated e-mails can quickly be considered spam by customers and thereby backfire (Batra & Keller, 2016).

The greater perceived risk involved with high value purchases was also thought to induce search channels such as SEO and SEA. To increase trust in a potential purchase decision, customers were expected to search for more information regarding high value purchases, making them more important in driving high-value conversions than in driving low-value conversions. Although prior research shows that SEA increases conversion value and resulting revenues (Wiesel, Pauwels, & Arts, 2011), results from this study are inconclusive. For technical supplies offered in a B2B market, results do show a significantly greater importance of SEA-branded and overall smaller importance of SEA-unbranded in driving high-value conversions. This indicates a mediating effect of the keywords used in a search. As high value products are more often related to a specific brand, it may be that especially searches in which branded key-words are used will appear more often in high-value conversions.

Again, results indicate that findings are highly dependent on industry context. Across both data sets and attribution measures, only two out of eight online channels that are used by both companies show similar results.

Finally, attribution estimates of direct type-ins for Company A and attribution estimates of the shopping channel for Company B drastically deteriorate when focusing on high-value conversions instead of low-value conversions. What causes these channels to be considered so much less important for driving high-value conversions is unknown.

6.2.2 Prior Buying Experience

This research finds that the share of returning customers is slightly greater in high-value conversions than in low-value conversions. Whereas new customers are found to be more risk averse (Moe & Fader, 2004) and therefore reluctant to make high value purchases, returning customers have gained experience from prior purchases, resulting in a higher level of trust in the company they are buying from and a reduction of uncertainty
and perceived risk (de Haan, Wiesel, & Pauwels, 2016; Morgan & Hunt, 1994). Research has shown that this increased trust from experience translates in returning customers to be of higher value (Monetate, 2015). However, prior studies find that returning customers are mainly of higher value because of increased purchase incidence with no observed increase in purchase size (Ansari, Mela, & Neslin, 2008). This study provides indications that they do, as the share of returning customers is slightly greater in high-value conversions than in low-value conversions.

Unfortunately, these findings can only be based on Data Set 1 as Data Set 2 provides no information about customer experience. Therefore, these findings may be company-specific and incidental.

6.2.3 Frequency of Interactions

In this research, frequency of interactions is found to be significantly greater for high-value conversions than for low-value conversions. This may be a result of the greater perceived risk that is involved in making purchases of high value, withholding risk averse customers of making such purchases after only having a few interactions. However, as each interaction in the customer journey has its own stochastic impact and the effects of interactions accumulate, the greater the amount of trust becomes (Morgan & Hunt, 1994). In return this increased trust reduces perceived risk and uncertainty and makes high value purchases more likely.

This finding is consistent throughout both data sets, but a greater difference in frequency of interactions between high- and low-value conversions can be observed for the online travel company in this research. In general, for this company, paths also consist of more interactions than paths of the technical supplies company. This may be due to the type of market these companies are active in. Whereas the travel company focuses on consumers, the technical supplies company focuses on business clients. This may affect browsing behavior as businesses behave differently in the buying process than consumers.

This may also have caused the greater difference in frequency of interactions between high-value conversions and low-value conversions for the B2C travel company than the B2B technical supplies company. Consumers have to invest their own money in buying products from the travel company, but business clients invest with capital that is not their own to buy technical supplies from Company B. When buying products where own personal belongings are at stake, customers may feel more hesitant in making high value purchases and therefore need relatively more interactions to feel comfortable with a purchase decision. Purchase value may be of less concern when these products are bought with company capital.
CHAPTER 7

Conclusion

This conclusion contains managerial implications by spelling out how findings resulting from this research can inform decisions and activities, and potentially change current practices and viewpoints. It concludes with limitations and future research suggestions.

7.1 Managerial Implications

First, this research provides novel insights into the impact of the dependent variable, i.e. the influence of the attribution measure, on attribution estimates. Although this study finds small differences in attribution estimates resulting from using conversion or conversion value as attribution measure, implying a minor influence of the attribution measure on attribution estimates, such differences may still result in suboptimal budget allocations across channels that can be of major impact considering companies’ large marketing budgets. Unfortunately, in attribution measurement, there exists no objective truth or evaluation set with which to measure any type of accuracy or loss function (Dalessandro, Perlich, Stitelman, & Provost, 2012). Hence, one cannot determine what attribution measure is “best”. However, channel attribution, more than anything, should align the incentive of the company with the choice for a certain attribution measure. The companies in this research are for-profit organizations, making it more seemly to take conversion value as attribution measure than the conversion event itself. Results from this study may at least address that serious thought should be given to the choice for an attribution measure.

Second, insights resulting from this study can guide companies in identifying and driving those conversions that are most likely to be of high value by revealing that high-value conversions differ from low-value conversions on the aspects of channel usage, customers’ prior buying experience and browsing behavior. Price comparison sites and e-mail are consistently found to induce high-value conversions, which implies that companies should invest more heavily in those specific channels to increase value of conversion. For company B, this also holds for SEA-branded. Furthermore, although prior research has shown that returning customers are of greater value because of increased purchase incidence (Ansari, Mela, & Neslin, 2008), this research indicates that they also more prominently drive high-value conversions. Hence, returning customers are of greater value to companies than new customers and should therefore require more attention. Finally, more frequent interactions are positively related to increased value of conversion. This finding may demonstrate that each interaction has its own stochastic impact and as the effects of interactions accumulate, high value purchases become more likely. This effect was found to be greater for customers in a B2C setting, such that especially in that environment interactions should be stimulated.

Finally, this study provides further indications that basic attribution models are not capable of accurately estimating online marketing channel performance. Both companies currently apply basic attribution models relying on simple heuristics. Basic attribution models have shown to impair attribution measurement, leading to biased inferences concerning channel performance and suboptimal allocation of marketing budgets (Li & Kannan, 2014). In order to obtain more accurate channel attribution estimates this may encourage the participating companies to start using algorithmic multi-touch attribution models instead.
7.2 Limitations and Further Research

This research has several limitations that may serve to stimulate future research. First, whereas Data Set 1 from Company A included non-conversions, Data Set 2 from Company B did not. When optimizing marketing operations, it is of special interest why certain customers have not converted, instead of only being able to optimize the journeys of converting customers. Hence, the inability to track non-conversions results in a large amount of valuable information being lost. Excluding non-conversions is also likely to impair channel attribution estimates of Company B. Further research may investigate to what extent channel attribution estimates are affected from such incomplete attribution analyses. This is particularly useful as basic attribution models commonly used in business practices disregard non-conversions, thereby convincing them of the superiority of algorithmic multi-touch attribution models.

Second, the data sets in this research do not include any information on offline marketing channels. Omitting offline channels does pose certain limitations for this study. Most importantly, prior research has shown that online channels do interact with offline channels and create synergies (Kannan & Li, 2016; Joo, Wilbur, Cowgill, & Zhu, 2014). This research does not deny that these interactions and synergies exist, but is simply limited by the unavailability of data about exposures and responses to offline marketing channels. Furthermore, sufficient methods to track exposures and responses to offline marketing channels are currently lacking. Without a way of joining online to offline data, quantifying offline impact remains difficult, if not impossible (Kaushik, 2008). Future research may contribute by developing such methods and including information about offline channels in the analysis as well.

Third, the two data sets that have been analyzed in this research contain individual-level, but not customer-level, clickstream data. Several customer paths in the data sets may belong to the same customer but general data protection regulation withholds companies from supplying personally identifiable information, and thus, prevents investigating such relations. Although for each customer path in Data Set 1 it is known whether a purchase was made in the last 12 months by the same customer, this information is limited. Additional customer-related information can provide interesting insights into customer buying behavior. Future research may therefore contribute by investigating customer-level data instead of individual-level data.

Finally, attribution modeling is endogenic: it measures the relative effectiveness of channels in a given setting (Li & Kannan, 2014), so results from this study may not apply for other companies in a different setting. Establishing causal relationships would require large-scale field experiments which are challenging to implement in practice. However, there are strong indications that industry context affects channel attribution estimates. This is likely to be caused by different customer buyer behavior in B2C markets compared to customer buyer behavior in B2B markets. Hence, future studies might more prominently anticipate and address such differences.
References


Dalessandro, B., Perlich, C., Stitelman, O., & Provost, F. J. (2012). Causally motivated attribution for online advertising. *Sixth International Workshop on Data Mining for Online Advertising and Internet Economy* (pp. 7:1-7:9). New York, NY, USA: ACM.


https://www.wsj.com/articles/retailers-scale-up-online-sales-distribution-networks-1447792869


The CMO Club & Visual IQ, Inc. (2014). Visual IQ. Opgehaald op October 9, 2017, van Big data, attribution & omni-channel: http://www.visualiq.com/sites/default/files/PDF-download-graphics/study\building\bridges\to\the\promised\land.pdf

https://www.ipria.org/events/conf/Competition_Conference/Measureabiiltiy_and_online_ads.pdf


# List of Figures

- **Figure 1.** The full customer journey ................................................................. 5
- **Figure 2.** Multi-touch attribution ....................................................................... 6
- **Figure 3.** (a) first-touch, (b) last-touch, (c) linear, and (d) time decay attribution ...................................... 7
- **Figure 4.** Exemplary first-order Markov graph .................................................. 17
- **Figure 5.** First-order Markov graph for Data Set 1 (Company A) ....................... 18
- **Figure 6.** ROC curves for Data Set 1 (left: within sample, right: out of sample) .......... 21
- **Figure 7.** Attribution estimates in conversion using different attribution models (Data Set 1) ........... 23
- **Figure 8.** Distribution of conversion value for Data Set 1 .................................. 26
- **Figure 9.** Distribution of conversion value for Data Set 2 .................................. 26
- **Figure 10.** First-order Markov graph for data set 2 (Company B) ....................... 46
- **Figure 11.** Attribution estimates in conversion using different models (Data Set 2) ........ 50
- **Figure 12.** Attribution estimates in conversion value using different models (Data Set 1) ............. 51
- **Figure 13.** Attribution estimates in conversion value using different models (Data Set 2) ............. 52
List of Tables

Table 1. Overview of multichannel marketing attribution studies ......................................................... 8
Table 2. Data descriptions...................................................................................................................... 15
Table 3. Channel distribution .............................................................................................................. 16
Table 4. Operationalization of variables of interest ............................................................................. 16
Table 5. Descriptive statistics ............................................................................................................. 16
Table 6. AUC values for Data Set 1 ................................................................................................... 21
Table 7. Robustness of removal effect: average standard deviation as % of average removal effect ...... 22
Table 8. Removal effects (%) taking conversion and conversion value as attribution measure .......... 24
Table 9. Removal effects (%) per type using conversion and conversion value as attribution measure ... 25
Table 10. Differences in removal effects (in %) between high- and low-value conversions (conversion) 27
Table 11. Differences in removal effects (in %) between high- and low-value conversions (conv. value) 27
Table 12. Returning customers (%) and amount of interactions for low- and high-value conversions ... 28
Table 13. Online marketing channels .................................................................................................. 43
Table 14. Online marketing channel classifications ............................................................................. 44
Table 15. Overview of early marketing channel attribution studies ....................................................... 45
Table 16. Conversion value for customer journeys in Figure 1............................................................... 48
Table 17. Removal effects for exemplary first-order Markov graph in Fig. 1........................................ 49
## Appendix A

<table>
<thead>
<tr>
<th>Online Channel</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Type-In</td>
<td>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.</td>
</tr>
<tr>
<td>Search (SEA/SEO)</td>
<td>A customer searching for a keyword in a general search engine (e.g. Google, Bing, Yahoo!) receives two types of results: organic search results ranked by the search algorithm called search engine optimization (SEO), and sponsored search results through search engine advertising (SEA).</td>
</tr>
<tr>
<td>Price comparison</td>
<td>Internet service platforms that allow users to compare prices and product information.</td>
</tr>
<tr>
<td>Retargeting</td>
<td>A special form of display advertising that uses a customer’s browsing history to deliver personalized banners.</td>
</tr>
<tr>
<td>Affiliate</td>
<td>Commission-based marketing in which a business (e.g. retailer) rewards the affiliate for referring a user towards the business’s website.</td>
</tr>
<tr>
<td>Referral</td>
<td>All traffic that is forwarded by external content websites (with or without remuneration) – for example, by including a text link.</td>
</tr>
<tr>
<td>Email</td>
<td>Sending marketing messages toward potential customers using email. This includes both ads within an email and entirely promotional emails. Also known as newsletter marketing.</td>
</tr>
<tr>
<td>Social media</td>
<td>Comprises a set of advertising platforms belonging to the field of social media, such as Facebook, Twitter and Instagram.</td>
</tr>
<tr>
<td>Application</td>
<td>An application of a company that can be downloaded on a mobile device</td>
</tr>
<tr>
<td>Shopping</td>
<td>A google service which allows users to search for products on online shopping websites and compare prices between different vendors.</td>
</tr>
</tbody>
</table>
## Appendix B

### Table 14. Online marketing channel classifications

<table>
<thead>
<tr>
<th>Channel</th>
<th>Classification approach</th>
<th>Contact origin</th>
<th>Brand usage</th>
<th>Browsing goal</th>
<th>Degree of content integration</th>
<th>Degree of personalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Type-In</td>
<td>Customer-initiated</td>
<td>Customer-initiated</td>
<td>Branded</td>
<td>Navigation</td>
<td>Content-separated</td>
<td>Personalized</td>
</tr>
<tr>
<td>SEO</td>
<td>Customer-initiated</td>
<td>n/a(^{20})</td>
<td>Branded</td>
<td>Navigation</td>
<td>Content-separated</td>
<td>Personalized</td>
</tr>
<tr>
<td>SEA-Branded</td>
<td>Customer-initiated</td>
<td>Branded</td>
<td>Branded</td>
<td>Navigation</td>
<td>Content-separated</td>
<td>Personalized</td>
</tr>
<tr>
<td>SEA-Unbranded</td>
<td>Customer-initiated</td>
<td>Generic</td>
<td>Branded</td>
<td>Navigation</td>
<td>Content-separated</td>
<td>Personalized</td>
</tr>
<tr>
<td>Price comparison</td>
<td>Customer-initiated</td>
<td>Generic</td>
<td>Generic</td>
<td>Information</td>
<td>Content-integrated</td>
<td>Non-personalized</td>
</tr>
<tr>
<td>Display</td>
<td>Firm-initiated</td>
<td>n/a</td>
<td>n/a</td>
<td>Information</td>
<td>Content-separated</td>
<td>Non-personalized</td>
</tr>
<tr>
<td>Retargeting</td>
<td>Firm-initiated</td>
<td>n/a</td>
<td>n/a</td>
<td>Information</td>
<td>Content-separated</td>
<td>Personalized</td>
</tr>
<tr>
<td>Affiliate</td>
<td>Firm-initiated</td>
<td>n/a</td>
<td>n/a</td>
<td>Information</td>
<td>Content-integrated</td>
<td>Non-personalized</td>
</tr>
<tr>
<td>Referral</td>
<td>Firm-initiated</td>
<td>n/a</td>
<td>n/a</td>
<td>Information</td>
<td>Content-integrated</td>
<td>Non-personalized</td>
</tr>
<tr>
<td>Email</td>
<td>Firm-initiated</td>
<td>n/a</td>
<td>n/a</td>
<td>Navigation</td>
<td>Content-separated</td>
<td>Non-personalized</td>
</tr>
<tr>
<td>Social media</td>
<td>Firm-initiated</td>
<td>n/a</td>
<td>n/a</td>
<td>Information</td>
<td>Content-separated</td>
<td>Non-personalized</td>
</tr>
<tr>
<td>Application</td>
<td>Customer-initiated</td>
<td>Branded</td>
<td>Branded</td>
<td>Navigation</td>
<td>Content-integrated</td>
<td>Non-personalized</td>
</tr>
<tr>
<td>Shopping</td>
<td>Customer-initiated</td>
<td>Generic</td>
<td>Generic</td>
<td>Information</td>
<td>Content-integrated</td>
<td>Non-personalized</td>
</tr>
</tbody>
</table>

\(^{20}\) For SEO it is unknown whether a search query was done using branded or generic keywords. Therefore, SEO cannot be classified along the dimension of brand usage.
## Appendix C

**Table 15.** Overview of early marketing channel attribution studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Online / Offline</th>
<th>Channels</th>
<th>Data Type</th>
<th>Method</th>
<th>Insights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papadimitriou et al. (2011)</td>
<td>Online</td>
<td>Display, SEA</td>
<td>Field experiment</td>
<td>Confidence interval analysis</td>
<td>Display advertising increases the amount of relevant search queries</td>
</tr>
<tr>
<td>Lewis and Nguyen (2014)</td>
<td>Online</td>
<td>Display, search</td>
<td>Quasi-experiment</td>
<td>OLS regression</td>
<td>Display advertising increases the amount of relevant search queries</td>
</tr>
<tr>
<td>Kireyev et al. (2013)</td>
<td>Online</td>
<td>Display, SEA</td>
<td>Field data</td>
<td>Persistence modeling</td>
<td>Display advertising increases the amount of SEA clicks and conversions</td>
</tr>
<tr>
<td>Nottorf (2014)</td>
<td>Online</td>
<td>Display, SEA</td>
<td>Field data</td>
<td>Binary logit (Bayesian)</td>
<td>Positive interaction effects between display advertisements and SEA influence consumer click probabilities</td>
</tr>
<tr>
<td>Xu et al. (2014)</td>
<td>Online</td>
<td>Display, SEA, other</td>
<td>Field data</td>
<td>Mutually exciting point model</td>
<td>Display ads stimulate subsequent visits through other channels, such as SEA</td>
</tr>
<tr>
<td>Joo et al. (2014)</td>
<td>Online + Offline</td>
<td>Television, search/SEA?</td>
<td>Field data</td>
<td>Regression (distributed lag)</td>
<td>Television advertisements influence the amount of relevant search queries</td>
</tr>
<tr>
<td>Agarwal et al. (2011)</td>
<td>Online</td>
<td>SEA</td>
<td>Field experiment</td>
<td>Hierarchical Bayesian model</td>
<td>The position of SEA influences the revenues and profits of these advertisements</td>
</tr>
<tr>
<td>Skiera and Nabot (2013)</td>
<td>Online</td>
<td>SEA</td>
<td>Field experiment</td>
<td>PROSAD (automated bidding decision support system)</td>
<td>The position of SEA influences the revenues and profits of these advertisements</td>
</tr>
<tr>
<td>Blake et al. (2015)</td>
<td>Online</td>
<td>SEA</td>
<td>Field experiment</td>
<td>Regression</td>
<td>SEA-branded is ineffective, but SEA-unbranded affect new and infrequent customers</td>
</tr>
<tr>
<td>Lambrecht and Tucker (2013)</td>
<td>Online</td>
<td>Display, retargeting</td>
<td>Field data</td>
<td>Proportional Hazard model</td>
<td>Effectiveness of retargeting depends on whether the customer has well-defined product preferences</td>
</tr>
<tr>
<td>Braun and Moe (2013)</td>
<td>Online</td>
<td>Display</td>
<td>Field data</td>
<td>Bayesian model</td>
<td>Effectiveness of display advertisements depends on the person’s advertisement impression history</td>
</tr>
</tbody>
</table>
Appendix D

Figure 10. First-order Markov graph for data set 2 (Company B)
Appendix E

Removal Effects for Conversion

The core of the removal effect is to remove each channel (C1, C2, C3, and C4) from the graph consecutively and measure how many conversions could be made without that channel. The probability of conversion of the complete model ($P_T$) in Figure 4 is:

$$P_T = (w_{start-C1} \times w_{C1-C3} \times w_{C3-conversion}) + (w_{start-C1} \times w_{C1-C3} \times w_{C3-C4} \times w_{C4-conversion}) + (w_{start-C1} \times w_{C1-C4} \times w_{C4-conversion})$$

$$P_T = (0.75 \times 0.67 \times 0.5) + (0.75 \times 0.67 \times 0.5 \times 1) + (0.75 \times 0.33 \times 1) = .75$$

The probability of conversion of the model after removing channel C1, C2, C3, and C4:

$$P_{T-C1} = (0) + (0) + (0) = 0$$
$$P_{T-C2} = (0.75 \times 0.67 \times 0.5) + (0.75 \times 0.67 \times 0.5 \times 1) + (0.75 \times 0.33 \times 1) = .75$$
$$P_{T-C3} = (0) + (0) + (0.75 \times 0.33 \times 1) = .25$$
$$P_{T-C4} = (0.75 \times 0.67 \times 0.5) + (0) + (0) = .25$$

The removal effects for C1, C2, C3, and C4 are calculated using the following formula

$$Removal\ \text{Effect} \ (s_i) = \Delta P = (P_T - P_{T-s_i})$$

$$Removal\ \text{Effect} \ (C1) = (P_T - P_{T-C1}) = .75$$
$$Removal\ \text{Effect} \ (C2) = (P_T - P_{T-C2}) = 0$$
$$Removal\ \text{Effect} \ (C3) = (P_T - P_{T-C3}) = .5$$
$$Removal\ \text{Effect} \ (C4) = (P_T - P_{T-C4}) = .5$$

After removal effects have been calculated for all channels ($s$), they have to be weighted. This is because the total sum of probability of conversion of the model after removing channels ($\sum P_{T-s_i}$) is bigger than the probability of conversion of the complete model ($P_T$). This provides the relative Removal Effects ($s_i$) in percentages:

$$Removal\ \text{Effect} \ (s_i) \ \text{in} \ % = \frac{Removal\ \text{Effect} \ (s_i)}{\sum \text{Removal\ \text{Effect} (s)}} \times 100\%$$

$$Removal\ \text{Effect} \ (C1) \ \text{in} \ % = \frac{.75}{(.75 + 0 + .5 + .5)} \times 100\% = 42.86\%$$
$$Removal\ \text{Effect} \ (C2) \ \text{in} \ % = \frac{0}{(.75 + 0 + .5 + .5)} \times 100\% = .00\%$$
$$Removal\ \text{Effect} \ (C3) \ \text{in} \ % = \frac{.5}{(.75 + 0 + .5 + .5)} \times 100\% = 28.57\%$$
$$Removal\ \text{Effect} \ (C4) \ \text{in} \ % = \frac{.5}{(.75 + 0 + .5 + .5)} \times 100\% = 28.57\%$$
**Removal Effect for Revenue**

The core of the removal effect for revenue is still to remove each channel (C1, C2, C3, and C4) from the graph consecutively. However, now we measure how much value can be earned after the removal of a certain channel. Table 16 presents additional information about the conversion value for each of the customer journeys in Figure 4.

<table>
<thead>
<tr>
<th>Customer journey</th>
<th>Sequence of channels</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journey 1</td>
<td>C1 – C3 - CONVERSION</td>
<td>$400,00</td>
</tr>
<tr>
<td>Journey 2</td>
<td>C1 – C4 - CONVERSION</td>
<td>$1000,00</td>
</tr>
<tr>
<td>Journey 3</td>
<td>C1 – C3 – C4 - CONVERSION</td>
<td>$500,00</td>
</tr>
<tr>
<td>Journey 4</td>
<td>C2 - END</td>
<td>$0,00</td>
</tr>
</tbody>
</table>

The conversion value for the complete model \( \text{Value}_T \) is:

\[
\text{Value}_T = (\text{Value}_{\text{journey1}}) + (\text{Value}_{\text{journey2}}) + (\text{Value}_{\text{journey3}}) + (\text{Value}_{\text{journey4}})
\]

\[
\text{Value}_T = 400,00 + 1000,00 + 500,00 + 0,00 = 1900
\]

When removing C1, only journey 4 can still take place. When removing C2, all journeys except journey 4 remain. When removing C3, journey 2 and 4 remain. When removing C4, journey 1 and 4 remain. Hence, the value of the models without C1, C2, C3, and C4 respectively:

\[
\text{Value}_{T-C1} = (\text{Value}_{\text{journey4}}) = 0,00
\]

\[
\text{Value}_{T-C2} = (\text{Value}_{\text{journey1}}) + (\text{Value}_{\text{journey2}}) + (\text{Value}_{\text{journey3}}) = 1900,00
\]

\[
\text{Value}_{T-C3} = (\text{Value}_{\text{journey2}}) + (\text{Value}_{\text{journey4}}) = 1000,00
\]

\[
\text{Value}_{T-C4} = (\text{Value}_{\text{journey1}}) + (\text{Value}_{\text{journey4}}) = 400,00
\]

Removal effects for C1, C2, C3, and C4 are still calculated in the same manner:

\[
\text{Removal Effect (s)} = \Delta \text{Value} = (\text{Value}_T - \text{Value}_{T-s})
\]

\[
\text{Removal Effect (C1)} = (\text{Value}_T - \text{Value}_{T-C1}) = 1900
\]

\[
\text{Removal Effect (C2)} = (\text{Value}_T - \text{Value}_{T-C2}) = 0
\]

\[
\text{Removal Effect (C3)} = (\text{Value}_T - \text{Value}_{T-C3}) = 900
\]

\[
\text{Removal Effect (C4)} = (\text{Value}_T - \text{Value}_{T-C4}) = 1500
\]

After removal effects have been calculated for all channels, they have to be weighted again. This provides the relative Removal Effects \( s_i \) in percentages:

\[
\text{Removal Effect (C1) in \%} = \frac{1900}{(1900 + 0 + 900 + 1500)} \times 100\% = 44.19\%
\]

\[
\text{Removal Effect (C2) in \%} = \frac{0}{(1900 + 0 + 900 + 1500)} \times 100\% = 0\%
\]

\[
\text{Removal Effect (C3) in \%} = \frac{900}{(1900 + 0 + 900 + 1500)} \times 100\% = 20.93\%
\]

\[
\text{Removal Effect (C4) in \%} = \frac{1500}{(1900 + 0 + 900 + 1500)} \times 100\% = 34.88\%
\]
Removal effects calculation results in terms of conversion and conversion value are summarized in Table 17. Results for conversion value are quite different from those results from conversion. Whereas the removal effect of channels C3 and C4 were similar in terms of conversion, they are very different based on their conversion value. This shows that removal effects are largely dependent on the attribution measure that is used.

**Table 17. Removal effects for exemplary first-order Markov graph in Fig. 1.**

<table>
<thead>
<tr>
<th>Channel</th>
<th>Conversion</th>
<th>Conversion value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Removal Effect ((s_i))</td>
<td>Removal Effect ((s_i)) in %</td>
</tr>
<tr>
<td>C1</td>
<td>.75</td>
<td>42.86%</td>
</tr>
<tr>
<td>C2</td>
<td>.00</td>
<td>.00%</td>
</tr>
<tr>
<td>C3</td>
<td>.5</td>
<td>28.57%</td>
</tr>
<tr>
<td>C4</td>
<td>.5</td>
<td>28.57%</td>
</tr>
</tbody>
</table>
Appendix F

Figure 11. Attribution estimates in conversion using different models (Data Set 2)
Figure 12. Attribution estimates in conversion value using different models (Data Set 1)
Figure 13. Attribution estimates in conversion value using different models (Data Set 2)