Keyword Segmentation, Campaign Organization, and Budget Allocation in Sponsored Search Advertising

Derek Visser, Master of Science in Management (Information Systems)

Management Sciences

Submitted in partial fulfillment of the requirements for the degree of

Master of Science in Management (Information Systems)

Faculty of Goodman School of Business, Brock University

St. Catharines, Ontario

Abstract

Sponsored search advertising, where search providers allow advertisers to participate in a real-time auction and compete for ad slots on search engine results pages (SERPs), is currently one of the most popular advertising channels by marketers. Some domains such as Amazon.com allocate in millions of dollars a month to their sponsored search campaigns. Considering the amount of money allocated to sponsored search as well as the dynamic nature of keyword advertising process, the campaign budget planning decision is a non-trivial task for advertisers. Budget constrained advertisers must consider a number of factors when deciding how to organize campaigns, how much budget to allocate to them, and which keywords to bid on. Specifically, they must decide how to spend budget across planning horizons, markets, campaigns, and ad groups. In this thesis, I develop a simulation model that integrates the issues of keyword segmentation, campaign organization, and budget allocation in order to characterize different budget allocation strategies and understand their implications on search advertising performance. Using the buying funnel model as the basis of keyword segmentation and campaign organization, I examine several budget allocation strategies (i.e., search *Volume-based*, *Cost-based*, and *Clicks-based*) and evaluate their performance implications for firms that may pursue different marketing objectives based on industry and or product/service offerings. I evaluate the simulation model using four fortune 500 companies as cases and their keyword advertising data obtained from Spyfu.com. The results and statistical analysis shows significant improvements in budget utilization using the above-mentioned allocation strategies over a *Baseline* strategy commonly used in practice. The study offers a unique insight into the budget allocation problem in sponsored search advertising by leveraging a theoretical framework for keyword segmentation, campaign management, and performance evaluation. It also provides insights for advertiser on operational issues such as

keyword categorization and campaign organization and prioritization for improved performance. The proposed simulation model also contributes a valid experimental environment to test further decision scenarios, theoretical frameworks, and campaign allocation strategies in sponsored search advertising.

Table of Contents

List of Figures

List of Tables

List of Equations

Acknowledgements

I would like to thank my supervisor Dr. Anteneh Ayanso for his guidance, help, support and patience through this entire journey. He has been much more than one would have expected from a supervisor. His passion, rigor, and encouragement significantly affected the quality of this thesis. I would like to thank the internal committee members, Dr. Martin Kusy and Dr. Reena Yoogalingam for their comments and suggestions. Their advice was invaluable in all the stages of my thesis. A special thanks go to my family and friends for their continuous support, caring, and encouragement. Without them, I would never have made it this far.

Chapter 1 Introduction

1.1 Background

The Internet has created many opportunities for advertisers to easily attract and retain customers. Internet advertising revenue now exceeds that of traditional advertising channels such as cable, broadcast television, and print media with $42.8 billion in 2013 (IAB, 2014). While the combined revenue of $74.5 billion for both cable and broadcast television exceeds that of Internet advertising, the growth in revenue for the Internet is far outpacing that of television with 17 percent and 3.3 percent, respectively (IAB, 2013; IAB, 2014). There is an abundance of information on the Internet today, and the volume can be overwhelming for users. Search engine providers such as Google and Yahoo! have gone to great lengths to reduce the burden of finding relevant information by offering their Web indexing services (search engines). These services actively crawl the Web and index sites providing users with the ability to input queries, and in return obtain relevant results (images, video, website listings, etc.) from the search engines. In doing so, the search engines simplify the process of locating the relevant information that the user is looking for. As of July 2015, the top three search providers, Google, Microsoft Bing, and Yahoo! each account for 88.4%, 4.1%, and 3.7% of search engine traffic, respectively (Statista, 2014).

Search providers deliver these services to Internet users free of charge. However, they have created mechanisms to monetize the utilization of these services by introducing sponsored search advertising, which is their substantial source of income (Edelman et al., 2007). Specifically, search providers allow advertisers to participate in a real-time auction and compete for ad slots on search engine results pages (SERPs). Advertisers bid on keywords and their advertisements are assigned a position on the SERPs in rank order based on a combination of the value of their bid and their quality score. Quality score is a combination of a number of factors, including landing page quality and expected click-through rate (CTR) (Jansen & Spink, 2009). In general, the top ranked position is worth the most to an advertiser as it generates the most attention. The attention garnered from positions below the first are often assumed to be monotone and decreasing in value, and in many instances, this is true for CTRs (Chitika, 2013; Ghose & Yang, 2009). However, this is not always the case for conversions or revenue generation (Agarwal et al., 2011; Naldi et al., 2010). An analysis of over 192 million impressions and over 2 million clicks in sponsored search advertising shows that the highest rate of conversion occurs in position 4 with approximately 12 per 10,000 impressions (Engine Ready, 2014). An impression is when an advertisement is served to a search user. A conversion is when a search user takes a desired action on the advertiser's website (fills a form, downloads, makes a purchase, etc.). In addition, position results tend to vary by level of competition as well as advertiser type, and product/service type (Ayanso & Karimi, 2015).

In recent years, sponsored search advertising has received significant attention both in academia and industry. A recent Forrester report indicated that, of the 81 online retailers surveyed, 85 percent say that search engine marketing is one of the top three ways to acquire customers online (Callard, 2014). Additionally, it was reported that 45 percent have online revenues exceeding $100 million and that 20 percent allocate 50 percent or more of their marketing budgets to search engine advertising. Of the total Internet advertising revenue in 2013, search advertising represented 43 percent ($18.4 billion). The fastest growing area was mobile advertising, which has increased from 9 percent ($3.4 billion) to 17 percent ($7.1 billion) in 2012 (IAB, 2014). Market leaders in spend on search advertising are focusing heavily on the AdWords platform. For example, Wow.com, Amazon.com, and Webcrawler.com have daily ad budgets of 19.2, 14.7, and 9.5 million, respectively (Spyfu, 2015). With such large sums of money being spent on search advertising, efficient allocation of the budget among different networks, campaigns, ad groups, and keywords is of great importance.

The campaign budget planning decision is a non-trivial task for advertisers, due to the dynamic nature of keyword advertising and the complexity of the search auction mechanism and bidding processes (Yang et al., 2014). Advertisers with limited budgets must consider a number of factors when deciding which keywords to bid on and how much to bid on them. Specifically, they must decide how to allocate budget across planning horizons, markets, or within the scope of a day (Yang et al., 2012) and across campaigns (Yang et al., 2014). In the sponsored search literature, the optimization of bid price and budget allocation has been studied in pursuit of different performance objectives, including the maximization of expected clicks (Feldman et al., 2007; Muthukrishnan et al., 2010; Yang et al., 2014), return on investment (Chaitanya and Narahari, 2010), profit (Cholette et al., 2012), revenue (Dayanik & Parlar, 2013; Özlük & Cholette, 2007; Zhang et al., 2012; Zhou et al., 2008), campaign reward (Zhang et al., 2014a), as well as the minimization of expenditure or loss (Yang et al., 2012). Nevertheless, most of the prior studies on budget optimization fail to consider the functionalities of the current ad platforms provided by search engines, the keyword auction rules, as well as the budget account structure (Yang et al., 2014). In addition, many of these studies lack a theoretical framework as the basis for integrating the issues of keyword segmentation, campaign organization, and budget allocation in sponsored search advertising. This thesis attempts to fill the research gaps between these issues by leveraging existing marketing frameworks as a theoretical guide.

1.2 Research Goal and Contributions

The goal of this thesis is to investigate whether advertisers can improve their campaign management and budgeting decisions by employing keyword segmentation and performance-based budget allocation strategies in sponsored search advertising. The study uses simulation modeling that integrates decisions involving keyword segmentation and campaign organization in order to characterize different budget allocation strategies and understand their implications on search advertising performance. Using the buying funnel model as the basis of keyword segmentation and campaign organization, the study examines several budget allocation strategies (i.e., search *Volume-based*, *Cost-based*, and *Clicks-based*) and evaluates their performance implications for firms that may pursue different marketing objectives based on industry and or product/service offerings. I use the buying funnel model as my theoretical foundation due to its commercial and advertiser-centric focus and prior statistical validation in capturing users' search intent (Jansen & Schuster, 2011). I use simulation as my methodology due to the dynamic nature and complexity of the budgeting decisions. It allows the modeller to capture the high number of decision variables and metrics in sponsored search advertising, which can then be empirically evaluated. Simulation is specifically chosen for its ability to test multiple decision scenarios and multiple performance based allocation strategies. Simulation also provides more flexibility to model the use of prior information in informing current decisions.

In sponsored search advertising, keyword segmentation, campaign organization, and budget allocation are challenging decisions for advertisers, due to uncertainty in demand, a highly competitive keyword market, as well as the current multi-factor ranking mechanisms adopted by major search providers (Ayanso & Mokaya, 2013; Ayanso & Karimi, 2015; Jansen & Spink, 2009). While previous literature has focused on mechanism design and bidding strategies in search advertising, there has been limited research on the campaign budget allocation problem. Considering the complexity of search advertising process, advertisers must manage which keywords to bid on, how to organize ad campaigns and ad groups, and decide on how much they are willing to spend per click for each keyword. In addition, the competitive landscape and fluctuations in keyword popularity in sponsored search make the budget allocation decision across multiple planning horizons a non-trivial task. For instance, an advertiser must decide how to allocate its total sponsored search budget across search markets; across temporal slots; or across ad campaigns, ad groups or individual keywords, continuously (Zhang & Feng, 2011).

Anecdotal evidence from search advertising practices show that most advertisers manage a large number of accounts, campaigns, ad groups, and keywords that vary in performance, budget consumption, and advertising goal orientation. Competition in keyword auctions is also increasingly growing. To remain competitive, firms need to know the various decisions involved in search engine advertising as well as competitors' campaign strategies in both organic (natural) and paid search advertising. The absence of systematic frameworks and decision support tools for many of the tasks involved in this process could lead to arbitrary decisions and resource inefficiencies. Therefore, this study contributes to the sponsored search literature by leveraging the buying funnel model as the theoretical foundation and integrating the campaign budget allocation decision with keyword segmentation and campaign organization issues.

Despite their intrinsic relationships and impact on all other search advertising strategies, previous research in search advertising mostly treated the keyword selection/segmentation problems (Ashkan & Clarke, 2013; Broder, 2002; Jansen et al., 2007; Jansen et al., 2008; Jansen and Schuster, 2011; Li, Pan, & Wang, 2010; Lu & Zhao, 2014; Rusmevichientong & Williamson, 2007; Zhang et al., 2014b) and the budget allocation problem (Cholette et al., 2012; Dayanik &

Parlar, 2013; Feldman et al., 2007; Muthukrishnan et al., 2010; Özlük & Cholette, 2007; Zhang et al., 2012; Zhang et al., 2014a; Zhou et al., 2008) as standalone decisions. While there are various ways to group keywords and their characteristics (Broder, 2002; Ghose & Yang, 2009; Jansen et al., 2007; Lu & Zhao, 2014), the buying funnel model (Jansen & Schuster, 2011) provides a theoretically sound framework to bring together issues related to search advertising campaign strategies, keyword categorization, budget utilization, and ad performance. This is primarily due to the well-defined stages of the model (i.e., Awareness, Research, Decision, and Purchase) and the ability to capture the consumer behaviors associated with each stage of the model using the various performance metrics in keyword advertising (e.g., impressions, clicks, cost-per-click, conversions, etc.). In addition, Jansen and Shuster (2011) empirically demonstrated that the stages of the funnel exist in sponsored search and have statistically significant variations across all the stages and performance metrics. With minor variations in the marketing literature (Caspari, 2004; Webb, 2006), the buying funnel represents a staged process for describing the way consumers make their buying decisions, starting from awareness of the existence of a need to the final purchase stage of a product or service that addresses this need (Jansen & Schuster, 2011). Although the empirical analysis of Jansen and Schuster (2011) indicates that the buying funnel stages may not be associated with expected consumer actions as predicted by the model, their results do suggest that the buying funnel stages can provide a more flexible framework for advertisers to develop campaign priorities and define keyword segments that serve their market and product contexts.

From a theoretical view, this thesis offers a unique insight into the budget allocation problem in sponsored search advertising by leveraging a theoretical framework for keyword segmentation, campaign management, and performance evaluation. Measuring keyword-level performance is important given that the basic mechanism of paid search advertising operates at the keyword level (Özlük & Cholette, 2007; Rutz et al., 2012). As opposed to advertising in traditional media such as television and newspapers, keywords are sold on a per-click basis, rather than on per impression or exposure. Given the significant amount of money spent on keyword advertising, marketers need guidance and insights into this process (Dhar & Ghose, 2010; Jansen, Sobel, & Zhang, 2011; Lu & Zhao, 2014). Therefore, effective keyword management requires identifying and creating a portfolio of keywords for improved budget utilization and the subsequent performance of advertising campaigns.

Inefficiency in the allocation of budget on keywords as well as campaigns is a constant challenge for advertisers. The complexity of the keyword selection and bid process requires a structured and systematic approach for segmentation and budget allocation across campaigns, ad groups, and individual keywords. Among the specific challenges advertisers face is the volatility in search demand, as it can be substantial and may have a detrimental effect on the performance of campaigns by causing budgets to run out early. For example, on any given day, an advertiser's most expensive keyword may be highly searched early in the morning, draining the daily budget, and leaving no money for ads to be shown for the remainder of the day. Thus, in order to better manage their productive keywords, advertisers need to measure performance on a continuous basis and assess their impact on budget utilization. Specifically, advertisers should take advantage of prior performance information to inform their current decisions. Fortunately, search advertising platforms such as AdWords provide various flexibilities to apply specific strategies and actions by advertisers. However, advertisers lack the decision support tools and the insights to evaluate and come up with the strategies and actions that improve their performance. Therefore, from a practical perspective, this thesis provides insights for advertisers on operational issues such as budget utilization, categorizing keywords based on different campaign objectives, and developing campaign strategies and prioritization in order to improve efficiency over time.

The rest of this thesis is organized as follows. Chapter 2 reviews the literature in search engine advertising (SEA). Chapter 3 focuses on the theoretical framework for the simulation model. Chapter 4 presents the description of the problem and the simulation model development. Chapter 5 introduces the empirical data and the performance measures utilized in this study. Chapter 6 presents the simulation results for four different cases along with the discussions. Chapter 7 provides the theoretical and practical implications of this study as well as limitations and future research directions, and finally Chapter 8 provides the study's conclusions.

Chapter 2 Literature Review and Research context

In 1998, Overture (originally Goto.com), which is now owned by Yahoo, invented pay-per-click search advertising with a first-price auction mechanism, by which advertising slots were allocated and sold based solely on the highest bid. Google moved from a first-price auction mechanism to a second-price auction design and integrated their proprietary AdRank algorithm that incorporated the click-through rates originally, and later the quality score, to increase the likelihood of relevant ads being displayed to users (Laffey, 2007). The second-price auction differs from a first-price auction in that, instead of bidders paying their exact bid price, they pay the bid price of the next highest bidder. Following Google, Yahoo and other search providers have since moved to the second-price auction design (Yao & Mela, 2009). This shift in ranking rule resulted from the desire by search providers to attract more advertisers and generate more revenue.

Search providers must also attract and maintain the interest of search users. Therefore, the dynamics of the relationships among advertisers, users, and the search providers are reliant on each participants' level of satisfaction within the system (Jansen & Mullen, 2008). In other words, the more users that utilize a given search provider's service, the larger and more attractive the platform is for advertisers; and the higher the quality of the search provider's service, the more likely it is to attract users and generate more revenue for search engines. Given the popularity of search advertising in recent years, academic research in this area has grown tremendously in the Marketing and Information Systems fields. Below I provide a comprehensive review of the literature in the following categories: mechanism design and equilibrium outcomes in keyword auctions; consumer search behaviour and advertising performance; keyword selection and advertising performance; and budget allocation and optimization.

2.1 Mechanism Design and Equilibrium Outcomes in Keyword Auctions

The early research in sponsored search advertising focused on mechanism design and analysis of equilibrium outcomes in keyword auctions. Sponsored search results differ from natural or organic search results in a number of ways. Sponsored results are located on top of and beside organic search results. The criteria for seeing a sponsored ad differs from organic results and also differ across search engines. As Google has established a dominating position with approximately 88.4% of the search market (Statista, 2015), this thesis focuses primarily on its marketplace. However, the results and implications of this study can be extended to other sponsored search networks due to similarities across platforms. Google uses its proprietary AdRank algorithm to calculate ad position based on bid value and quality score. The quality score is a composition of a number of factors including relevance, quality of landing page, quality of text ads, and prior CTR (Lahaie, 2006; Liu & Chen, 2006). Google's AdWords employs a Generalized Second Price (GSP) auction mechanism that is an extension of the Vickery Clarke Groves (VCG) mechanism (Clarke, 1971; Groves, 1973; Vickrey, 1961). This is in contrast to a first price auction that awards the top slot to the highest bidder.

The incorporation of the quality score and Ad Rank algorithm into Google's GSP auction creates unique challenges and opportunities for advertisers. Such that, advertisers with more experience and a keen eye for analytics are better able to use this characteristic to their advantage. Not all advertisers follow the same strategy. Maximization of revenue, top position, lead generation, and exposure are a few examples of the different goals advertisers compete for.

A number of researchers have studied sponsored search auctions. Feng et al. (2007) compared several ranking mechanisms, including Overture's and Google's slot allocation rules and found that their performance depends on the correlation between advertisers' willingness to pay and their relevance to the search term. In an intermediated search market, where ads are shown via paid referral, Weber and Zheng (2007) found that to generate maximum revenue, the mechanism should place ads based on a weighted average of bid amount and product performance. Studying auction equilibria under the GSP rule and under a complete information setting, Edelman et al. (2007) and Varian (2007) have shown that the auctioneer's equilibrium revenue under the GSP rule is at least as high as that under the VCG mechanism. Additionally, under the complete-information setting, Edelman et al. (2007) has proved that bidders will not bid their true valuation and that there is no dominant strategy equilibrium in GSP auctions. Kim et al. (2012) analyzed the optimal slot design problem and concluded that varying the number of slots based on the expected valuation distribution of advertisers by keyword is superior to varying the minimum bid as proposed by Edelman and Schwarz (2010).

Considerable research has been conducted on the design of keyword auctions with advertisers' past performance information in various settings (Chen, 2010; Lahaie, 2006; Liu & Chen, 2006, Liu et al., 2010). Comparing ranking rules that consider only bid price as well as the product of CTR and bid price, Lahaie (2006) and Liu and Chen (2006) found that the combined measure is more efficient, but neither of the rules is revenue-maximizing. Liu et al. (2010) extended the study by Liu and Chen (2006) in a generalized multi-slot setting to study the use of differentiated minimum bids, together with the weighted ranking rule, as a way of exploiting ex ante information on advertisers. Additionally, Chen (2010) further extended this to a situation in which advertisers may improve their performance for their own interest in reaction to the performance-based ranking policy.

Zhu and Wilbur (2011) empirically examined the effects of a hybrid auction that incorporates bids for both cost-per-click (CPC) and cost-per-one-thousand impressions (CPM).

They examined two types of advertisers: those who are interested in immediate results such as increased sales, and those who are brand focused and interested in creating goodwill through exposure that can be used persuasively in the future. From an ROI maximization standpoint, Auerbach et al. (2008) examined advertisers' behavior within the Yahoo! Webscope on the top 1000 keywords, and tested a model on both first and second-price auctions. They classified bidders by levels of ROI maximization and found that within both auction types, the majority of advertisers are pursuing a simplified form of ROI maximization, or full ROI maximization strategies.

Considering ad quality in sponsored search advertising, Animesh et al. (2010) examined keyword auction designs to determine whether low-quality firms are able to mimic and execute the strategies of high-quality firms. They identified that low-quality firms are in fact able to attain high positions within sponsored search markets that only consider bid price for allocating advertising slots on SERPs. However, when a search provider incorporates an intervening mechanism such as Google's AdRank algorithm that applies a quality score to each advertiser, this same effect was not observed. Thus, the incorporation of additional attributes that measure quality of an advertiser within the auction mechanism design aids in the market's ability to protect advertisers from direct threats of new entrants and low-quality firms. These results also further emphasize the importance of attaining and maintaining a good quality score in the competition for ad slots. In general, most of the studies in this category provide the theoretical foundations in search advertising research.

2.2 Consumer Search Behaviour and Advertising Performance

User perception and behaviour in sponsored advertisement has been another area of focus in the sponsored search literature. The purpose of investigation is to examine the factors that impact the user's actions such as clicking on ads, or purchasing a product (i.e., conversion). By identifying how users interact and react to advertisements within the search engine results pages (SERPs), advertisers are better able to tailor their sponsored search initiatives to achieve marketing goals. The return advertisers seek need not be specifically in the form of increased revenues or profits, as it may also be their goal to create awareness, and further persuade and convince search users to trust their brands.

Most of the research on users search behavior is focused on the interactions of users within search engines that include separate and distinguishable ads (i.e., sponsored and organic ads). For example, to analyze user behaviour when the lines between organic and sponsored search results are not clear, Jansen and Spink (2007) analyzed data from a meta-search engine (that combines sponsored and organic results) web log of nearly 2 million queries. Utilizing Broder's (2002) categories; informational, transactional, and navigational, the researchers ordered the query patterns of the roughly 7 hundred thousand users chronologically to predict CTR. Interestingly, they found that users are no more likely to click on sponsored ads when the ads are mixed with organic results, and the number of clicks on sponsored ads actually decreases. Due to the character and spacing limits placed on sponsored ads, in contrast to the more information rich organic results, this result may not be surprising. Thus, advertisers should consider richness of ad format when selecting the type of advertisement to use. Additionally, the study found that the percentage split into categories is heavily skewed toward the informational category, as opposed to the other two categories (Jansen & Spink, 2007).

On the other hand, Yang and Ghose (2010) examined the effects that organic results have on sponsored ad click-through behaviour and the inverse relationship. Employing aggregate consumer level data, the researchers empirically tested for interdependence between organic and sponsored search engine results and their effects on profitability. Their findings show that there is a positive interdependence between both listings, with a stronger relationship from organic to sponsored results, which suggests that both Search Engine Optimization (SEO) and Pay-per-Click (PPC) advertising will increase a firms' revenues by increasing the likelihood a user will interact with a link. Furthermore, they tested to see the combined effect of having both organic and sponsored ads and found that there is a significant increase in profit potential by having both present on SERPs. This finding indicates that an advertiser will benefit from balancing SEO and PPC, while the absence of SEO, or good organic results placement mostly diminishes profit potential.

Ghose and Yang (2009) also found that landing page quality positively impacts the cost-per-click (CPC) that the advertiser pays as well as the likelihood of conversion. This implies that the better an advertiser's landing page is with respect to all aspects of Google's Ad Rank algorithm, the less expensive a higher position is on the SERP. Thus, advertisers with limited budgets should allocate more resources to improving quality score. Additionally, they found that consumers are more likely to click on retailer-specific keywords and less likely to click on longer or brand-specific keywords.

To address limitations in extant literature, Braun and Moe (2013) investigated the effects of ad creative heterogeneity by incorporating numerous ad creative characteristics into their model. Frequency and timeliness of ad exposure can enhance or hinder an advertiser's goal of generating conversions online. Braun and Moe (2013) explored this issue by examining ad impression histories while varying ad creatives and incorporate advertising goodwill, a construct that allows for accumulation and decay of "goodwill". Through simulation, it was shown that the proposed model increases expected visits by 12.7 percent and increases conversions by 13.8 percent (Braun

& Moe, 2013, p. 754). Thus, by varying ad creatives and by creating ad impression history policies, the advertiser can significantly improve acquisition and conversion of their target audiences. The study by Brettel and Spilker-Attig's (2010) on differences across national cultures further supports this argument, as their results indicate that it is also beneficial to vary ads based on location. This, in combination with the frequency adjustment tool available to advertisers in Google's AdWords, can aid in the pursuit of revenue generation while limiting loss from overexposure. "Conventional wisdom is that the optimal frequency for most campaigns is around 4-7 ad exposures, and much beyond that, results hit a point of diminishing returns for both brand and direct response objectives" (Agashe, 2010, p. 11). Therefore, by varying and tracking multiple ad creatives, an advertiser can increase the likelihood of a user clicking.

Kim and Sundar (2010) empirically examined whether the negative effects of ad clutter are moderated by ad relevance and that perceived intrusiveness of an ad mediates the attitudes of consumers towards ads and ad hosting websites. Their results indicate that relevance of an ad plays an extremely important role in reducing negative perceptions of ads in cluttered and non-cluttered environments. However, perceived intrusiveness does not operate as a mediator of consumer attitudes. This study was limited to a small sample of respondents and may not replicate in varying conditions. However, these results suggest that an advertiser should ensure that relevance is a priority when creating advertisements. Similar to Kim and Sundar (2010), Goldfarb and Tucker (2011a) utilized a survey and examined user behaviour, empirically testing for factors that cause users to balk at online advertisements. The results indicate that user's privacy concerns play a crucial role in determining the likelihood of clicking on an advertisement. Additionally, the analysis showed that an ad with a combination of being highly targeted (to the website or results pages it is displayed on) and very obtrusive (bigger and bolder details) turns users off. On the other hand, each setting separately increases the likelihood of a user interacting with an ad. This is relevant to both of Google's advertising platforms, AdWords (PPC) and AdSense (Google's display network). Specifically, this suggests that well targeted plaintext ads will benefit the advertiser in both domains and the deployment of obscure but boldfaced and attractive ads on AdSense will also garner more interaction.

The ability of the advertiser to track ads that impact a consumer's purchase raises additional challenges. To address this issue Abishek et al. (2012) proposed a dynamic hidden Markov model using the stages of the conversion funnel (buying funnel) to model the attribution problem. They found that display ads behave as a primer for the movement of consumers from the initial stage of dormancy to the latter stages of the model. In general, the studies dealing with consumer search behaviour have a number of implications for advertisers in their pursuit of efficient allocation of resources as well as maximizing advertising performance.

2.3 Keyword Selection and Advertising Performance

The volume of keywords to choose from is enormous, considering the combinations and permutations of virtually unlimited strings of words and symbols that may be relevant for an advertiser. Advertisers can select keywords using a variety of keyword suggestion tools available online (see, for example, Chen (2010) for a list). Nevertheless, the identification and selection of high utility keywords for a specific marketing goal is a non-trivial problem. Yang et al. (2013) formulated the selection of keywords as an optimization problem with a multi-level computational framework that creates two distinct pools of keywords (Domain-specific and Market-level), that are subsequently assigned to campaigns and ad groups and adjusted dynamically. Ji et al. (2010) empirically examined a set of keywords using characteristics such as keyword length, rank, and

CTR to predict the identification of high potential candidate keywords and generate additional and more relevant keywords. Abhishek and Hosanagar (2007) developed a keyword generation system called Wordy that seeks out similar, highly relevant, and potentially under-utilized keywords using semantic similarity. Others have looked at this problem using existing keywords, ad quality, and the advertiser's propensity to pay (Feng et al., 2007), and by selecting keywords from a profit-to-cost ratio ranking (Rusmevichientong & Williamson, 2006). While each of these approaches can benefit the advertiser by identifying keywords of higher value, the processes of generating keywords is computationally intensive or require significant amount of historical data.

Furthermore, the value of a keyword depends on the criterion an advertiser chooses to use. For example, the popularity, competition, bid valuations, and conversion value to advertiser may dictate the value of a keyword. In addition, there are keyword combinations that can lead to negative effects on campaign spend, and certain words are more open to invalid clicks from customers seeking unrelated products. To address this issue, Kiritchenko and Jiline (2008) used classification and machine learning techniques to identify keywords and their combinations along with adding the removal of negative keyword associations into their model to improve the keyword selection process. The use of negative keywords is an important task as it reduces the number of invalid clicks that the advertiser must pay for (i.e., clicks by users who may not be actually looking for product or service offered by the advertiser). As a simple example, a horse shoe sales company would likely bid on "horse shoes" and a user searching for "horse shoe pit" would trigger its ad. If the user clicks, this results in an invalid click and would cost the advertiser.

Not all firms are equally efficient in practicing PPC advertising. As Ayanso and Mokaya (2013) found through their analysis of 200 of the top 500 Web retailers, many firms are characterized by inefficiencies in terms of different resources and performance measures.

Employing data envelopment analysis combined with principal component analysis (PCA-DEA), they identified significant variations in the efficiency of firms that are Web-only versus multi-channel. They found that efficiency for multi-channel retailers (MCRs) is primarily explained by the performance metrics of sales, conversion rate, impressions, and CTR; whereas, organic and sponsored search ranks mainly reveal the differences in efficiency of web-only retailers (WORs). This suggests that each type of retailer must focus their efforts and competition strategies accordingly. Ghose and Yang (2009) also empirically examined keyword-level effects and found results that challenge the previous assumption that value per click is uniform across ad positions (Aggarwal et al., 2007; Edelman et al., 2007; Varian, 2007). They also found that retailer-specific keywords increase click-through behaviour by 14.27%, whereas brand-specific keywords decrease click-through by 56.6%, and keyword length decreases clicks by 13.9%. These results suggest that keywords should be managed in segments as their performances depend on different categorizations.

This thesis addresses the keyword management issue in terms of keyword segmentation that is aligned with the specific campaign objectives of an advertiser. More specifically, I draw on the buying funnel framework (Jansen & Schuster, 2011) for segmenting keywords and organizing campaigns based on this categorization to serve an advertiser’s product and market context.

2.4 Budget Allocation and Optimization

Despite being the antecedent to all other sponsored search decisions, budget and its allocation has mostly been overlooked in the literature. Given that no advertiser has an unlimited pool of capital to spend on sponsored search campaigns, the problem of determining how to allocate available resources is very important. Advertisers must determine the maximum daily amount they are willing to spend on sponsored search, how many and which keywords to bid on, and how much to bid on them. In recent years, many researchers have recognized the potential for improving the allocation of budget within the search advertising paradigm.

The budget allocation problem has been investigated as part of other sponsored search decisions where budget is defined as a constraint for these decisions (Chakrabarty, Zhou, & Lukose, 2008; Kitts & LeBlanc, 2004). Other related studies have looked at optimizing the allocation of budget across keywords using bid price settings (Chaitanya & Narahari, 2012; Cholette et al., 2012; Dayanik & Parlar, 2013; Feldman et al., 2007; Muthukrishnan et al., 2010; Özlük & Cholette, 2007; Zhou et al., 2008), optimizing bid price settings and budget allocation over temporal slots or markets simultaneously within a single campaign (Yang et al., 2012; Zhang et al., 2014), jointly optimizing bid price settings and campaign budgets (Zhang et al., 2012), and optimizing budget allocation over several campaigns while considering substitution and complementarity effects (Yang et al., 2014).

By adjusting bid prices, advertisers are able to attain different positions on the SERPs and different click-through performance based on the resulting positions. For some keywords such as the ones that include brand name or highly competitive keywords, the top position may be the most desirable. Due to the variation in demand, competition, and performance across keywords, determining the optimal bid price for each keyword under budget constrain is a complex task that many researchers have attempted to address. For example, under a budget constraint, Özlük and Cholette (2007) created a deterministic model that compares and sets the bid price of two separate keywords based on their expected performance outcomes. They generalized the problem to N keywords and consider a condition of adding a new keyword to the pool (N+1), where if the value of the keyword being introduced is sufficient, some of the budget used for the other keywords is allocated to it. Cholette et al. (2012) improved upon this previous model by introducing the probability to rank as a function of bid price, and showing how this affects the trade-off between profit and risk and its efficiency frontier on a single keyword. They also expanded their model to multiple keywords, but found that the problem quickly becomes computationally too expensive with a probabilistic budget constraint. The results of their analysis showed that by adding a probabilistic budgeting constraint, an advertiser's expected revenue actually falls.

Dayanik and Parlar (2013) formulated bid price setting as a dynamic stochastic net revenue optimization problem and compared alternative bidding strategies varying in computational intensity and policy with a hard budget constraint in different advertising scenarios. They found that the dynamic bidding strategy, solved using dynamic programming that continuously adjusts bid price based on remaining daily budget, outperforms all other models in net revenue. Feldman et al. (2007) examined complex and simple bidding strategies and found that a two-bid uniform bidding strategy, where all keywords are bid on with the same price or all with a second price, outperforms the optimum by 63 percent. Chaitanya and Narahari (2012) considered budget and bid optimization problems via linear programming, employing two different bid strategies called OPT and MAX, both using a greedy approach that is designed to reach an envy-free equilibrium between which keywords to select and how much to bid for each. Each model performs better than baseline comparisons under different circumstances of synchronicity. While they do solve this analytically, it becomes extremely complex and beyond the capability of most advertisers to employ it practically. Additionally, the setting of bid price problem has been analysed with multiple dynamic click estimation models (Muthukrishnan et al., 2010) and as a multiple choice knapsack problem optimized with stochastic and deterministic bidding algorithms (Zhou et al., 2008). While these studies improve upon the bid price setting problem, they do not consider the hierarchy of budget accounts. In addition, some only consider a single advertising slot, which does not reflect the actual PPC bidding tools currently supported by search providers.

Zhang et al. (2012) analyzed p-accounts (partially running out) and proposed the joint optimization of both bid price and campaign budget allocation using sequential quadratic programming. They reported the significance of lost opportunity of the partially running out accounts from both the advertiser and search provider perspective. In addition, they highlighted the challenge of allocating campaign budgets across the account with reported average number of campaigns of 15 and a maximum of over 2000. The proposed joint optimization of bids and campaign budgets achieves improved performance across all performance measures, such as average ad impressions, average expected clicks, advertiser revenue, and search engine revenue. However, the focus is on the maximization of utility or value per click (VPC) and does not consider the purpose of any of the campaigns or of the sponsored search initiative as a whole.

Zhang et al. (2014) examined the budget allocation and bid price setting problem as real time adjustments of bid prices and budget to control how spend is allocated across a temporal period (day). While demonstrating improvement over a baseline methods, the proposed method of allocation ignores the context of keywords and the underlying purpose of the campaigns they are placed in. The study also focused on the way budget is spent over time under a single campaign and does not consider the interdependencies of the keywords or marketing goals.

Yang et al. (2014) proposed a method of campaign budget allocation using marketing concepts of substitution relationships or overlapping degree. Using a tri-dimensional space to show the degree of overlap among promotional periods, campaign contents, and target areas, they demonstrate that a high degree of overlap reduces the overall advertising performance in terms of revenue and that advertising effort can be diminished if the degree of overlap between campaigns is ignored. These findings serve as a good motivation for this thesis and the problem it attempts to address. While their study identifies the importance of substitution effects such as the degree of complementarity between two campaigns, it does not incorporate keyword segmentation as the basis for campaign organization and budget allocation decisions.

The challenges faced by advertisers in optimal bid price setting, keyword selection and segmentation, and budget allocation across campaigns, make sponsored search advertising a lucrative area for research. The determination of optimal bid price across an array of diverse keywords is complex and often computationally expensive. While this solves bid price problem for each campaign separately, it does not take into consideration the performance of all campaigns simultaneously and the allocation of the account level budget across the different campaigns and ad groups to ensure better overall advertising performance. Motivated by these challenges and the findings of the related studies reviewed above, this thesis addresses the sponsored search budget allocation problem by examining different budget allocation strategies through simulation modeling. Using the buying funnel framework as the basis of keyword segmentation and campaign organization, this thesis attempts to bridge the research gaps between the issues of keyword segmentation, campaign organization, and budget allocation in the search advertising body of knowledge.

Chapter 3 Theoretical Frameworks

Advertisers face several challenges when planning sponsored search campaigns, which include deciding how much budget to allocate over a timed horizon, and determining which keywords to bid on as well as how much to bid on them. For example, a company with a single product/service that operates in a single region or market may have numerous reasons for having multiple adverting campaigns. Specifically, the company's product or service may have different price points, serve more than one purpose or have a variety of benefits, target multiple demographic groups, or may be marketed via multiple distribution channels (online or offline), or via different promotions. Campaigns may be organized based on any number of strategies such as positioning in different regions or countries, positioning amongst competitors, or any other form of market segmentation, all of them providing a rationale for keyword categorization, campaign organization and prioritization, and resource allocation.

Recognizing this challenge, Fischer et al. (2014) developed a decision support model for allocating budget across countries, products, and marketing channels such as direct mail, and print media. By dissecting the global budget (in sponsored search, the equivalent of account level budget) into country-product-marketing activity level, they are able to incorporate multiple criteria into the allocation heuristic, such as size of business, profit contribution margin, growth potential, short term potential, time value of money, and carryover effects. In doing so, their model increased overall profit improvement potential by approximately 50%. Of particular note is the incorporation of carryover effects and a growth function based on the profile of a product, which is modelled according to diffusion along the product development life cycle. Their model also incorporates a sponsored search market specific adaptation of these two criteria and the use of a modified heuristic with keyword related criteria that capitalize on the characteristics of the sponsored search auction environment.

There are a number of strategic marketing frameworks that provide firms with meaningful ways to plan their marketing initiatives. This thesis is particularly motivated by prior empirical consumer search behavior research that provides the basis for keyword categorization, resource allocation, and performance improvement (Broder, 2002; Jansen et al., 2007; Jansen et al., 2008; Jansen & Schuster, 2011; Nottorf & Funk, 2013). These frameworks include the taxonomy of search user intent (Broder, 2002; Jansen et al., 2007; Jansen et al., 2008), spillover effects from generic to branded search (Nottorf & Funk, 2013), and more importantly, the buying funnel model that is empirically tested by Jansen and Schuster (2011).

3.1 Taxonomy of Search User Intent

The literature on the classification and categorization of keywords is focussed on either the user's or advertiser's perspective. The bulk of the literature to date has been conducted on the former. The expansion of this area of research benefits the advertiser by allowing heterogeneity across users' search behaviours or identifying criteria for selecting more profitable keywords. One of the early taxonomies used for classification of search user intent was introduced by Broder (2002). It includes three categories: Informational, Navigational, and Transactional. In the Informational category, the user is finding information expected to be present on a page or multiple pages; in the Navigational category, the user is interested in reaching a particular site; and in the Transactional category, the user intends to further engage with the internet by shopping or attaining downloads or accessing additional web enabled tools. The results of the random Alta Vista survey were that roughly 20 percent of queries are navigational, 48 percent are informational, and 30 percent are transactional. However, the sample for the survey was small and the categories were not clearly defined (Broder, 2002).

To further refine the definitions and classification of user intent using these categories, Jansen et al. (2008) derived additional attributes of each. In addition to creating a more comprehensive classification schema, they also manually classify a random sample of 400 search queries and develop an algorithm for automatic classification of web queries. The data consisted of a Web transaction log with over one and a half million queries, and the automatic classification algorithm resulted in an overall accuracy of 74 percent. They concluded that approximately 80 percent of all Web searches are informational, 10 percent navigational, and 10 percent transactional. This taxonomy and the operationalized definitions offer additional value to an advertiser by providing a basic structure of how users are querying search engines. An additional benefit is that this ontology can be directly applied to an existing search advertising campaign and used to pursue objectives related to each category. However, this taxonomy only identifies a small portion of transactional queries using a Web transaction log that advertisers typically do not have access to, and the results do not offer actionable marketing insights beyond informing advertisers that the majority of search queries are informational in nature. Further, the authors acknowledge that roughly 25 percent of the queries were misclassified and that many of the queries could be classified as one or more type. Operationalizing this taxonomy in the sponsored search domain may not be also a straightforward process as it does not primarily focus on understanding the commercial intent of a search user (i.e., purchase a product or sign up for a service).

Ashkan and Clarke (2013) developed the classification of search user intent further to include a more comprehensive hierarchical ontology. They extended the categories by developing a subcategory called commercial intent, and subsume transactional queries into either navigational or informational for simplicity. User's queries are split into "major dimensions of query intent", commercial/non-commercial, where the commercial label is applied when it is assumed that the user is planning to purchase at the time of the query or in the future (Ashkan & Clarke, 2013, p. 431). The "subcategories of commercial intent" are retailer, product, and brand. These subcategories effectively differentiate between the user's purpose of search for products/services (p. 431). Utilizing a Web transaction log obtained from the Beyond Search program (developed by Microsoft Research and Microsoft adCenter in 2007) of roughly 100 million search impressions, the researchers identified variation in estimated CTRs across categories and subcategories of commercial intent. While this study provides an advertiser with additional context in which to classify users across keywords, it does not provide additional insights beyond identifying keywords with commercial intent. Moreover, the authors did not define explicitly the characteristics of each category – commercial and non-commercial.

Bridging the gap between the search user intent taxonomy and its commercial/non-commercial extension is the work of Rutz and Bucklin (2011), which shows that there is a positive and asymmetric advertising effect from generic to branded terms, and its expansion by Nottorf and Funk (2013). Using a dynamic linear model with Bayesian estimation, Rutz and Bucklin (2011) showed that there is a positive and asymmetric advertising effect from generic to branded terms. Utilizing two latent constructs, awareness from impressions, and awareness from clicks, Nottorf and Funk (2013) also proposed a model that accounts for generic impressions and clicks in addition to the carryover and spillover effects of previous searches with their respective parameters. The model captures impressions, clicks, conversions and keyword-level data that is classified into branded or generic and then into industry related subgroups. They also controlled for seasonality and measure for weekly, daily, and hourly effects. As the user begins to search for a product, the search can be very broad and lacks specific context such as searching for “new car”, and after subsequent searches and exploration, the user ends up searching for “mid-size Nissan 4 door”. As the user progresses to the final search term, an advertiser may be able to attract attention by advertising its product/services at different levels of broadness. Nottorf and Funk (2013) developed dynamic linear models and determined that there are significant differences in spillover effects across industries and suggested that high involvement and longer term purchase decisions induce more spillover. While this framework does consider the commercial intent of search users, it does not provide a comprehensive categorization of search queries.

Motivated by the works of Rutz and Bukilin (2011) on spillover effects, and Ghose and Yang’s (2010) examination of cross-category purchasing in sponsored search, Lu and Zhao (2014) categorized keywords into three categories: general, specific, and “other” (or irrelevant). General and specific categories reflect the stages of the shopping goals theory as defined by Lee and Ariely (2006). Shopping goals theory recognizes the development of more concrete shopping goals as the consumer proceeds through their shopping process. The theory has two stages. In the first one, the consumer is broadly defining and developing their consideration sets and shopping goals. In this stage they are susceptible to persuasion from external sources. Once the consumer has formed solid shopping goals, in the second and final stage, they seek out satisfying these goals. This suggests that the consumer is more likely to remain focussed on these goals and be less inclined to change their mind via external stimuli such as advertisements. The researchers argue that general keywords are used more by consumers in the early stage without shopping goals, and that specific keywords are used by those in the later stages with more concrete shopping goals. Using a dataset from Taobao.com for the “digital camera” product category, Lu and Zhao (2014) identified products as either main or accessory, and attributed sales to either direct sales or indirect sales.

They model the problem using two structural equation models, one with direct sales as the dependent variable, and the other one with indirect sales as the dependent variable. Using product type as a moderator, they found that shopping goals theory can be used to explain consumer behavior in search engine advertising. Their model confirms that direct sales via specific keywords are more likely from consumers with well-formed shopping goals and that they are less likely to be persuaded. In contrast, general keywords lead more to indirect sales which suggests that consumers using general keywords may not have well-formed consideration sets or shopping goals and are more likely to change their minds and purchase other products. In addition, they found that product type is a moderator of specific keywords on direct and indirect sales.

Each of these ontologies provides a lens into the way consumers interact with search engines. However, the search user intent model does not sufficiently explain or operationalize how advertisers should be selecting keywords. Specifically, it does not help an advertiser to decide how to attract, motivate, or convert consumers from informational or navigational to transactional. It only offers a means to segment keywords based on the type of searcher they are assumed to attract. The latter research on generic to branded keywords and from general to specific keywords do recognize the commercial or purchase intent of search users and confirm that consumers do progress through different stages of intent. However, while they are empirically validated, the models do not provide advertisers with sufficient actionable descriptions of their classification schemes. They also provide limited diversity across keywords, with only two categories in each case.

In contrast, the buying funnel model, which is empirically validated by Jansen and Shuster (2011), provides four distinct categories and captures the spillover effect or progression of consumers into different stages of purchase intent. The buying funnel model is well known in practice and it does not assume indirect or direct sales typical of multiproduct or multiservice providers. Thus, the model is applicable to any advertiser in a B2B or B2C context. Considering its comprehensive classification scheme and popularity in practice, in this thesis I employ the buying funnel model as the main theoretical framework for keyword segmentation and campaign organization.

3.2 The Buying Funnel Model

The buying funnel is a popular marketing paradigm that is commonly utilized in search marketing initiatives (Jansen & Schuster, 2011). Sometimes referred to as the buying cycle or sales funnel (Caspari, 2004; Webb, 2006), it is considered to be a sequential process consumers follow when making a purchase (Ramos & Cota, 2008; Seda, 2004), similar to the progression as suggested in the shopping goals theory (Lu & Zhao, 2014). It is a relatively simple yet powerful model of how consumers behave when interacting with advertisements, and is derived from traditional models such as the marketing funnel (Howard & Sheth, 1969; Myerson & Scarborough, 2007; Young et al., 2006), conversion funnel (Abhishek, Fader, & Hosanagar, 2012), AIDA model (Lancaster & Withey, 2006), and hierarchy of effects model (Lavidge & Steiner, 1961). Jansen and Shuster (2011) also emphasize that the buying funnel is founded on information processing theory and consumer behaviour models (Bettman et al., 1998). It is referred to in the sponsored search literature as the purchase process as early as 2009 (Naldi et al., 2009), and as early as 1996 in the online advertising literature (Hoffman & Novak, 1996). Indicating the presence of several minor variations in the marketing literature, Jansen and Schuster (2011) describe the buying funnel model as a four-stage model, which includes the stages of awareness, research, decision, and purchase. According to their descriptions, awareness refers to a consumer noticing a product or service that will satisfy a need or desire, which is then followed by a research stage in which the consumer actively seeks out additional information related to the product type in order to become more knowledgeable about it before making a decision. In the decision stage, the consumer decides between brands of a specific product by forming a choice set before making a purchase in the final stage. In the purchase stage, the consumer knows what product and brand they want to purchase, but the consumer is doing price or bundle comparison prior to making a purchase.

Based on the above descriptions, Jansen and Schuster (2011) classified keywords into the buying funnel and empirically validated that each stage (Awareness, Research, Decision, and Purchase) does in fact exist. However, while each stage is statistically different by all measures of performance, they noted that consumer progression through the funnel may not follow the traditional approach from awareness through purchase, but instead consumers may choose to purchase at any stage of the funnel. This phenomenon is also noted in Lu and Zhao's (2014) study and may be explained by indirect sales from general keywords. Additionally, Abishek et al. (2012) used a hidden Markov model with the buying funnel as its stages and found differences in consumer behaviour and contributions by online marketing channels (display vs. search). For example, the number of orders, number of items ordered, and average sales revenue were found to be highest in the awareness stage which shows that the bulk of purchases are made in the awareness stage, not the purchase stage (Jansen & Schuster, 2011). In addition, awareness queries were found to have the highest number of clicks and impressions, but the lowest average CPC. This is counter to the expectation that the purchase queries would have the largest revenue generation and most orders. Nevertheless, their classification scheme and the results of their analysis provide the motivation as well as the theoretical foundation of this study, due primarily to the flexibility of the framework in applying the stages of the buying funnel to an advertiser's unique market and product context. These results are also in line with the results of Nottorf and Funk's (2013) study of spillover effects from generic to branded searches, where they suggested that depending on industry and product type, it is in the advertiser's interest to look into broader terms as keyword choices.

Other researchers also identified that product type plays a role in consumer purchase behaviour at different purchase intent stages (Lu & Zhao, 2014). Therefore, the buying funnel model provides a more comprehensive theoretical basis for this study to develop keyword segmentation as well as campaign management and prioritization in the sponsored search domain. While valuation based on conversions or clicks may seem logical as the performance measure to evaluate sponsored search outcomes, their sole pursuit may not apply to all advertisers when an advertiser is not attempting to sell a service or good, and instead may be seeking to increase awareness of an event or brand (Naldi et al., 2010). Jansen and Schuster (2011) acknowledged counter-intuitive results in the flow of searchers from one cognitive stage to the next within their data, with more searches happening in the research stage and more revenue generated at the awareness stage. This may be due to the significant differences in how a shopper is provided information on a product when seeking online versus offline. Specifically, search engine result pages and the landing pages the ads point to are information rich ad spaces.

Chapter 4 Keyword Advertising Decision Scenario

Search advertising has become a dynamic and more complex process, where advertisers select keywords, create multiple ad campaigns and ad groups, and repeatedly submit bids based on their maximum cost-per-click (most they are willing to pay) for keywords in their account (inventory). The results of the auction process are dynamic throughout the day, and the CPC charged and the resulting ad positions may vary from time to time (auction to auction). An advertiser must decide how to allocate its total sponsored search budget across search markets (i.e., different search providers, or search versus content networks within a given search provider such as Google's AdWords versus AdSense); across temporal slots (i.e., monthly, daily budget, and promotion periods, etc.); or across ad campaigns, ad groups or individual keywords dynamically to limit the opportunity costs of running out of budget early (Zhang & Feng, 2011). As many advertisers' accounts are currently organized by products/services, the ability to track marketing objectives by campaign is limited in this setting. Further, the allocation of daily budget is done by the advertiser either at the account level or campaign level, as shown in Table 4-1.

An advertiser can choose to set the total budget or allow the budget to be automatically adjusted by Google in order to ensure that the daily budget does not run out early or that maximum performance is achieved (Google, 2015). This automatic adjustment favours click maximization which is in Google's interest and may not account for the advertiser's specific marketing objectives. Therefore, the use of this method may be ineffective for advertisers that seek other performance goals such as exposure or high search volume. In this research, I use the buying funnel framework as a basis for synchronizing the keyword segmentation and campaign organization efforts and facilitating the allocation of an advertiser's account-level budget to multiple sponsored search campaigns. Using simulation modelling and keyword advertising data of four different

Fortune500 companies as cases, I evaluate the outcome of cost-based, search volume-based, and click-based allocation heuristics against a baseline allocation.

Table 4-1 The current state of Google's AdWords account hierarchy and budget setting options

<table>
<tr><th colspan="4">ACCOUNT</th></tr>
<tr><td colspan="4">Budget
Settings</td></tr>
<tr><th colspan="2">CAMPAIGN</th><th colspan="2">CAMPAIGN</th></tr>
<tr><td colspan="2">Budget
Settings</td><td colspan="2">Budget
Settings</td></tr>
<tr><th>AD GROUP</th><th>AD GROUP</th><th>AD GROUP</th><th>AD GROUP</th></tr>
<tr><td>Ads
Keywords</td><td>Ads
Keywords</td><td>Ads
Keywords</td><td>Ads
Keywords</td></tr>
</table>

4.1 Description of Problem

In this thesis, the budget allocation problem is formulated at the search network level (e.g., AdWords account), where the objective is to determine how the search network account level budget is allocated to multiple ad campaigns. Using the buying funnel model as the basis of keyword segmentation as well as campaign or ad group organization, the performances of different allocation strategies (search volume-based, cost-based, and click-based) and their implications on different marketing objectives are investigated. Simulation modeling is used to capture the search advertising process, events, the advertiser's actions, and performance outcomes in a given time horizon. Multiple search advertising performance measures that are defined in the sponsored search literature and/or industry practice are used, along with competitive keyword data from Spyfu.com.

Characteristics of a product/service such as the stage of its development life cycle can drive the specific types of keywords that may be used in different campaigns. The product development life-cycle is a good example for illustrating why campaigns may have different types of keywords representing each stage of the buying funnel and why a budget allocation strategy is important. Consider, for example, an advertiser's AdWords account with two campaigns. One campaign, designed to introduce a new product into the market, may have a large number of awareness keywords and fewer keywords representing the other stages of the buying funnel. In this case, impressions (the typical measure of exposure and awareness in sponsored search) may be what the advertiser is more likely to maximize. While a second campaign, designed around a promotion for a discounted mature product, may have a large number of keywords representing decision or purchase stages and fewer keywords representing the other stages of the buying funnel. In this case, clicks and conversions may be the metrics the advertiser is more likely to maximize. Given this variation in campaign organization and performance prioritization, different strategies for allocating the account level budget will impact the outcomes of the aforementioned campaigns in different ways. Thus, characterising the performance variations and examining the implications on different marketing objectives provide actionable insights for advertisers to improve their resource utilization. The following simulation comprises 16 total scenarios (i.e., four experiments tested under four budget allocation scenarios, including the baseline).

4.2 Simulation Model Development

To capture the dynamics of the sponsored search environment, the simulation model is formulated. Using historical keyword advertising data from Spyfu.com, the daily search volume, CPC, and the positions of ads are projected. The notation used in the simulation model is summarized in Table 4-2.

Table 4-2 Notation

Parameter	Description
j	Campaign identifier, where $j = 1,2,..., m$.
i	Keyword identifier in campaign j, where $i = 1,2,..., n_j$.
k	Allocation strategy identifier, where $k = 1,2,3,4$; 1 represents *Baseline,* 2 *Cost-based,* 3 *Volume-based,* and 4 *Clicks-based.*
B_t	The overall advertising budget set for time period t.
C_{jkt}	The amount of advertising budget allocated to campaign j using allocation strategy k during time period t, where $\sum_{j=1}^{m} C_{jkt} = B_t$.
λ_{jkt}	The allocation proportion calculated for campaign j using allocation strategy k during time period t, where $\sum_{j=1}^{m} \lambda_{jkt} = 1$.
Q_{ijt}	A random variable with $Min(Q_{ijt}), E(Q_{ijt}), and\ Max(Q_{ijt})$, representing the number of searches of keyword i from campaign j during time period t.
CPC_{ijt}	A random variable with $Min(CPC_{ijt}), E(CPC_{ijt}), and\ Max(CPC_{ijt})$, representing the cost-per-click of keyword i from campaign j during time period t.
POS_{ijt}	A random variable with $Min(POS_{ijt}), E(POS_{ijt}), and\ Max(POS_{ijt})$, representing the ad position of keyword i from campaign j during time period t.
CTR_{ijt}	Click-through-rate of keyword i from campaign j during time period t, dependent on position POS_{ijt}. See Equation 4-2.
$\sum_{i=1}^{n_j} Q_{ijt}$	Total volume of searches from all keywords in campaign j during time period t.
$\sum_{j=1}^{m}\sum_{i=1}^{n_j} Q_{ijt}$	Total volume of searches from all keywords at the account level during time period t.

$\sum_{i=1}^{n_j} COST_{ijt}$	The total cost of all keywords in campaign j during time period t.
$\sum_{j=1}^{m}\sum_{i=1}^{n_j} COST_{ijt}$	The total cost of all keywords at the account level during time period t.
$\sum_{i=1}^{n_j} CLICKS_{ijt}$	Total clicks of all keywords in campaign j during time period t.
$\sum_{j=1}^{m}\sum_{i=1}^{n_j} CLICKS_{ijt}$	The total clicks of all keywords at the account level during time period t.

I consider a single advertiser with m campaigns and n_j keywords in each campaign j, where each campaign is defined as one of the four buying funnel segments for simplicity (i.e., m = 4 in the simulation demonstrated in this thesis). Each campaign j contains specific keywords from a stage of the buying funnel model; 1 represents *Awareness*, 2 *Research*, 3 *Decision*, and 4 *Purchase*. Furthermore, each campaign is defined to represent a single ad group. In other words, the budget hierarchy is defined from an account level to a campaign level, and from a campaign level to individual keywords, as shown with notation in Table 4-3 and in an AdWords account in Figure A-3.

Table 4-3 The account hierarchy and budget setting options with notation

ACCOUNT			
BUDGET B_t			
CAMPAIGN 1 **C1: AWARENESS**	**CAMPAIGN 2** **C2: RESEARCH**	**CAMPAIGN 3** **C3: DECISION**	**CAMPAIGN 4** **C4: PURCHASE**
C_{1kt}	C_{2kt}	C_{3kt}	C_{4kt}
Keywords $i = 1,2,..., n_1$	Keywords $i = 1,2,..., n_2$	Keywords $i = 1,2,..., n_3$	Keywords $i = 1,2,..., n_4$

The company's daily budget (B_t) is allocated to the four campaigns using the four allocation strategies as depicted in Table 4-4 (the model is generalizable to any number of campaigns). In addition, the time period t can be set at any level of granularity, such as a week, a day, or an hour. However, considering the current state of search advertising platforms, advertisers typically set a daily budget on a monthly basis, and are limited in the number of times budgets can be adjusted during a day. Therefore, to follow the practice and for ease of modeling and understanding, t represents a day, and the simulation model is run for a planning horizon of 30 days (one month). This is illustrated in Table 4-4 for each allocation strategy.

The segmentation of keywords based on the stages of the buying funnel model allows companies to create or modify different compositions of keywords (number of keywords in each stage), track and prioritize the performance of multiple marketing objectives (generate "awareness", introduce consumers conducting "research" to its product/service offerings, persuade "decisions" when comparing brands, and increase "purchase" conversions), and determine how each keyword type effects budget utilization. In addition, the segmentation of keywords using the buying funnel provides companies with the ability to identify, within the context of a marketing objective, why certain types of keywords offer different value and vary in competitiveness, and decide if higher or lower budget should be allocated or if keywords should be added or removed. is intended to illustrate the analysis and presentation of the simulation model. It shows a 4 x 4 grid where each column represents an allocation strategy and each row represents a campaign. For each column, the sum of the rows equals the account level budget (B_t).

Table 4-4 Budget Allocation Strategies

C_{jkt}: AMOUNT OF BUDGET ALLOCATED to campaign j using allocation strategy k during time period t, where $j = 1,2,3,4$; $k = 1,2,3,4$; $t = 1,2,...,30$, and BUDGET= B_t				
	ALLOCATION STRATEGY			
CAMPAIGN	*Baseline* $k = 1$	*Cost-based* $k = 2$	*Volume-based* $k = 3$	*Clicks-based* $k = 4$
C1: AWARENESS $j = 1$ (CAMPAIGN ONE)	Portion of BUDGET ($) allocated to C1 using BASELINE	Portion of BUDGET ($) allocated to C1 using COST	Portion of BUDGET ($) allocated to C1 using VOLUME	Portion of BUDGET ($) allocated to C1 using CLICKS
C2: RESEARCH $j = 2$ (CAMPAIGN TWO)	Portion of BUDGET ($) allocated to C2 using BASELINE	Portion of BUDGET ($) allocated to C2 using COST	Portion of BUDGET ($) allocated to C2 using VOLUME	Portion of BUDGET ($) allocated to C2 using CLICKS
C3: DECISION $j = 3$ (CAMPAIGN THREE)	Portion of BUDGET ($) allocated to C3 using BASELINE	Portion of BUDGET ($) allocated to C3 using COST	Portion of BUDGET ($) allocated to C3 using VOLUME	Portion of BUDGET ($) allocated to C3 using CLICKS
C4: PURCHASE $j = 4$ (CAMPAIGN FOUR)	Portion of BUDGET ($) allocated to C4 using BASELINE	Portion of BUDGET ($) allocated to C4 using COST	Portion of BUDGET ($) allocated to C4 using VOLUME	Portion of BUDGET ($) allocated to C4 using CLICKS

This thesis uses simulation modeling for three reasons. First, simulation modeling provides the analyst with the ability to recreate phenomenon and behaviour of interest for study (in this case, the sponsored search program of a single advertiser). Second, simulation provides the modeller with the flexibility to analyse multiple scenarios over time (i.e., different budget allocation strategies). As shown in Table 4-4, I conduct four experiments under four budget allocation strategies. Third, simulation allows statistics gathering through meaningful output and results that can be statistically validated (Ree, 2000). Taking this into account, it is important to ensure that the system does in fact reflect the reality and is validated through testing. The process that is captured in the simulation is shown conceptually in Figure 4-1 and Figure 4-2. I specifically used ARENA® 14 simulation software for model development.

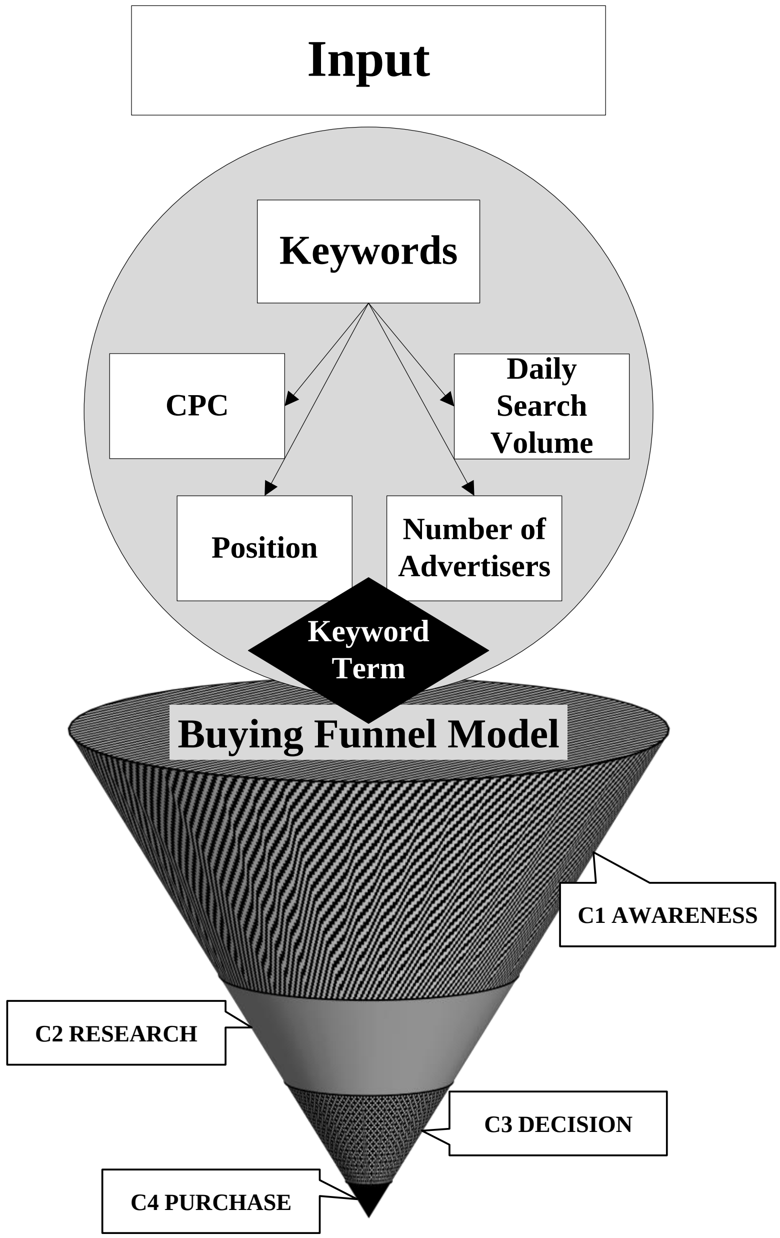

Figure 4-1 Input classification

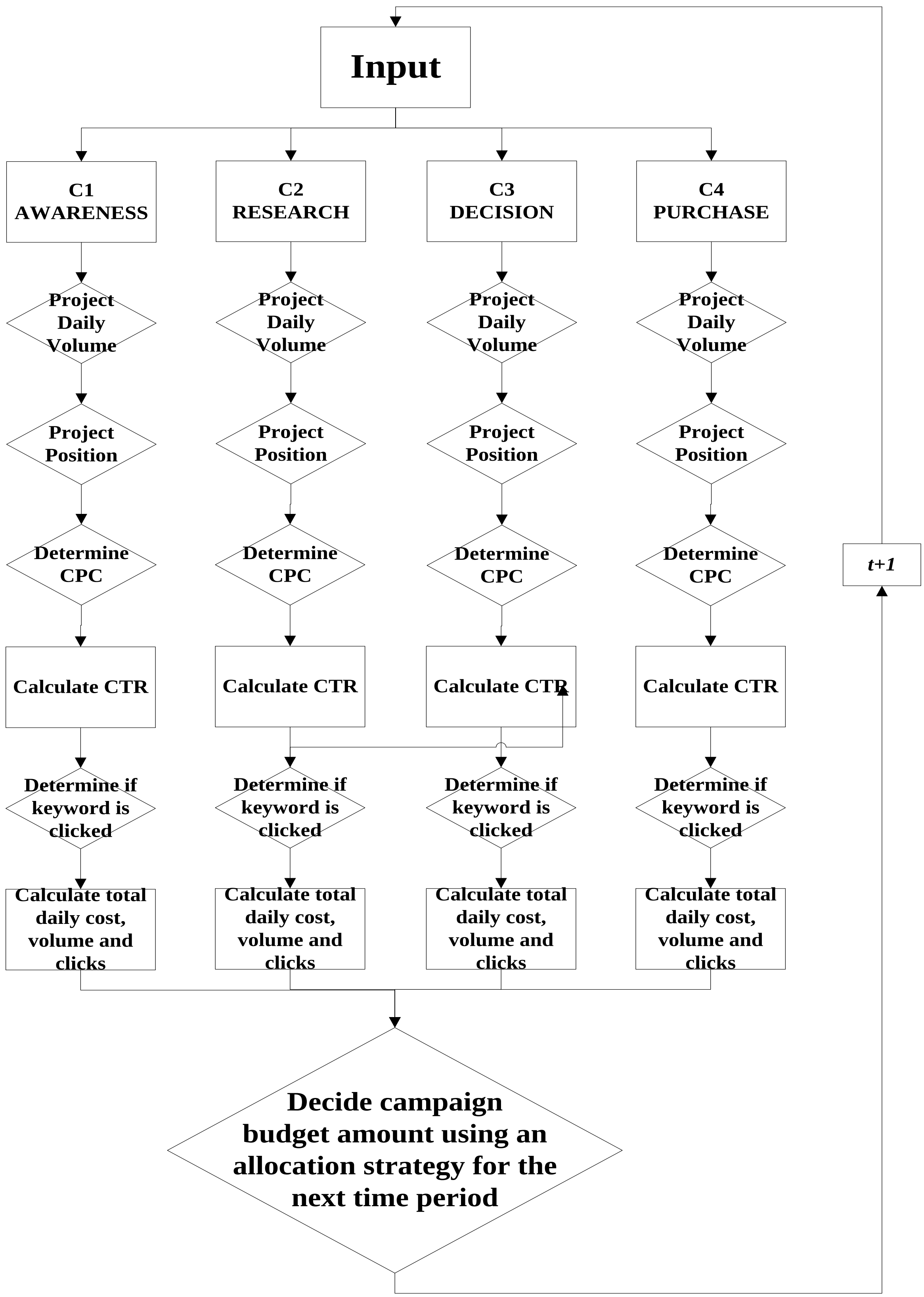

Figure 4-2 Conceptual framework of the simulation model

The amount of budget allocated to campaign j during time period t using a given allocation strategy is shown in Equation 4-1 below. The amount of budget allocated (C_{jkt}) equals the allocation proportion (λ_{jkt}) multiplied by the total account level budget (B_t). The allocation proportion (λ_{jkt}) represents the ratio of a campaign's total outcome of volume, cost, or clicks to the corresponding total outcome at the account level from $t-1$.

Equation 4-1 Amount of budget allocated to campaign j using an allocation strategy k

$$C_{jkt} = \lambda_{jkt} * B_t, \text{ for } j = 1,2,3,4;\ k = 1,2,3,4;\ \text{and } t = 1,2,...,30.$$

For the *Baseline* allocation strategy, $\lambda_{j1t} = \frac{1}{4}$, and $C_{j1t} = \frac{1}{4} * B_t$;

For the *Cost-based* allocation strategy, $\lambda_{j2t} = \frac{\sum_{i=1}^{n_j} COST_{ijt-1}}{\sum_{j=1}^{4}\sum_{i=1}^{n_j} COST_{ijt-1}}$, and;

For the *Volume-based* allocation strategy, $\lambda_{j3t} = \frac{\sum_{i=1}^{n_j} Q_{ijt-1}}{\sum_{j=1}^{4}\sum_{i=1}^{n_j} Q_{ijt-1}}$, and;

For the *Clicks-based* allocation strategy, $\lambda_{j4t} = \frac{\sum_{i=1}^{n_j} CLICKS_{ijt-1}}{\sum_{j=1}^{4}\sum_{i=1}^{n_j} CLICKS_{ijt-1}}$.

The search volumes, positions, and CPCs of each keyword are determined using a triangular distribution, with minimum, expected, and maximum values from the SpyFu data. The triangular distribution is chosen for its flexibility in approximating the historical data. Instead of relying on one single parameter such as average, the triangular distribution allows the use of the minimum, most likely, and maximum values of a given variable to approximate its distribution based on historical data. The input parameters specifically used for search volume are based on a minimum of one, and expected and maximum values of the "exact global daily search volumes" for each keyword from the SpyFu data. The position of a keyword is determined using the theoretical minimum position of one as the minimum, and expected and maximum positions based on the historical average position and the number of competing advertisers, respectively, from the SpyFu data. To determine CPC, the "broad cost per click", "phrase cost per click" and "exact cost per click" values from the SpyFu data are ordered from lowest to highest and used to represent minimum, expected, and maximum values in the triangular distribution. According to previous empirical studies, the top ranked position generates the largest CTR. The CTR for positions below the first are monotone and decreasing in value (Chitika, 2013; Ghose & Yang, 2009). Considering this, and to capture the decay in clicks as position increases, CTR_{ijt} is formulated in the simulation model using the Zipf distribution (Balakrishnana & Kambhampati, 2008; Naldi et al., 2010; Regelson & Fain, 2006; and Xie and O'Hallaron, 2002). Equation 4-2 shows this formulation which represents the probability a user clicks the ad of keyword i from campaign j during time period t. The effects of different rate of decay values $\alpha = 1, 1.5, or\ 2$ are shown in Figure 4-3. The higher the value of α, the faster the decay.

Equation 4-2 Click through rate of keyword i from campaign j during time period t

$CTR_{ijt} = \frac{1}{(POS_{ijt})^{\alpha}}$, where the rate of decay of CTR_{ijt} is tested at $\alpha = 1, 1.5, or\ 2$.

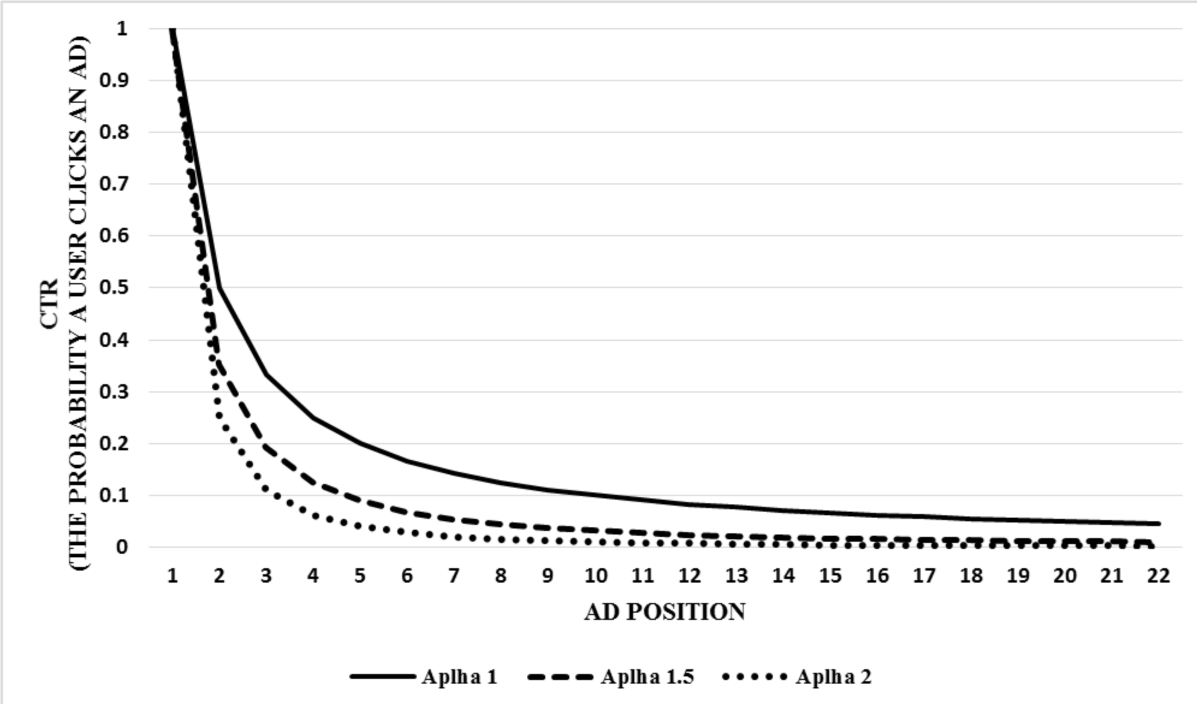

Figure 4-3 The rate of decay of CTR using different α values

The performance implications of each allocation strategy is assessed based on two direct measures, budget overage/underage (DEV_{jkt}) and budget absolute deviation ($ABSDEV_{jkt}$). First, I use the budget overage/underage at the campaign level. According to keyword auction rules, once the daily budget runs out, ads are not shown for the remainder of the day. Thus, advertisers miss opportunities due to the shortage of budget and DEV_{jkt} would be a negative value (underage). Alternatively, if the daily budget is left unutilized, it represents poor budget allocation or ineffective campaign or keyword performance, and DEV_{jkt} would be a positive value (overage). In particular, at campaign levels, poor allocation is characterized by budget overage in some campaigns and budget underage in other campaigns. Thus, I track this performance for the different allocation strategies. Equation 4-3 measures this performance at the campaign level, which is calculated by subtracting the total cost spent for campaign j from the amount of budget allocated to it.

Equation 4-3 Overage/underage in dollars of campaign j and strategy k during time period t

$$DEV_{jkt} = C_{jkt} - \sum_{i=1}^{n_j} COST_{ijt} \text{ , for } j = 1,2,3,4 \text{ , } k = 1,2,3,4 \text{ , and } t = 1,2,...,30 \text{ .}$$

To show the relative efficiency of each allocation strategy, I also calculate the percent overage/underage of each campaign at the end of each time period. This percentage is calculated by dividing the overage/underage of campaign j by the amount of budget allocated to it. Finally, the overall efficiency of each strategy k and campaign j for all 30 days is calculated by taking the average of the percent overage/underage across the planning horizon. An example is shown graphically in Figure 4-4.

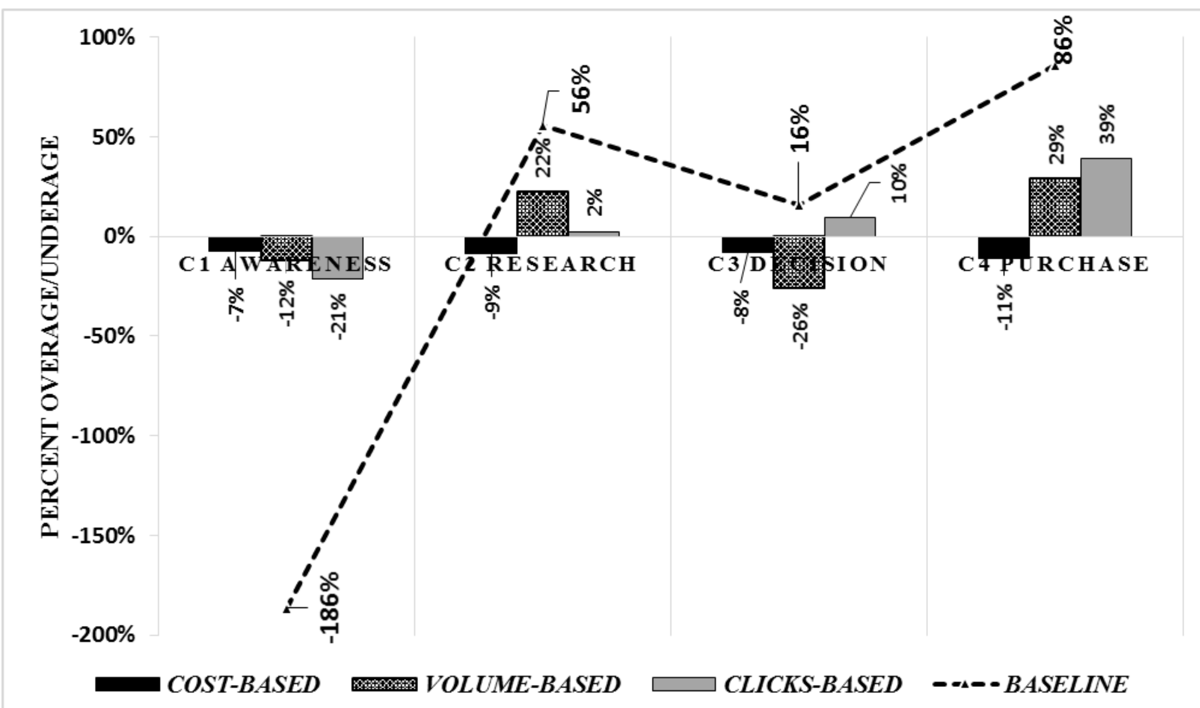

Figure 4-4 Example of average percent *overage/underage* by segment

Furthermore, to assess the variance in dollars from the allocated budgets at the campaign level, I report the absolute value of the overage/underage for each campaign and allocation strategy, as shown in Equation 4-4.

Equation 4-4 Absolute deviation in dollars of campaign j and strategy k during time period t

$$ABSDEV_{jkt} = \left|DEV_{jkt}\right|, \text{ for } j = 1,2,3,4,\; k = 1,2,3,4, \text{ and } t = 1,2,...,30.$$

To assess the impact of each allocation strategy across the duration of the planning horizon (i.e., 30 days), I report the average of the absolute deviation for 30 days. Furthermore, I calculate the total overage/underage across the entire account. This is simply the sum of the average absolute deviations from all campaigns which is used to compare the overall efficiency of the allocation strategies as shown in Figure 4-5.

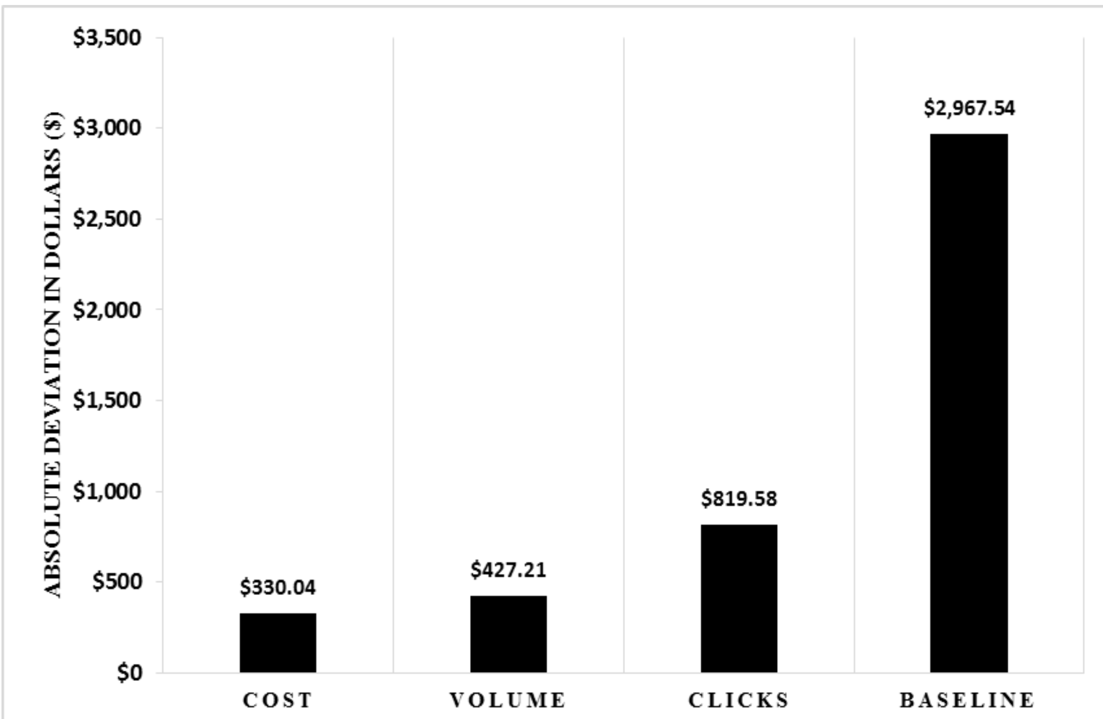

Figure 4-5 Example of total average absolute deviation

Chapter 5 Experimental Setting

5.1 Data

Before selecting the sample cases used for the simulation, keyword advertising data from 76 companies spanning 16 industries, as defined by Bloomberg (2015), was first collected during the month of May 2015 from Spyfu.com, a leading provider of keyword research technology and competitive intelligence to search engine advertisers (Ayanso & Karimi, 2015) (See Table A-1 in the appendix). Many major companies use SpyFu, and it has been endorsed by major media outlets such as Forbes, The Washington Post, and The Wall Street Journal (SpyFu, 2015). Bloomberg (2015) lists 49 distinct industries. Many of these industries (e.g., mining, Aerospace, or Defence) do not use sponsored search advertising, thus they were not chosen for comparisons. Of the 49 industries, 16 were chosen (shown in Table 5-2) for cross-industry comparisons and for selecting sample companies for the simulation and detailed analysis. The files obtained from Spyfu.com for each company contain a domain's keywords, the URL it is linked to, its position, exact local and exact global daily searches, broad, phrase, and exact cost per click, clicks per day, and cost per day values. Full description of each measure is given in Table A-2 in the appendix.

Google provides three different keyword match types for advertisers to select from: "Broad" (the default setting in AdWords and the widest casting net), "Phrase", and "Exact" (the most specific and smallest casting net). Each match type has an impact on the number of similar keywords an advertiser's ad can possibly be linked to for a single keyword bid in a campaign. This directly impacts the average CPC due to varying CPCs for every variation of a keyword. Table 4-1 shows Google's definition of each match type with examples. All tree match types are used in this thesis.

Table 5-1 Match type definitions according to Google

Match type	Special symbol	Example keyword	Ads may show on searches that	Example searches
Broad match	none	women's hats	include misspellings, synonyms, related searches, and other relevant variations	buy ladies hats
Broad match modifier	+keyword	+women's +hats	contain the modified term (or close variations, but not synonyms), in any order	hats for women
Phrase match	"keyword"	"women's hats"	are a phrase, and close variations of that phrase	buy women's hats
Exact match	[keyword]	[women's hats]	are an exact term and close variations of that exact term	women's hats
Negative match	-keyword	-women	are searches without the term	baseball hats

5.2 Data Pre-processing

The keyword advertising data was initially pre-processed to ensure the quality of the input data. Of the 76 companies, five were found to have corrupted records that were removed before aggregate statistics were obtained. The remaining records represent the output from Google's keyword planner tool with an average target position of 2. Thus, values of zero for measures of cost per click and search volume were acceptable. However, prior to segmenting the keywords into the buying funnel, additional records were removed. Any record with global search volumes of less than one or broad cost per click values equal to zero was removed. There were no values below zero for broad cost per click. These records were removed due to the simulation model's input data specifications. Specifically, as global search volume is used as input for the number of searches in the model and a searcher is defined as an entity, a value below one or non-integer value is invalid. Moreover, using keyword records with CPC values of zero does not make sense. Thus, the broad cost per click variable must be greater than zero. Once the records were segmented into the buying funnel, the distributions of global exact volume and broad cost per click were visualized by segment, and any outliers were removed. The CPC values by match type do not follow a sequential pattern where Broad CPCs can be said to be less than or greater than Phrase or Exact

CPCs consistently. Therefore, the results were ordered sequentially by actual values instead of match type in order to set minimum CPC, expected CPC, and maximum CPC values for the triangular distribution.

5.3 Sampling

In order to examine detailed profiles of the companies in each industry, the top five companies in terms of market share from each of the 16 industries were used if they were found to be actively advertising on Google's AdWords platform (SpyFu had competitive keyword data available). This resulted in some industries with fewer than five observations. Additionally, in some cases, Bloomberg lists conglomerates and parent companies (e.g., ACCOR SA) that do not sell products or services. In the latter case, if a subsidiary company was found to be advertising, it is used to represent the parent company.

Table 5-2 Sixteen industries and the top companies by market share

INDUSTRY	COMPANY NAME	URL
Airlines	United Continental Holdings	United.com
Apparel, Footwear, Acc Design	Nike Inc.	nike.com
Automobile OEM	Toyota Motor Corp	toyota.com
Banking	Bank of America Corp	bankofamerica.com
Beverages	Coca-Cola Co/The	coca-colacompany.com
Computer Hardware	Lenovo Group Ltd	lenovo.com
Department Stores	Macy's Inc	macys.com
IT Services	International Business Machines	ibm.com
Internet Media	Google Inc	google.com
Life Insurance	Prudential Financial Inc	prudential.com
Lodging	Accor SA	novotel.com
Mass Merchants	Wal-Mart Stores Inc	walmart.com
Restaurants	Starbucks Corp	starbucks.com
Retail Discretionary	Home Depot Inc/The	homedepot.com
Software	Microsoft Corp	microsoft.com
Telecom Carriers	AT&T Inc	att.com

After examining the profiles of the companies in each industry, four industries were selected to represent the final sample. The four industries selected for keyword classification and for the simulation model are Airlines, Computer Hardware, Department Stores, and Life Insurance. Each industry was selected based on its distinct product/service offerings, relevance to searchers (most searchers are likely to have purchased from one, if not all the industries), and different expected advertising goal orientations as defined by the buying funnel. This is done to demonstrate the applicability of different budget allocation methods in different industries. For example, prior work has investigated insurance keywords (Goldfarb & Tucker, 2011b) and has shown that the keywords are mostly in the "awareness" and "research" categories of the buying funnel. Computer hardware was chosen because of the high possibility of "decision" and "purchase" oriented keywords related to specific product model names and numbers and the high volume of cross-product or cross-model comparisons. Department stores were chosen because of the broad selection of products they carry and their large brick and mortar presence. Additionally, given the amount and variety of products department stores carry (from convenience goods to luxury), their keyword portfolios are more likely to form the buying funnel. Finally, the airlines industry was chosen because most sales occur online and high level of competitions are expected depending on the markets or locations. The top company by market share is selected from each of these four industries. These companies are Lenovo from the Computer and Hardware industry, Macy's from Department Store industry, Prudential from the Life Insurance industry, and United Continental Holdings from the Airline industry. Although these companies have top market share and financial performance, inefficiencies in keyword advertising may still occur. However, these companies provide richer keyword data pool to demonstrate the proposed budget allocation strategies. Being the top performing companies in their industries, competing companies would try to emulate similar business and marketing strategies, manufacture or sell similar products, and therefore bid on subsets of similar keywords. This competition provides a more appropriate setting for comparisons and further data analyses.

For each company, a random sample of 1000 keywords were selected and classified into the buying funnel model using the keyword classification schema proposed by Jansen and Schuster's (2011) as shown in Table 5-3 below. Example keywords by each stage of the buying funnel are shown in Table 5-4 for each of the four companies.

Table 5-3 Classification Schema

Awareness:
▪ Does not contain a brand name ▪ Could contain partial product name/type ▪ Could contain problem to be solved ▪ Could contain, gift certificate amount, generic company special, or generic company code
Research
▪ Does not contain brand name ▪ Contains specific product
Decision
▪ Contains specific product and partial brand name ▪ Does not contain full brand/company
Purchase
▪ Contains specific product and full brand name/company

Table 5-4 Example keywords classified according to the buying funnel

LENOVO			
AWARENESS	**RESEARCH**	**DECISION**	**PURCHASE**
laptops sale	discount laptops for students	battery lenovo	lenovo ideapad y530 battery
labtop	thin laptop computers	thinkpad x series	lenovo thinkpad x120e
laptop best	best high performance laptops	lenovo charger	lenovo ac adapter 65w 20v
MACY'S			
AWARENESS	**RESEARCH**	**DECISION**	**PURCHASE**
elegant plates	top rated rice cooker	melior coffee press	krups b100 beertender
men gold watch	24k gold watch	mens diesel watch	bulova men s marine star watch
cologne fragrance	best deodorant and antiperspirant for women	mori perfume	christian dior addict
PRUDENTIAL			
AWARENESS	**RESEARCH**	**DECISION**	**PURCHASE**
401 retirement	401k direct rollover	prudential investment	bank of america wealth management banking
whats an annuity	pension and annuity income	ing annuties	the hartford variable annuity
life insurance search	affordable term life	life insurance fidelity	suze orman whole life insurance
UNITED CONTINENTAL HOLDINGS			
AWARENESS	**RESEARCH**	**DECISION**	**PURCHASE**
fly to florida	cheap florida airfare	disney florida package	disney world orlando florida tickets
is travel to cancun safe	cheap flight to cancun	barcelo los cabos resort	riu caribe all inclusive cancun
one way airline ticket	airfares hawaii	hawaiian airlines international flights	hawaiian airline reservation

A summary of the input data and the final count results of the classification by stage are shown in Figure 5-1, where each campaign represents a stage of the buying funnel. The campaigns are ordered sequentially according to the stages of the buying funnel. C1 represents “awareness” keywords, C2, “research”, C3, “decision”, and C4, “purchase” keywords.

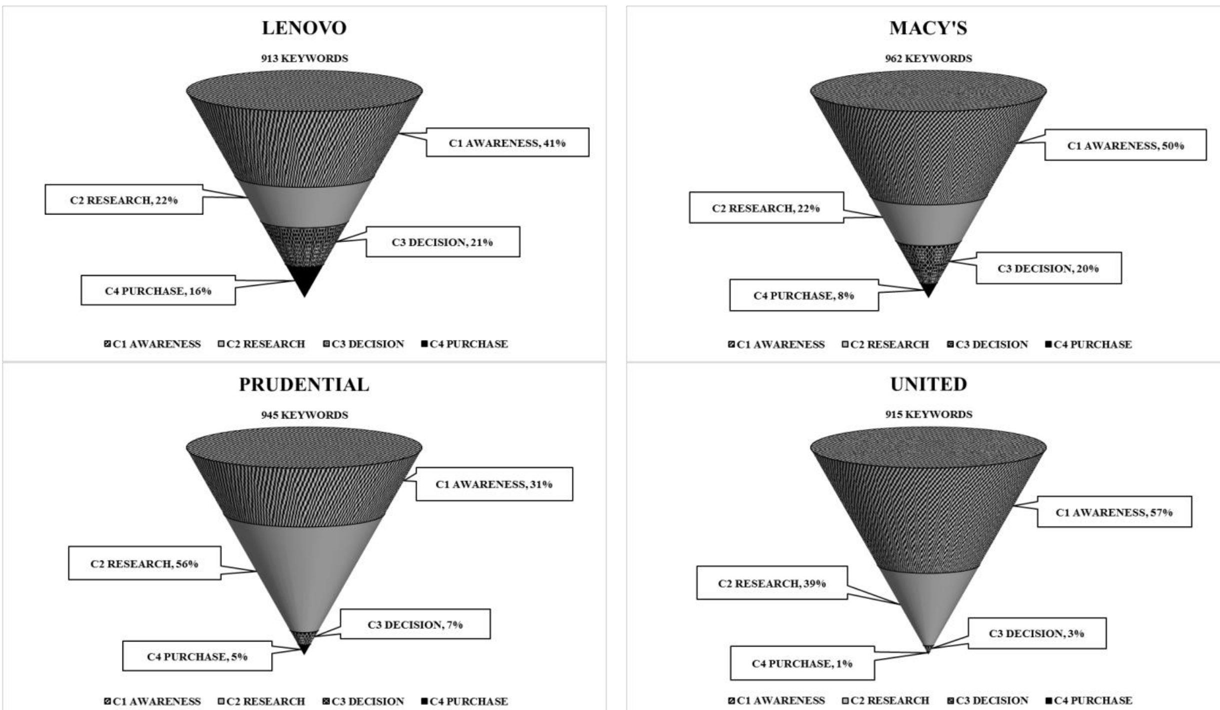

Figure 5-1 Keyword counts by campaign for each company

With the exception of Prudential, where there is more “research” keywords than any other, each company’s number of keywords follow the funnel shape, where most keywords are “awareness”, and the number of keywords diminishing from “awareness” to “research”, “research” to “decision”, and from “decision” to “purchase”. Nevertheless, the funnel shapes vary for these example companies. Macy’s has a near perfect funnel shape. Lenovo’s funnel is the second most balanced with over one third of the keywords in the latter stages of the funnel and almost similar number of “research” and “decision” keywords. United Continental Holdings has a heavily skewed funnel with 96 percent of the keywords belonging to either “awareness” (57%) or “research” (39%).

Table A-3 in the appendix presents the descriptive statistics of the keyword data for the four example companies. Lenovo has the highest average volume at 18.71, but the lowest average position and average number of advertisers (i.e., indicator of the degree of keyword competition), with values of 5.19 and 11.58, respectively. This is in contrast to United Continental Holdings, whose average volume is the lowest at 7.51; average position is the second highest at 7.53; and average number of advertisers is the highest at 17.09. Prudential has the highest average CPCs. Summary statistics are also shown by campaign/segment for average volume in Figure 5-2, CPC in Figure 5-3, position in Figure 5-4, number of advertisers in Figure 5-5, and combined results in Table A-4 in the appendix.

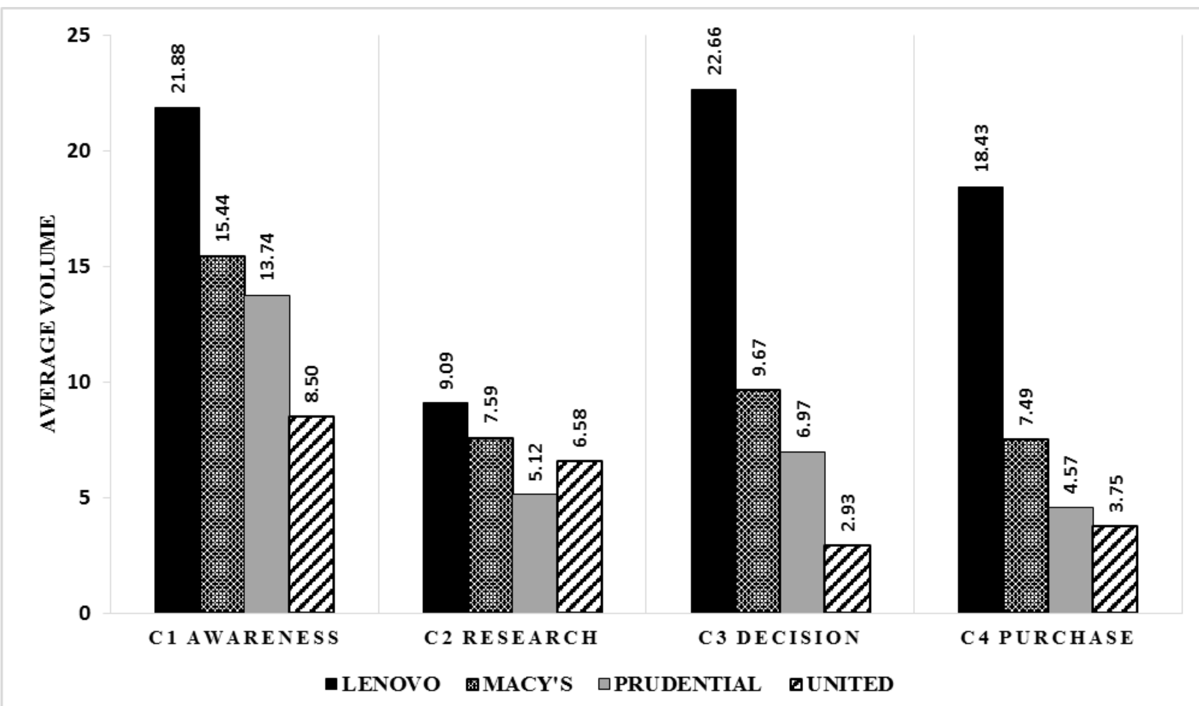

Figure 5-2 Average volume by segment

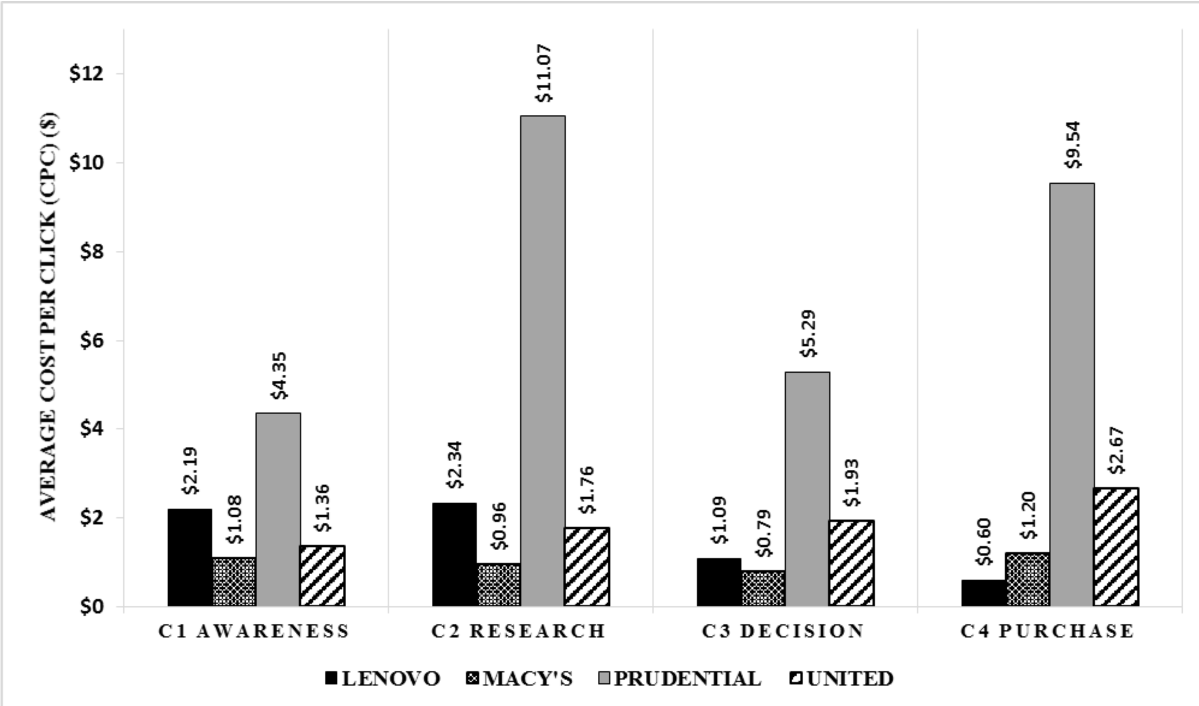

Figure 5-3 Average cost per click by segment

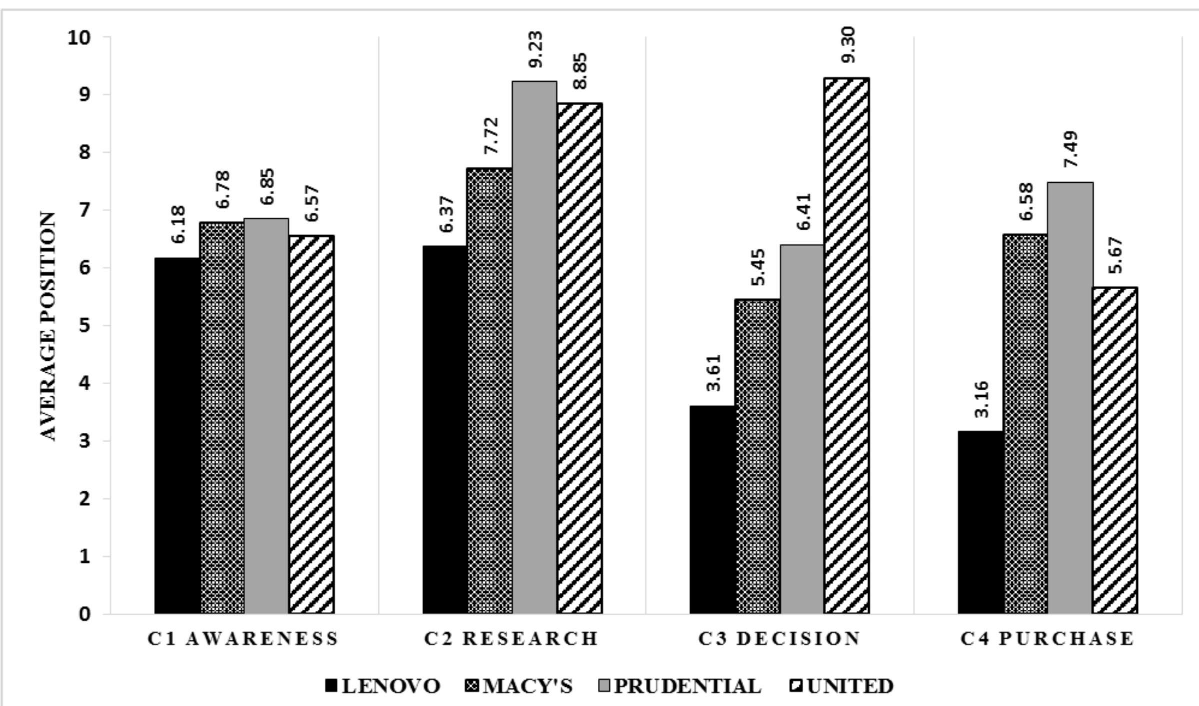

Figure 5-4 Average position by segment

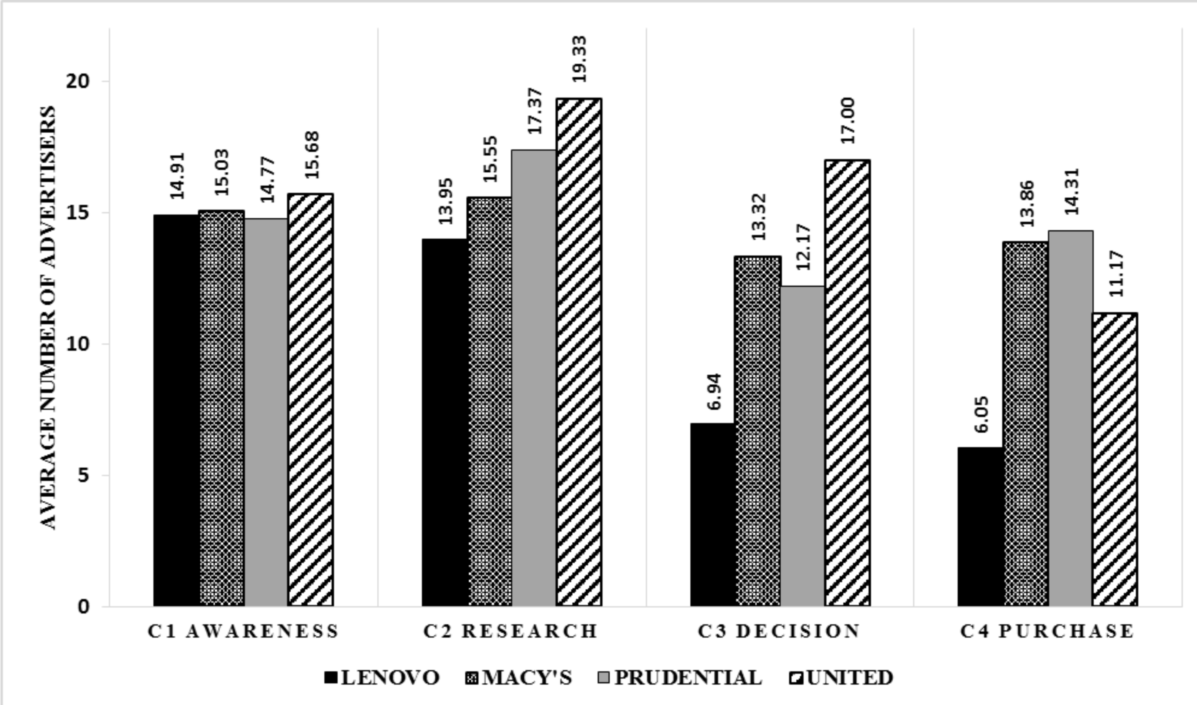

Figure 5-5 Average number of advertisers by segment

5.4 Performance Measures

I conducted four simulation experiments, one for each company (Lenovo, Macy's, Prudential, and United Continental Holdings). Prior to the experiments, the companies' keywords were segmented into one of four campaigns, with each campaign representing a stage of the buying funnel for simplicity. To demonstrate the benefits of the allocation strategies, I compare them with a baseline strategy. The *Baseline* allocation strategy is a commonly used strategy in practice and is also used in prior literature for similar purpose (Yang et al., 2012; Zhang et al., 2012; Zhang et al., 2014). For example, in Zhang et al. (2012), the advertising budget (search market or account level) is fixed and spread equally to all campaigns and is not changed from day to day. In practice, advertisers employ this method due to its simplicity, or possibly from a lack of decision support tools to apply more systematic allocation methods. The second strategy is *Cost-based* where the amount of advertising budget allocated to a campaign is determined by that campaign's cost in proportion to the total cost spent for all campaigns during the previous time period. The third strategy is *Volume-Based*. The advertising budget is allocated to campaigns based on the amount of volume a campaign generated in proportion to the total volume generated across all campaigns during the previous time period. Finally, the fourth allocation strategy is *Clicks-based*, and the advertising budget is allocated to a campaign based on the campaign's click performance in proportion to the total clicks for all campaigns during the previous time period.

In order to understand the performance of an allocation strategy, the actual costs spent against the allocated budget are tracked and the budget *overage* or *underage* are computed for each campaign during each time period (see Equation 4.3 and Equation 4.4). Corresponding results in percent overage/underage are also reported. For each campaign, an average absolute *overage/underage* value is computed for a 30-day planning horizon. In addition, to capture the overall performance of an allocation strategy, the sum of the average absolute *overage/underage* values of the four campaigns is used. This represents the overall performance of each allocation strategy, which is then compared to the performance using the *Baseline* strategy. Examples of *overage/underage* values in dollars for the four campaigns over a 30-day planning horizon are shown in Figure 5-6 and the corresponding results in percent *overage/underage* are shown in Figure 5-7.

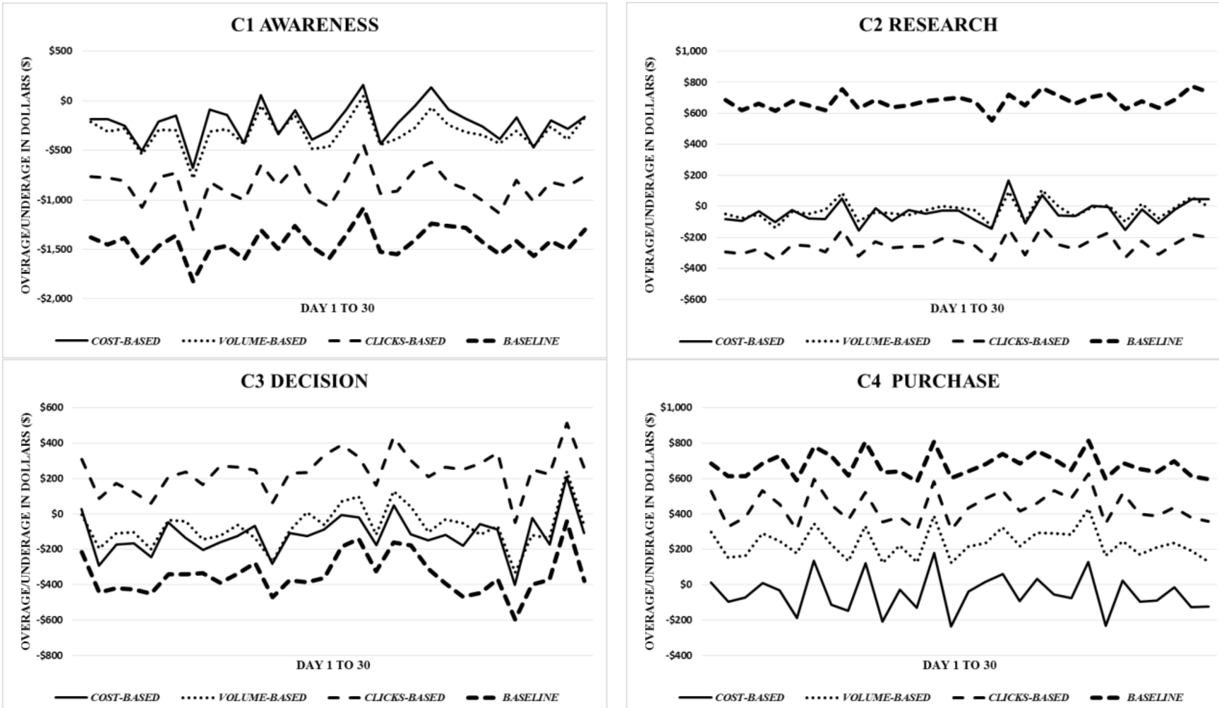

Figure 5-6 Example o*verage/underage* in dollars ($) over a 30 day period

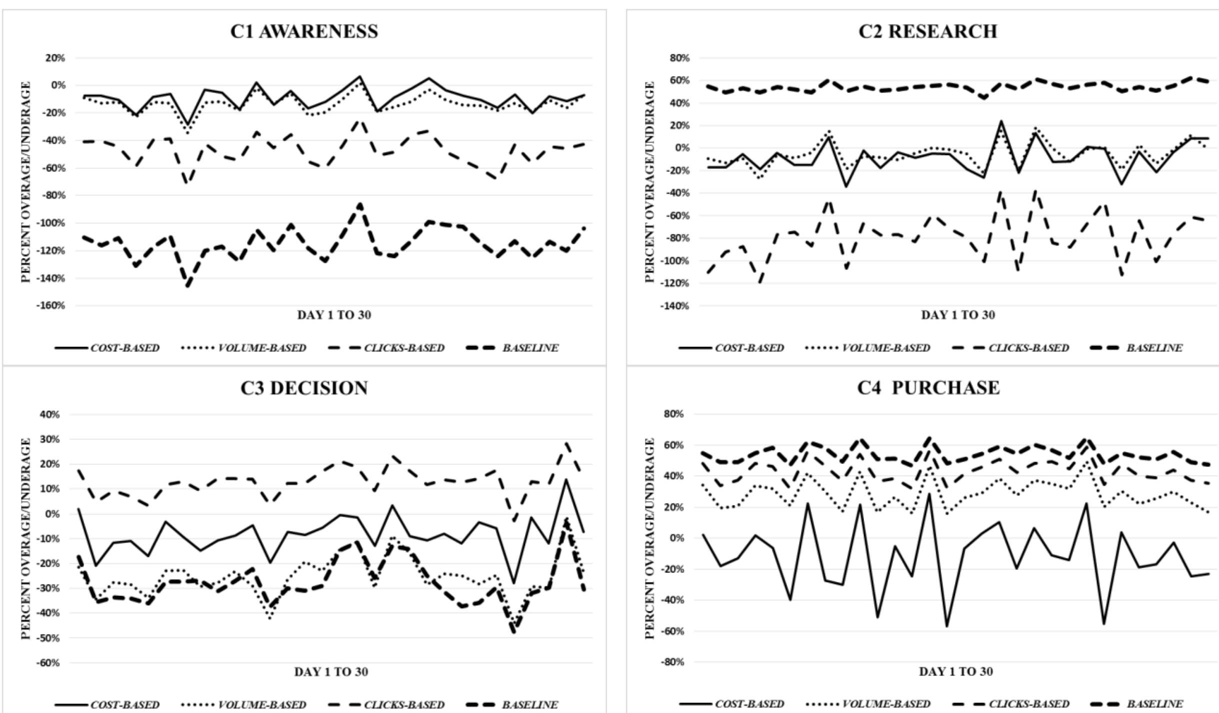

Figure 5-7 Example percent *overage/underage* over a 30 day period

Performance is defined as the total average absolute deviation of an allocation strategy, where the lower the value (closer to zero), the better the strategy. These results are used in conjunction with the graphical representation of the percent *overage/underage* by campaign for each allocation strategy compared to the *Baseline*. A company's estimated budget value is determined by multiplying the broad clicks per time period by the broad CPC and dividing this number by the number of keywords in the file. This value is then multiplied by the number of keywords included in the final simulation input file. Equation 5-1 shows this estimation.

Equation 5-1 Estimated budget

$$\frac{\sum \text{Broad Clicks} \times \text{Broad CPC}}{\text{Total \# of Keywords}} \times \text{\# of Keywords in simulation}$$

Chapter 6 Simulation Results

For each sample company, I profile and compare the outcome of the simulation experiments under each allocation strategy with that of the baseline strategy. First, I present the outcome of total cost, volume, clicks, and the number of keywords, and each campaign's profile in each. Second, I show performance as the total average absolute deviation by allocation strategy and the average percentage of *overage/underage* by campaign.

6.1 Simulation Cases

6.1.1 Case-1: Lenovo

The classification of Lenovo's keywords in the buying funnel presents some interesting outcomes. As expected, relative to the other companies, there is more focus on the latter stages of the buying funnel (37%). It is a well-established brand, and the products it carries vary across the product development life cycle with new models of laptops and tablets (sometimes entire product categories) introduced on an annual basis similar to the automotive industry. Due to the expensive nature of its products, the flow-through stages of the buying funnel are expected to be slower, as laptops are longer term purchase decisions. For this reason as well as because of the high tech and oligopolistic nature of the Computer hardware industry, Lenovo is present in all stages of the buying funnel. The earlier stages of keywords, where there is more competition (see Figure 5-5), focus on the introduction and explanation of new types of products, in addition to promotions such as "cheapest laptops online" and "cheap laptops for college students. The decision stage is focussed on the brand, "laptops lenovo", product categories, "lenovo thinkpad reviews", and its accessories, "lenovo batteries". Finally, in the "purchase" stage, the majority of the keywords are for specific models such as, "lenovo thinkpad x120e".

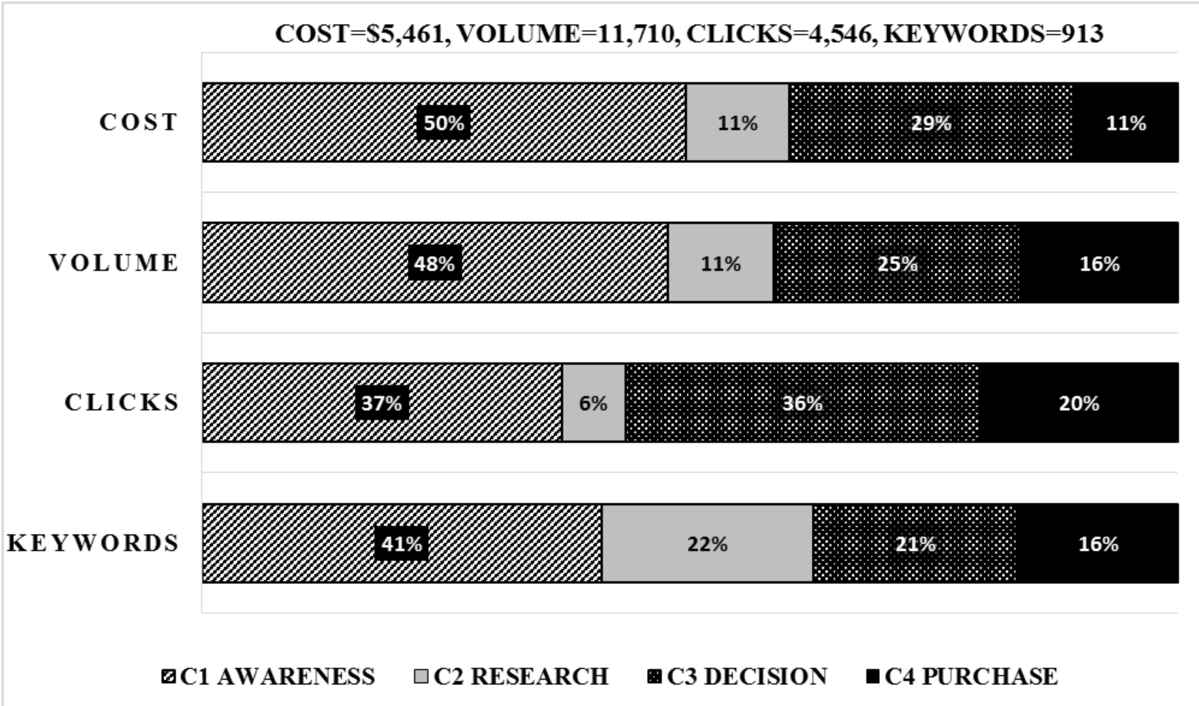

Figure 6-1 Lenovo's average contribution of cost, volume, clicks and keywords by segment

The percent contribution of each segment to total cost, volume, clicks, and the number of keywords, along with the corresponding totals, is shown for Lenovo in Figure 6-1. From the figure, the following can be observed: (i) the outcomes of each measure do not match the percentage contribution of the number of keywords. As expected, due to the varying CPCs and average positions, some campaigns contribute more or less to a measure's total, (ii) of particular note is Lenovo's "decision" and "purchase" keywords represent only 37 percent of all keywords, 40 percent of the cost, and 41 percent of the volume, but generate a much higher percentage of clicks (57%). This is caused by lower average CPCs and positions Lenovo attains in these two segments, (iii) this is in contrast to "awareness" keywords which represent 40 percent of the keywords, but generate a higher percentage of cost (50%) and a lower percentage of clicks (37%), (iv) "Research" keywords are the poorest contributors to total clicks with only 6 percent. This is caused by higher average CPCs and Positions, in addition to lower average volumes in this segment. With little contribution to overall cost, volume, or clicks, more research keywords are required to restore the funnel shape. However, the disparity in percent contribution of cost (11%) and clicks (6%) would suggest that in terms of click generation, “research” keywords are poor investment, (v) on the other hand, “purchase” keywords with low cost contribution (11%) and high click contribution (20%) would be a good investment. (vi) for a click maximization strategy, Lenovo should invest in more “decision” and “purchase” keywords. Surprisingly, in terms of exposure, “purchase” keywords provide the most volume (16%), while contributing the smallest cost (11%).

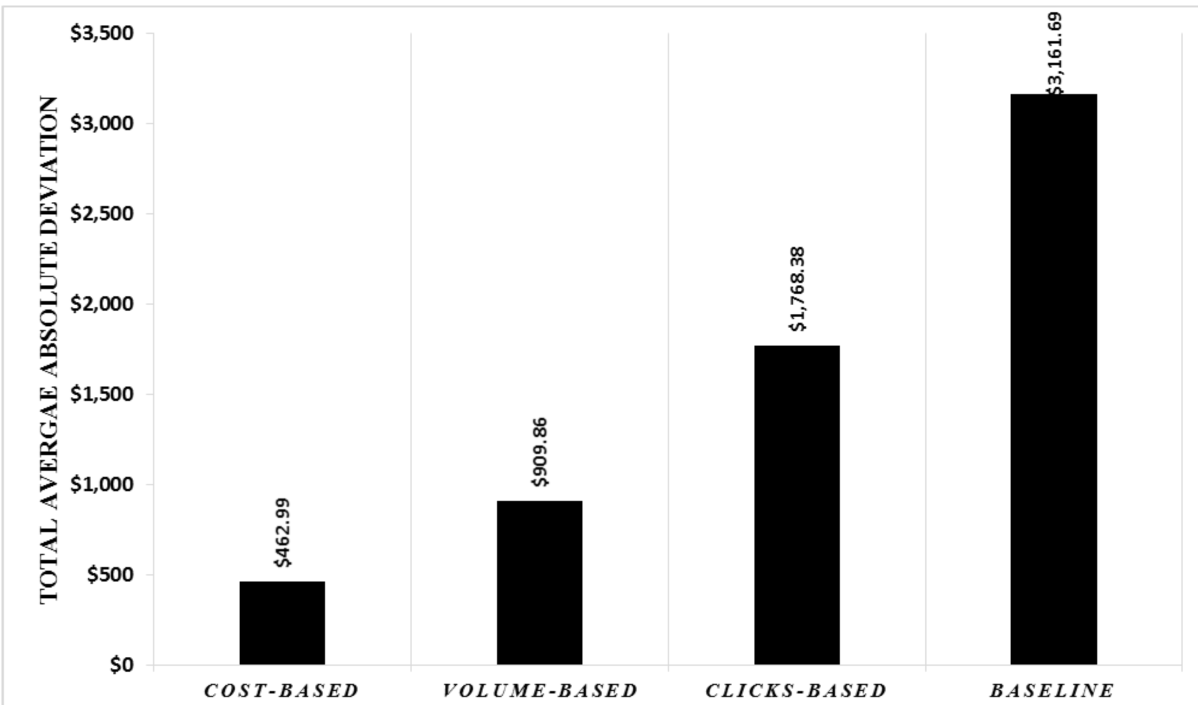

Figure 6-2 Lenovo’s total average absolute deviation in dollars ($)

The total average absolute deviation of all four strategies is shown for Lenovo in Figure 6-2. From the figure, the following can be observed: (i) all allocation strategies significantly outperform the *Baseline*. This suggests that by using the buying funnel and any of these simple allocation strategies, Lenovo can achieve better budget utilization, (ii) the *Cost-based* strategy performed the best overall and is nearly seven times better than the baseline, (iii) the *Clicks-based* strategy performed the poorest of all the proposed strategies, but is still twice as efficient as the *Baseline*.

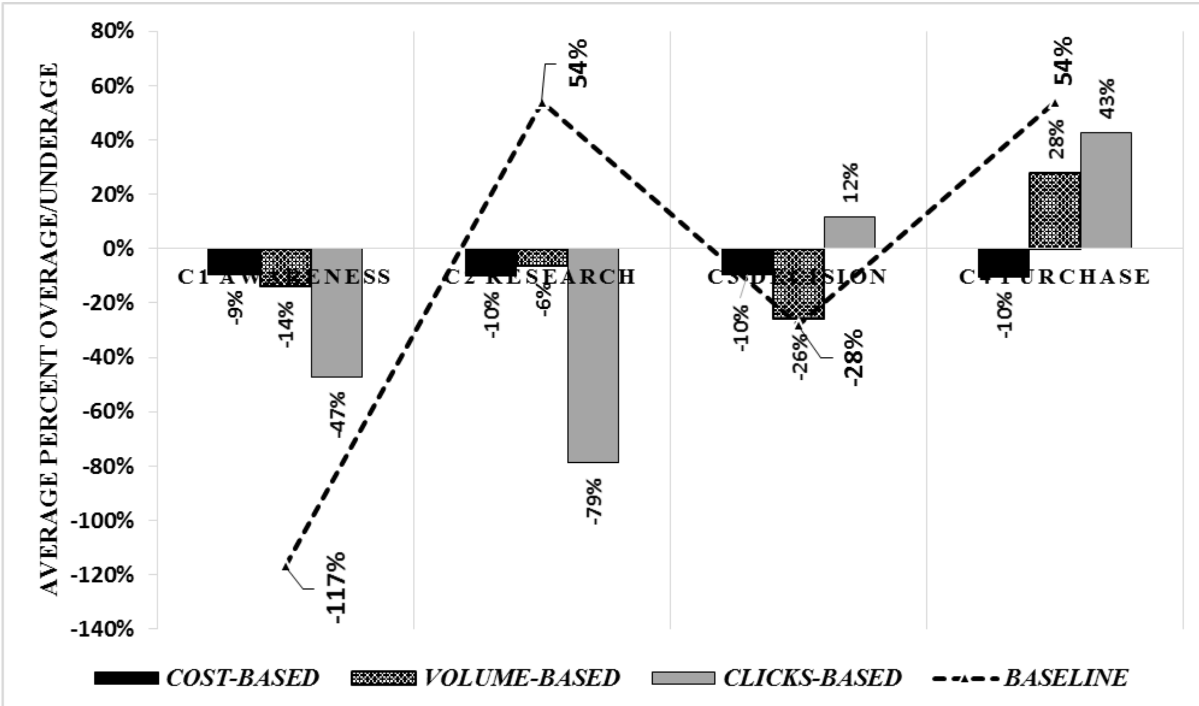

Figure 6-3 Lenovo's average percent overage/underage by segment

Lenovo's average percent *overage/underage* by segment is shown in Figure 6-3. From the figure, the following can be observed: (i) in all cases, except campaign two using the *Clicks-based* strategy, the proposed allocation strategies outperform the baseline, (ii) with the current keywords, more budget is needed for campaign one for all allocation strategies, (iii) the results clearly show that the *Cost-based* strategy is best able to allocate the funds and that about 10 percent increase in budget is required to not exceed budget across all campaigns, (iv) due to the disparity between percent of total cost (11%) and percent of total clicks (6%) for "research" keywords, there is a large misallocation in campaign two using the *Clicks-based* strategy.

6.1.2 Case-2: Macy's

As expected, the classification of Macy's keywords resulted in close to the funnel shape. As a large department store, Macy's carries a vast array of products varying across the product development life cycle and in purchase intensity (e.g., "men's footwear" or "womens underwear for men" and "14k pearl necklace"). As a consequence of this extreme variety, Macy's invests heavily in sponsored search with a monthly budget of 2.33 million across 678,000 keywords. For the majority of its products, search users likely transition through the cognitive stages of the buying funnel rapidly. Thus, making it important that Macy's is present at each stage to inform the searcher that it provides or carries the information or products they seek. Macy's "awareness" keywords reflect the necessity of the broad net that it must cast, with keywords such as "holiday top", "room couch", and "cool mens clothes". This logic follows through each stage, with "research" keywords such as, "sleeveless tops for women", "chrome wall mirror", and "professional makeup brush", "decision" keywords such as, "jessica simpson denim", "adidas polo shirts", and "clayton marcus sectional", and finally, "purchase" keywords such as "conair illumina lighted makeup mirror", "lebron james 3 shoes", and "krups b100 beertender".

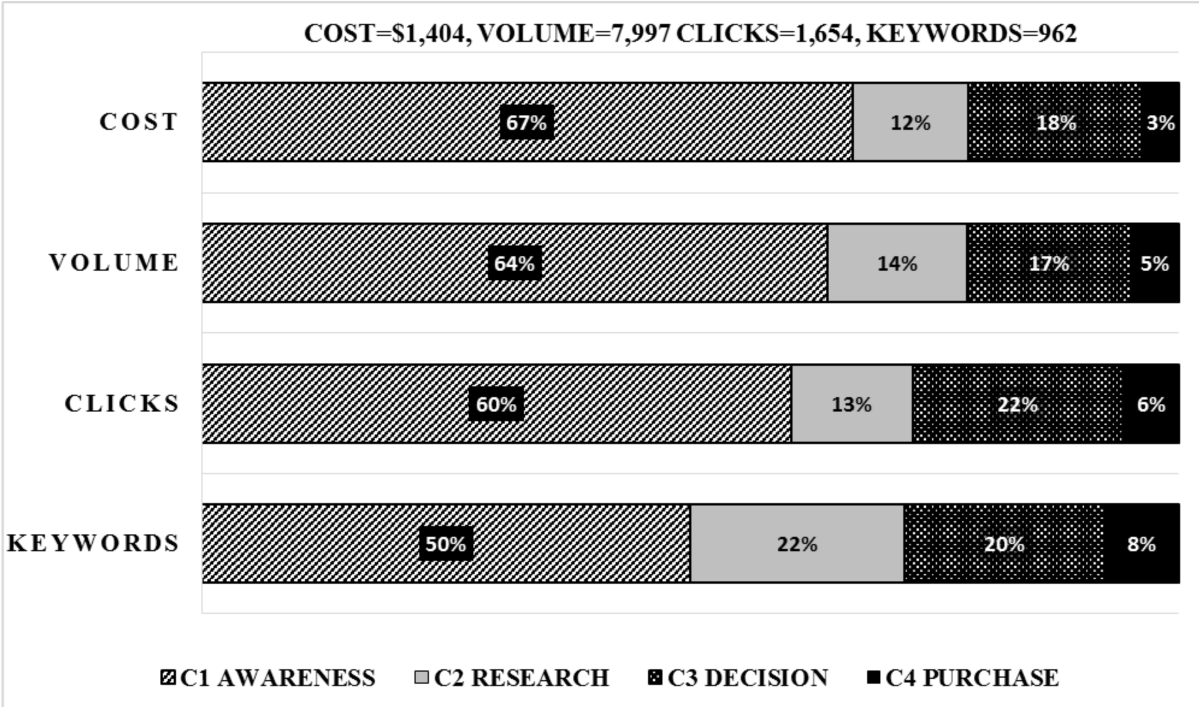

Figure 6-4 Macy's average contribution of cost, volume, clicks and keywords by segment

The percent contribution of each segment to total cost, volume, clicks, and the number of keywords, along with the corresponding totals, is shown for Macy's in Figure 6-4. From the figure, the following can be observed: (i) "Awareness" keywords contribute the most across all measures. However, the cost contribution (67%) is larger than the contributions to volume (64%) and clicks (60%), suggesting that "awareness" keywords are a poor investment in terms of volume or clicks per dollar relative to the other segments, (ii) the most productive keywords in terms of clicks and volume are "purchase" keywords. While only contributing three percent of the total cost, "purchase" keywords account for nearly double that, for both volume (5%) and clicks (6%), (iii) "Decision" keywords are the only keywords that contribute lower costs (18%), but provide more clicks (22%) than its percentage of keywords (20%).

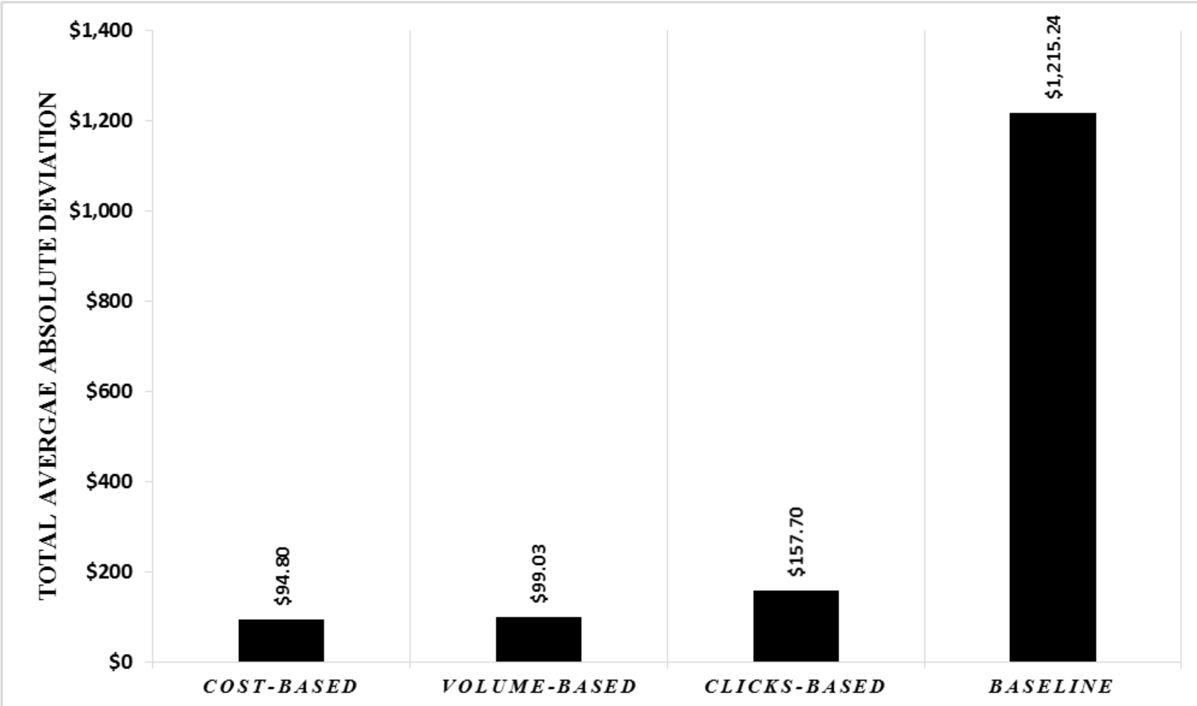

Figure 6-5 Macy's total average absolute deviation in dollars ($)

The total average absolute deviation of all four allocation strategies is shown for Macy's in Figure 6-5. From the figure, the following can be observed: (i) all allocation strategies significantly outperform the *Baseline*. This suggests that by using the buying funnel and any of these simple allocation strategies, Macy's can reduce the opportunity costs of overage/underage resulting from poor allocation, (ii) the *Cost-based* and *Volume-based* strategies performed the best overall and the *Clicks-based* strategy performed slightly worse, (iii) the *Baseline* strategy has the poorest performance, and at $1,215.24 total average absolute deviation, it is over 12 times worse than the *Cost-based* and *Volume-based* strategies.

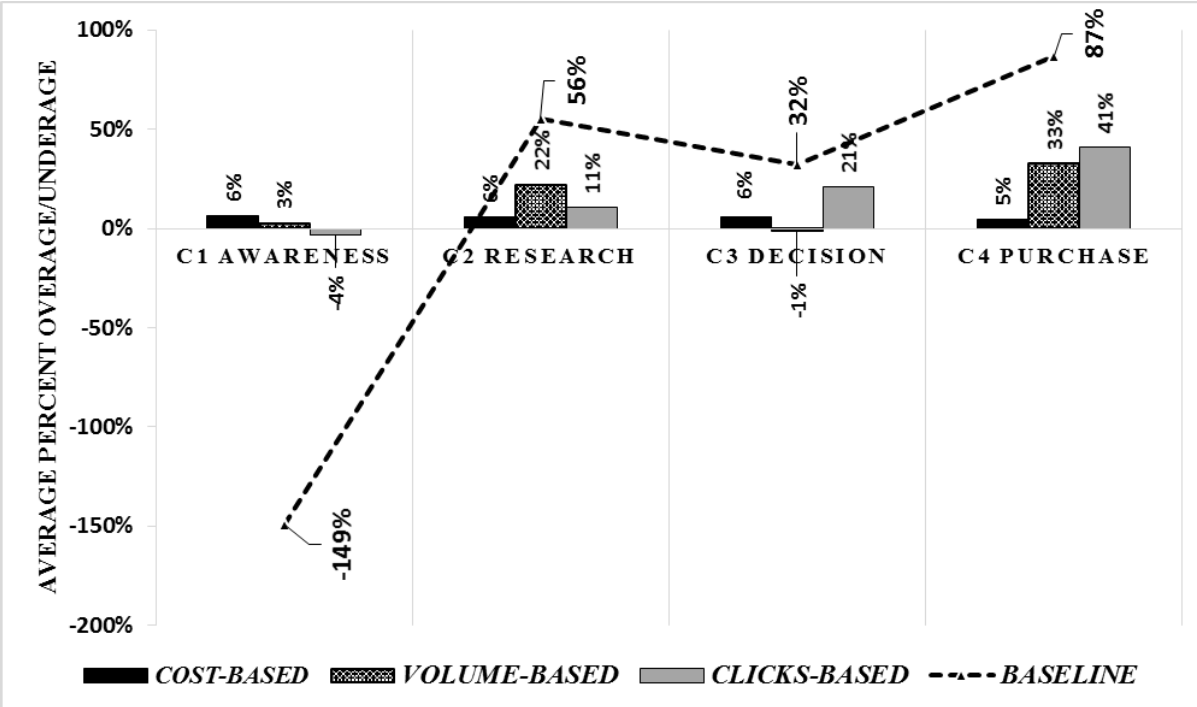

Figure 6-6 Macy's average percent overage/underage by segment

Macy's average percent overage/underage by segment is shown in Figure 6-6. From the figure, the following can be observed: (i) in all cases, the *Baseline* allocation strategy is outperformed, (ii) in this instance, the choice between the *Cost-based* and *Volume-based* allocation strategies is not straightforward. Using the *Cost-based* allocation strategy, there is consistent over-allocation of about 6%, and a reduction to the overall account budget of that amount would bring it closer to zero *overage/underage* in all campaigns. On the other hand, the *Volume-based* strategy performs better on the most expensive segments, "awareness" (3%) and "decision" (-1%), but poorer in the less expensive segments, "research" (22%) and "purchase" (33%). Due to the one percent underage observed for "research" keywords using the *Volume-based* strategy, there is potential lost opportunity. Therefore, the *Cost-based* strategy is the best strategy, as it has lowest average *overage/underage* value in all segments, (iii) the baseline strategy is unable to account for the variation in costs of the segments, and is severely under-budgeted for "awareness" and over budgeted for the three other segments.

6.1.3 Case-3: Prudential

As Goldfarb & Tucker (2011) suggested, the emphasis of Prudential's sponsored search initiative is in "research" and "awareness" stages, with 57 percent and 31 percent of its keywords, respectively. This is not surprising given the nature of the life insurance industry, with products and services that are invested in over extended periods of time, and high involvement purchase decisions. Additionally, the highly personal nature of life insurance products and services, and the individual level customization of terms, conditions, and policies make it difficult to select "purchase" or final conversion keywords. Of particular note is the power of the buying funnel in showcasing the differences in CPCs (see Figure 6-7) and the high intensity of competition in the research stage as shown in Section 5.4, where there is a low average CPC of $4.35 for "awareness" keywords and a high of $11.07 for "research" keywords. The types of keywords in the "awareness" stage are related to financial and life insurance queries such as "pension planning", "retirement and financial planning", and "advice on life insurance". Prudential's "research" keywords are also heavily focused on financial and life insurance questions such as "whole life vs term life insurance", "asset liability management insurance, and "single premium annuity calculator". Its "decision keywords are focussed on competitors' services and promotion of its own, with keywords such as, "north carolina mutual life insurance company", "hartford retirement services", and "prudential life insurance quote". Finally, Prudential's "purchase" keywords mostly target competitor's specific products and services such as, "rollover 401k to roth", "metlife investors variable annuity", and "the hartford variable annuity".

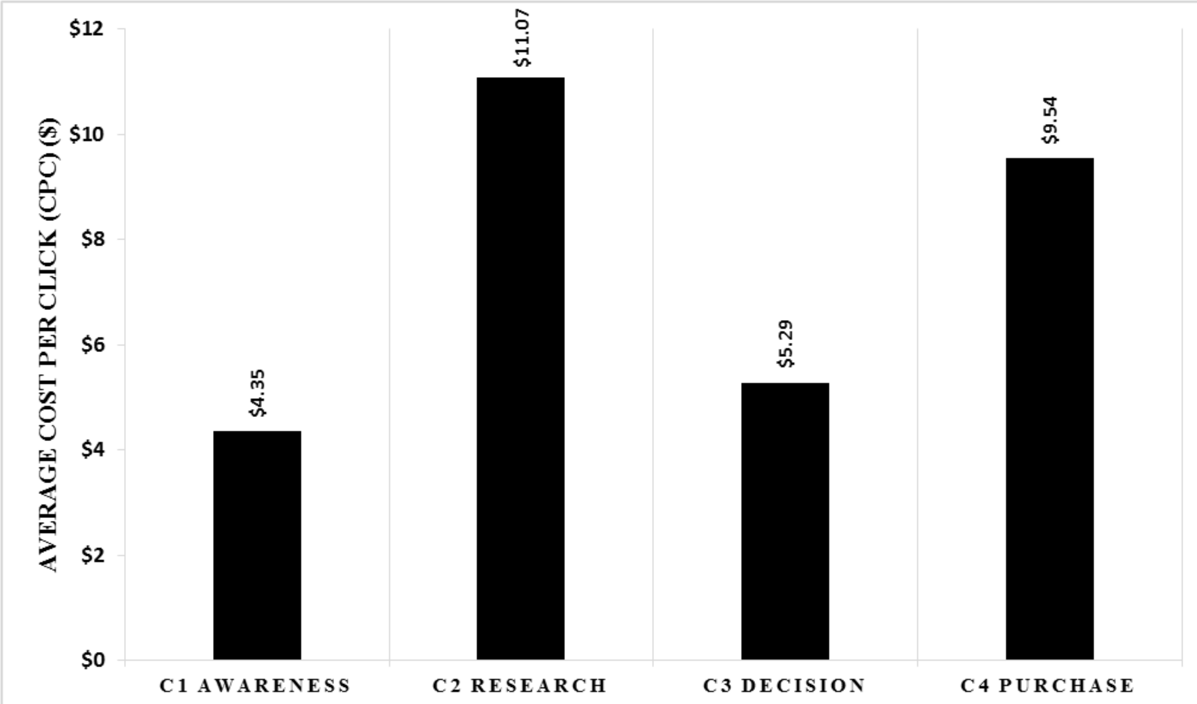

Figure 6-7 Prudential's average CPCs

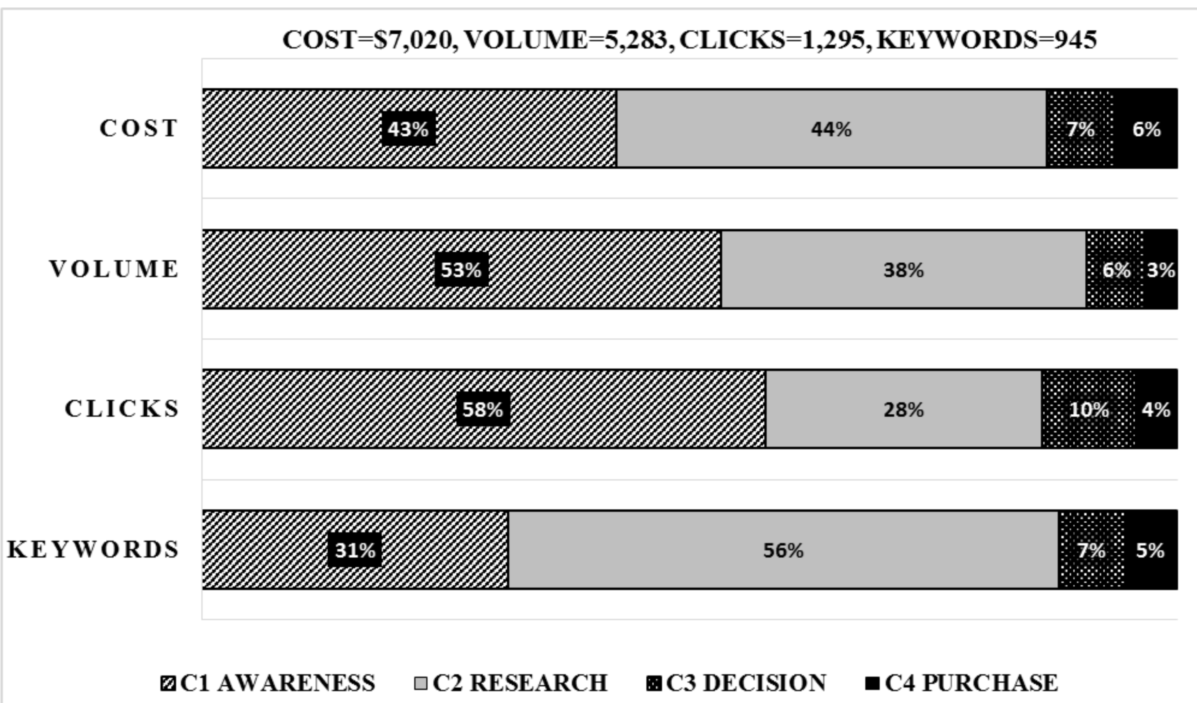

Figure 6-8 Prudential's average contribution of cost, volume, clicks and keywords by segment

The percent contribution of each segment to total cost, volume, clicks, and the number of keywords, along with the corresponding totals, is shown for Prudential in Figure 6-8. From the figure, the following can be observed: (i) “Awareness” keywords offer the greatest value in terms of cost, volume, and clicks. While contributing 43 percent of the total cost, they account for a higher percentage of total volume (53%) and total clicks (58%), (ii) the “decision” and “purchase” keywords contribute far less than “awareness” and “research” keywords across all measures, (iii) Prudential has invested heavily in the number of “research” keywords (56%), but they only provide 28 percent of total clicks and contribute the most cost (44%). A possible explanation for Prudential’s large investment in research keywords is that they lead to the most valuable conversions.

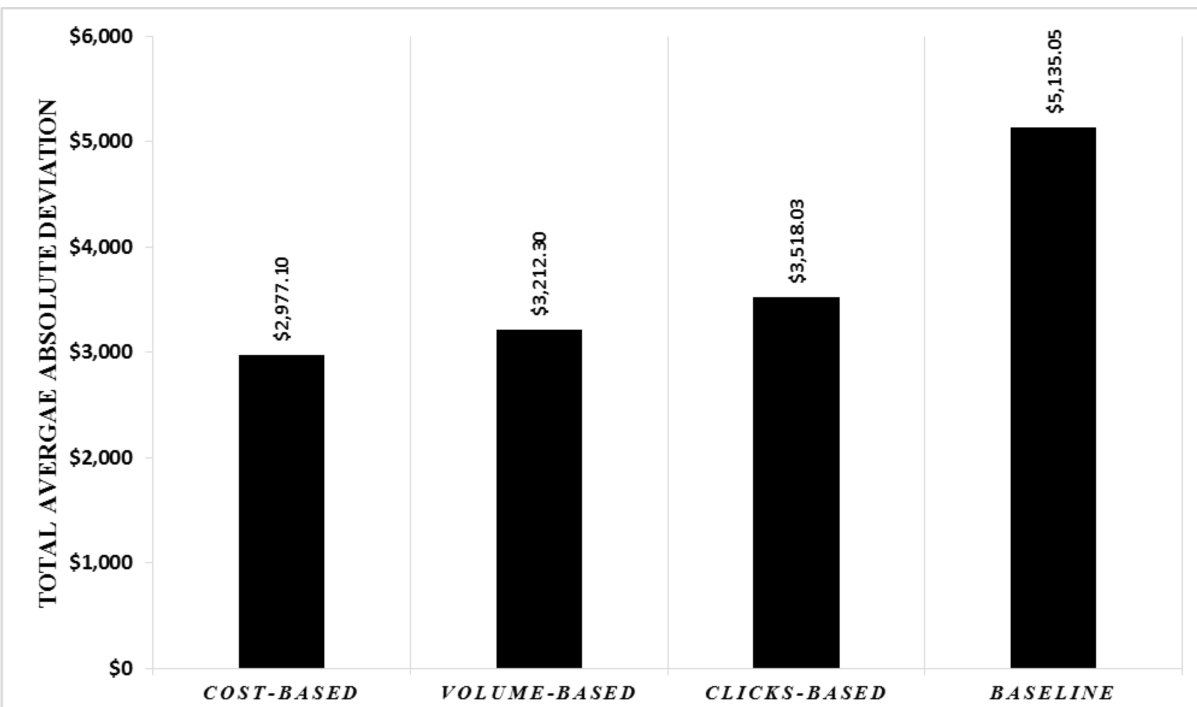

Figure 6-9 Prudential’s total average absolute deviation in dollars ($)

The total average absolute deviation is shown for Prudential in Figure 6-9. From the figure, the following can be observed: (i) at the expected budget, each allocation strategy has a large total average absolute deviation, (ii) however, the proposed allocation strategies still outperform the *Baseline* with roughly 60 percent the total average absolute deviation, (iii) the *Volume-based* and *Clicks-based* strategies are second and third best options, respectively, (iv) a reduction in budget would produce lower values.

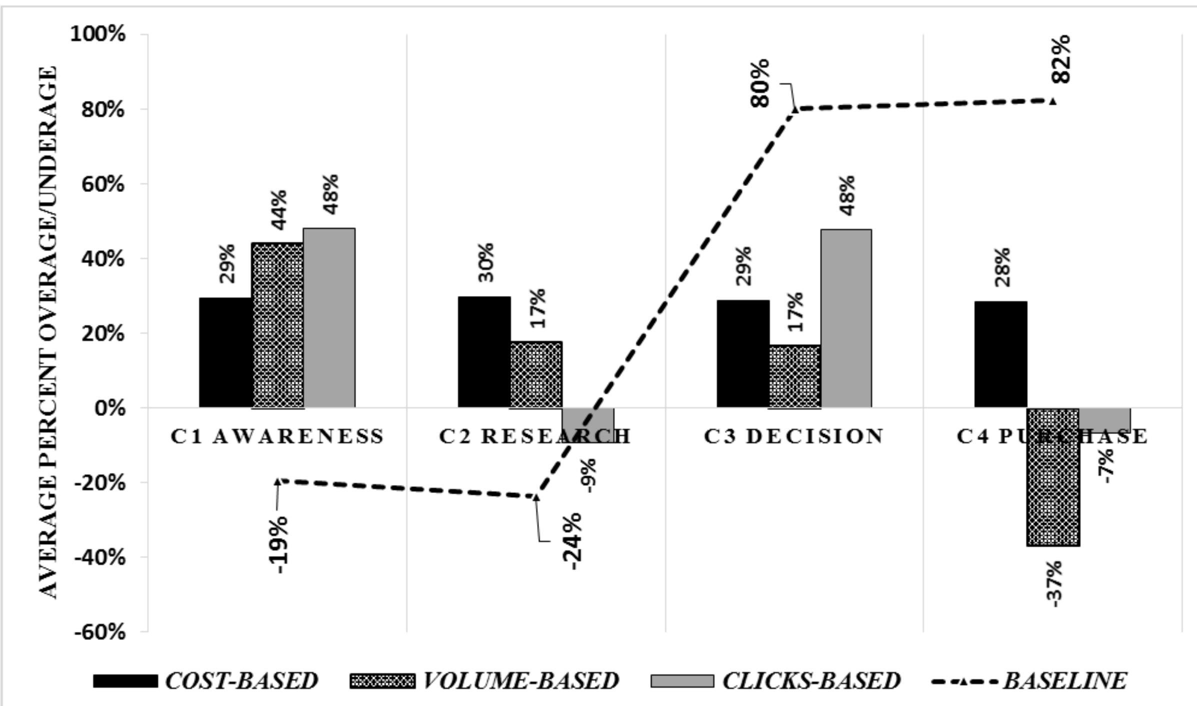

Figure 6-10 Prudential's average percent overage/underage by segment

Prudential's average percent *overage/underage* by segment is shown in Figure 6-10. From the figure, the following can be observed: (i) by using the *Cost-based* allocation strategy, the budget is consistently about 29% higher and a reduction of this percent to the total account budget would bring the deviation closer to zero, (ii) the *Volume-base*d strategy resulted in a shortage of budget for "purchase" keywords. This is due to the low average volume of 4.57 and high CPCs of $9.54 of the "purchase" keywords relative to the other keywords as shown in Figure 5-2 and Figure

5-3, (iii) the *Baseline* strategy outperforms all proposed allocation strategies on “awareness” keywords and shows better performance than the *Cost-based* strategy on “research” keywords, but is significantly over budgeted for the “decision” and “purchase” keywords, (iv) the *Clicks-based* allocation strategy resulted in the highest average percent *overage/underage* with “awareness” keywords and had an equal percent in campaign three. It is the only strategy at this budget level to have two campaigns (C2 and C4) under budgeted.

6.1.4 Case-4: United Continental Holdings

Interestingly, the classification of United Continental Holdings’ keywords resulted in a heavily skewed funnel shape with very little invested in the “decision” and “purchase” stages of the buying funnel (4% combined). This is possibly due to flights and airline travel being perceived as a commodity good, where most consumers are only interested in the price of travel and not differences in service quality or amenities. Considering this, the low investment in the latter stages of the buying funnel may also be driven by the nature of the airlines industry and ease of finding relevant information once on a travel agent website or airline website. Additionally, when entering a search term on Google such as “flight to …”, it provides the user with its’ own sponsored results in special windows in varying positions among the other advertisements as shown in Figure 6-11, as well as Figure A-1 and Figure A-2 in the appendix. This leads to a separate service at google.com/flights. The effect of this is highlighted by the diminishing level of volume observed in the latter stages of the buying funnel as shown in Figure 6-12.

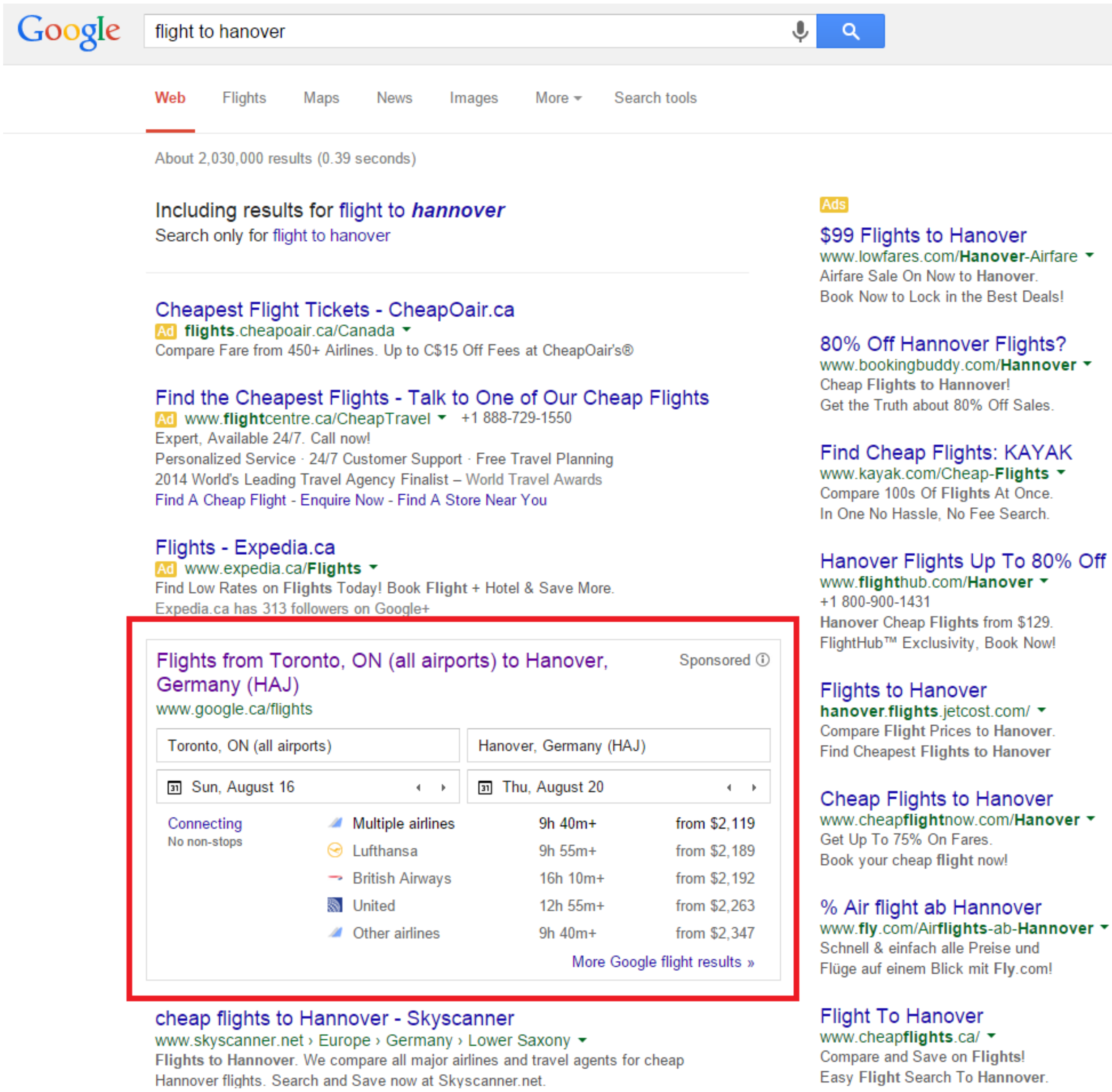

Figure 6-11 Google's search results for the term "flight to hanover"

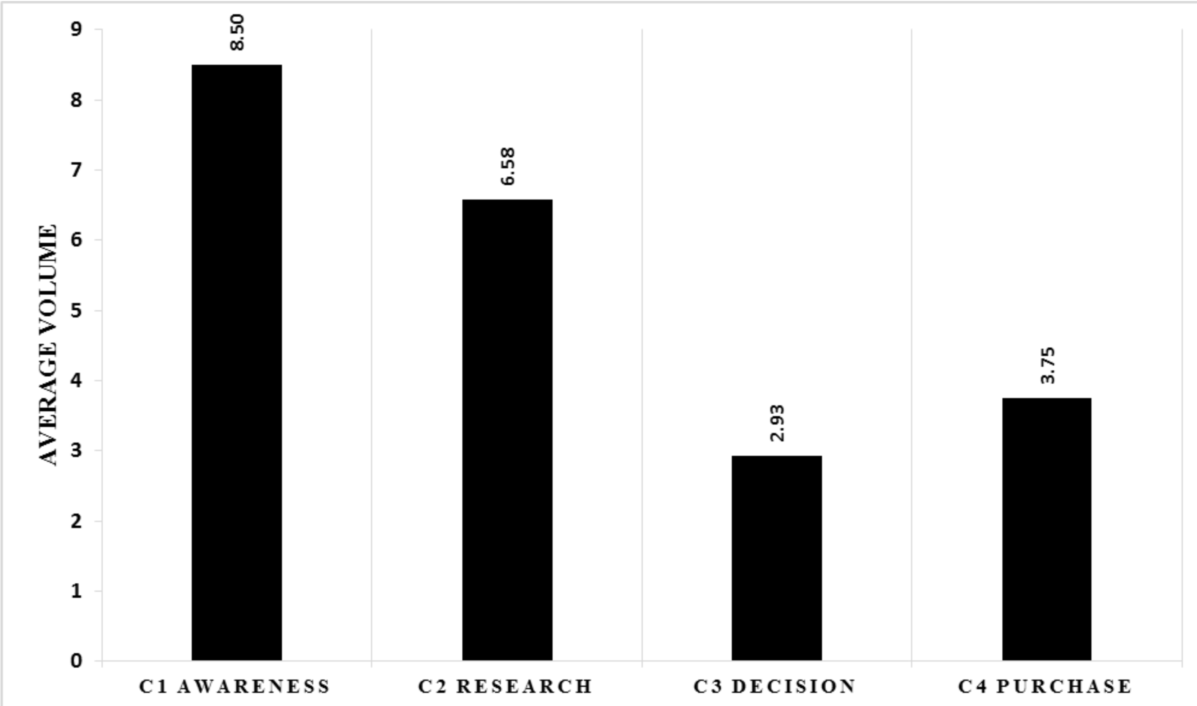

Figure 6-12 United Continental Holdings' average volume by segment

United Continental Holdings' "awareness" keywords are mostly a destination along with the mode of travel desired either for determining if it is available or to monitor the progress of a current flight such as "flights Jamaica", "shuttle to sfo", "new york to tel aviv flights". Its "research" keywords mostly include terms that indicate intentions to find out more about the purchase of airfare such as, "vancouver flight deals", "cheap flights san francisco to London", and "cheapest airline to dubai". United's "decision" keywords are almost entirely focused on vacation packages such as, "disney vacations packages with airfare", "apple vacations las vegas", and "best disney world vacation packages". Finally, its "purchase" keywords are full destination and specific vacation packages and a single product such as, "riu caribe all inclusive cancun", "princess cruise tour Alaska", and "united business credit card".

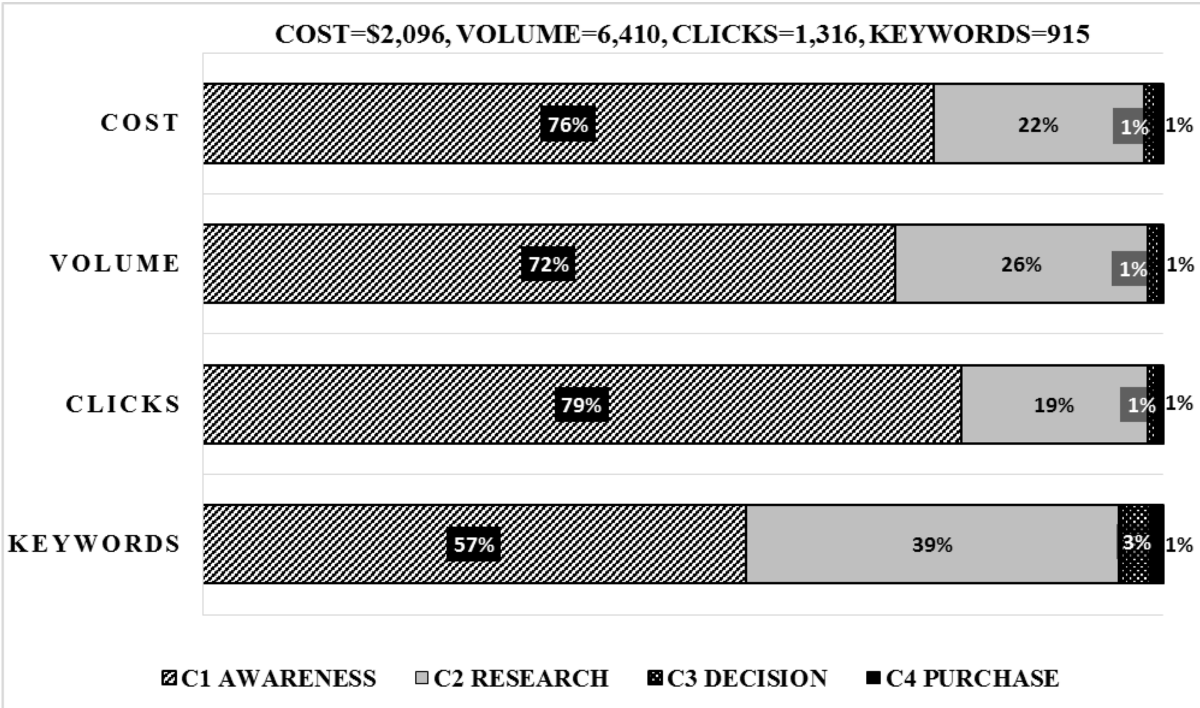

Figure 6-13 United Continental Holdings' average contribution of cost, volume, clicks and keywords by segment

The percent contribution of each segment to total cost, volume, clicks, and the number of keywords, along with the corresponding totals, is shown for United Continental Holdings in Figure 6-13. From the figure, the following can be observed: (i) "Decision" and "purchase" keywords contribute minimally to overall outcomes for all measures, (ii) "Awareness" keywords contribute more clicks (79%) than costs (72%) and account for more cost, volume, and clicks than percent of keywords (57%), (iii) this is in contrast to the "research" keywords, which account for more volume (26%) than cost (22%), and contribute less cost, volume, and clicks than the percent of keywords (39%).

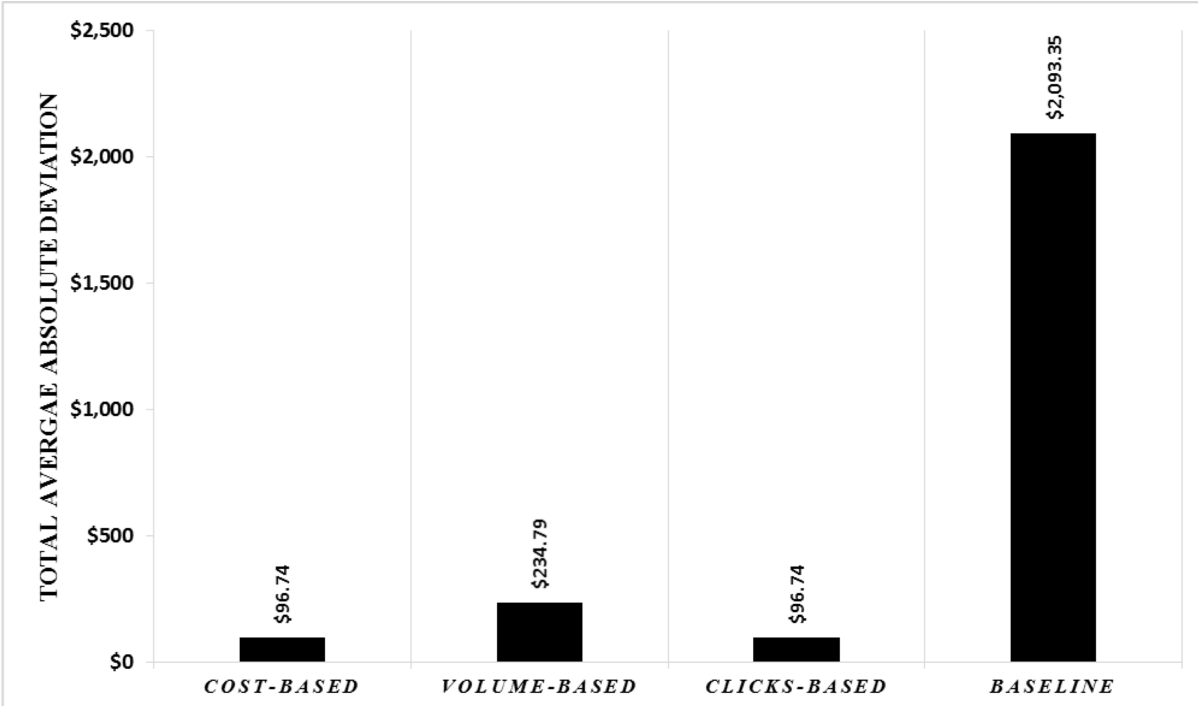

Figure 6-14 United Continental Holdings' total average absolute deviation in dollars ($)

United Continental Holdings' total average absolute deviation is shown in

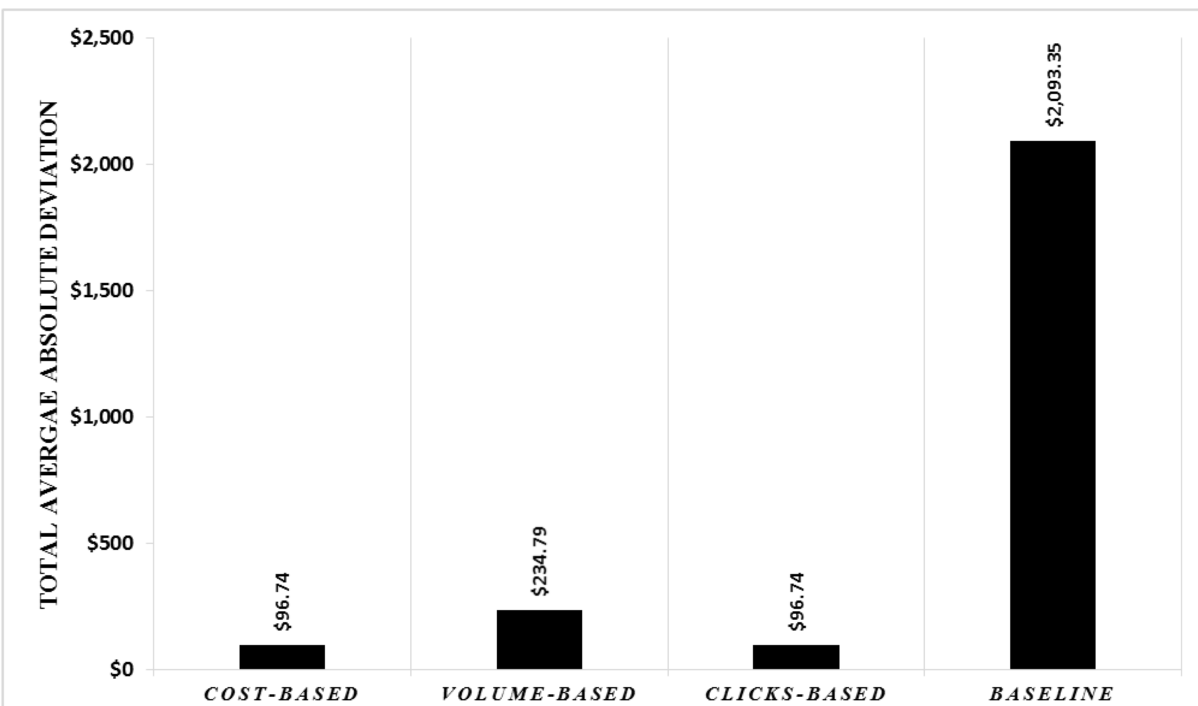

Figure 6-14. From the figure, the following can be observed: (i) all allocation strategies significantly outperform the *Baseline*, (ii) Interestingly, the total average absolute deviation for both the *Cost-based* and *Clicks-based* strategies are identical, (iii) the *Baseline* strategy resulted in a total average absolute deviation equal to the budget. This is caused by the heavily skewed funnel with significant investments in both "awareness" and "research" and minimal investment in "decision" and "purchase", (iii) The *Volume-based* strategy performed the poorest, but is still significantly more efficient at allocating budget as the *Baseline*.

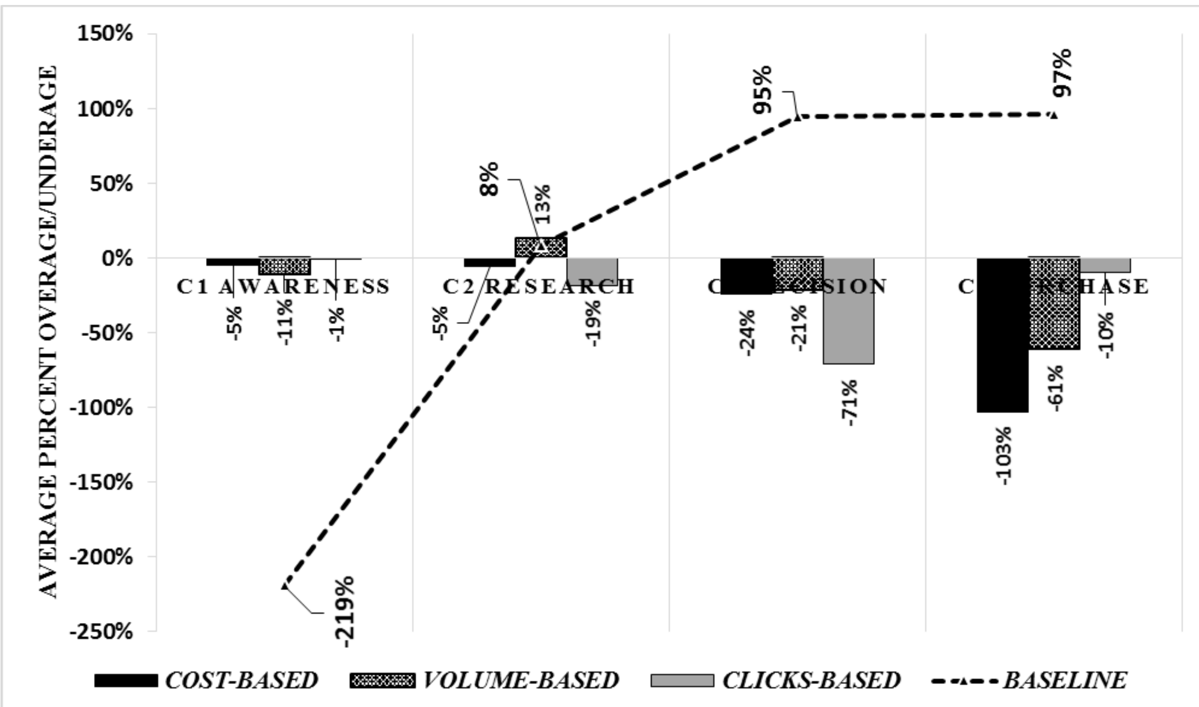

Figure 6-15 United Continental Holdings' average percent overage/underage by segment

United Continental Holdings' average percent overage/underage by segment is shown in Figure 6-15. From the figure, the following can be observed: (i) the *Baseline* is outperformed in all cases, except for research keywords. The spread of budget equally across segments led to a budget overage of 219 percent on "awareness" keywords, and left 95 percent of "decision", and 97 percent of "purchase" campaigns' budgets unutilized, (ii) of note is that the *Cost-based* strategy has equal levels of underage for both “awareness” and “research” at five percent, but is 24 percent and 103 percent under-budgeted for “decision” and “purchase”, respectively. This disparity resulted from a small portion of the budget being allocated to these segments, with average *overage/underage* values of -$1.21 for “decision” and -$0.79 for “purchase”. This is also the case for the *Volume* and *Clicks-based* strategies, (iii) the *Clicks-based* strategy resulted in the same total average absolute deviation as the cost-based strategy, but it is better for campaign one and four.

6.2. Statistical Analysis

In order to assess the simulation results, statistical tests were conducted. Specifically, I run paired t-tests to compare the total absolute deviations of each of the proposed allocation strategies to the *Baseline* strategy. Paired t-tests are used to compare the means of two populations when samples are correlated or matched pairs, or when it is a case-control study, and is therefore the best suited method for comparing the proposed allocation strategies to the *Baseline* strategy. The paired t-tests are run using 3000 observations, from 100 replications of 30 days. The summary results are shown in the following sections and the raw output from SPPS® 20 is shown in Appendix B.

6.2.1 Case 1: Lenovo

As shown in the previous section, there were noted differences between the total average absolute deviation of each proposed allocation strategy and that of the baseline. In confirmation of the significance of each of these differences, I formulate the following three hypotheses:

Hypothesis 1-A: There is a significant difference between the *Cost-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Lenovo's case.

Hypothesis 1-B: There is a significant difference between the *Volume-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Lenovo's case.

Hypothesis 1-C: There is a significant difference between the *Clicks-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Lenovo's case.

Table 6-1 Case 1: Lenovo paired t-test results

LENOVO		Paired Differences					
		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	*COST-BASELINE*	-2598.7168	177.482942	3.2403804	-801.98	2999	0.000
Pair 2	*VOLUME-BASELINE*	-2225.43826	126.195125	2.3039972	-965.9	2999	0.000
Pair 3	*CLICKS-BASELINE*	-1387.70413	295.630354	5.3974471	-257.1	2999	0.000

As Table 6-1 above shows all three hypothesis 1-A, 1-B, and 1-C, are confirmed. Of note is that each is significantly different at the 99 percent confidence level. Therefore, in terms of total absolute deviation, each allocation strategy is statistically significantly better than the Baseline allocation strategy in Lenovo's case.

6.2.2 Case 2: Macy's

Similarly, there were noted differences for the total average absolute deviations in Macy's case. In confirmation of the significance of each of these differences, I formulate the following three hypotheses:

Hypothesis 2-A: There is a significant difference between the *Cost-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Macy's case.

Hypothesis 2-B: There is a significant difference between the *Volume-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Macy's case.

Hypothesis 2-C: There is a significant difference between the *Clicks-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Macy's case.

Table 6-2 Case 1: Macy's paired t-test results

MACY'S		Paired Differences			t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean			
Pair 1	*COST-BASELINE*	-1088.90045	104.603625	1.9097922	-570.17	2999	0.000
Pair 2	*VOLUME-BASELINE*	-1075.61683	89.4299984	1.6327609	-658.77	2999	0.000
Pair 3	*CLICKS-BASELINE*	-1029.95874	63.5116918	1.1595595	-888.23	2999	0.000

As Table 6-2 above shows all three hypothesis 2-A, 2-B, and 2-C, are confirmed. Of note is that each is significantly different at the 99 percent confidence level. Therefore, in terms of total absolute deviation, each allocation strategy is statistically significantly better than the *Baseline* allocation strategy in Macy's case.

6.2.3 Case 3: Prudential

For Prudential, in confirmation of the significance of the differences between the allocation strategies and the baseline, I formulate the following three hypotheses:

Hypothesis 3-A: There is a significant difference between the *Cost-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Prudential's case.

Hypothesis 3-B: There is a significant difference between the *Volume-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Prudential's case.

Hypothesis 3-C: There is a significant difference between the *Clicks-based* allocation strategy and the *Baseline* strategy in total absolute deviation in Prudential's case.

Table 6-3 Case 1: Prudential paired t-test results

PRUDENTIAL		Paired Differences					
		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	*COST-BASELINE*	-2161.9368	736.397822	13.444723	-160.8	2999	0.000
Pair 2	*VOLUME-BASELINE*	-1924.71571	746.980069	13.637928	-141.13	2999	0.000
Pair 3	*CLICKS-BASELINE*	-1511.24835	685.855634	12.521953	-120.69	2999	0.000

As Table 6-3 above shows all three hypothesis 3-A, 3-B, and 3-C, are confirmed. Of note is that each is significantly different at the 99 percent confidence level. Therefore, in terms of total absolute deviation, each allocation strategy is statistically significantly better than the *Baseline* allocation strategy in Prudential's case.

6.2.4 Case 4: United Continental Holdings

Finally, for United Continental Holdings, in confirmation of the significance of the differences between the allocation strategies and the baseline, I formulate the following three hypotheses:

Hypothesis 4-A: There is a significant difference between the *Cost-based* allocation strategy and the *Baseline* strategy in total absolute deviation in United Continental Holdings' case.

Hypothesis 4-B: There is a significant difference between the *Volume-based* allocation strategy and the *Baseline* strategy in total absolute deviation in United Continental Holdings' case.

Hypothesis 4-C: There is a significant difference between the *Clicks-based* allocation strategy and the *Baseline* strategy in total absolute deviation in United Continental Holdings' case.

Table 6-4 Case 1: United Continental Holdings paired t-test results

UNITED		Paired Differences					
		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	*COST-BASELINE*	-1933.04272	90.7584468	1.657015	-1166.6	2999	0.000
Pair 2	*VOLUME-BASELINE*	-1851.9043	56.0287747	1.0229408	-1810.4	2999	0.000
Pair 3	*CLICKS-BASELINE*	-1928.69796	124.116937	2.2660549	-851.13	2999	0.000

As Table 6-4 above shows all three hypothesis 4-A, 4-B, and 4-C, are confirmed. Of note is that each is significantly different at the 99 percent confidence level. Therefore, in terms of total absolute deviation each allocation strategy is statistically significantly better than the *Baseline* allocation strategy in United Continental Holdings' case.

The statistical test results clearly show that the proposed budget allocation heuristics significantly outperform the *Baseline* strategy. Therefore, the use of any of these allocation strategies, along with the segmentation of keywords as well as campaign organization into the buying funnel framework, improves advertiser's overall budget utilization. Additional paired t-tests were also conducted comparing each of the proposed performance-based allocation strategies to one another, and the results show significant differences (see Appendix B). These variations indicate the options advertisers have in performance-based allocation strategies, depending on their product/service contexts and marketing goal orientations.

Chapter 7 Theoretical and Practical Implications

Overall, one major observation from the results of the simulation experiments for all allocation strategies is that they all consistently outperform the *Baseline* strategy in terms of total average absolute deviation. The *Cost-based* allocation strategy performs the best in all cases. The implication of this is that using historical information to inform the current time period leads to better performance than spreading the budget equally across all campaigns. Therefore, using past performance as an input potentially reduces an advertiser's lost opportunity and inefficiency in budget utilization. The *Cost-based* strategy, as it is the measure that is directly related to budget spending, adapts to the changes in spending pattern of each campaign and is better able to allocate available funds.

It is also important to highlight the interrelationships among the allocation strategies. For example, total cost of a campaign depends on the number of clicks and CPCs, and the number of clicks depends on the volume of searchers. These relationships are also affected by the company (industry it competes in), its inferred marketing strategy, and its product/service offerings. While the Cost-based strategy shows consistent performance across the segments (campaigns), it does not consistently outperform the other strategies in each campaign. Table 7-1 below summarizes the four cases, the four allocation strategies, as well as the findings and their implications.

Table 7-1 Summary of case findings and implications

	Lenovo	**Macys**	**Prudential**	**United Continental Holdings**
Industry	Computer Hardware ▪ high tech and oligopolistic	Department Stores ▪ Competitive and majority brick and mortar presence	Life Insurance ▪ Highly competitive and a mixture of online and offline channels	Airlines ▪ Highly competitive with many online booking avenues
Basic company profile	▪ Well-established brand ▪ Main products are laptops	▪ Vast array of products ▪ large department store with heavy brick and mortar presence	▪ Many financial and life insurance products ▪ Long term investments	▪ Airline tickets varying in price and price elasticity ▪ Middle of the pack service
Product development life cycle	▪ Expensive longer term purchase decisions ▪ Products vary across the product development life cycle ▪ New models of laptops and tablets (sometimes entire product categories) introduced on an annual basis	▪ Large variety of products varying across the product development life cycle and in purchase intensity ▪ Majority of products are low involvement purchases, such as, clothing and accessories	▪ The product development life cycle does not apply unless it refers to the consumer and stage of life they are in. ▪ Long term, high involvement purchase decisions	▪ The product development life cycle does not apply to airline tickets. Tickets can be purchased at virtually any period of time from months in advance to day of.
Inferred marketing strategy and focal stages	Competitive and conversion based. Focused more on click maximization and consumer attention grabbing in the "decision" and "purchase" stages	Generate awareness (volume focus) for both online and offline channels, with more emphasis on making consumers aware that products are available in the brick and mortar stores where the consumer can be cross- and up-sold	Generate awareness and convince consumers to choose Prudential in the "research" stage	Generate clicks and bookings through United's reservation system

Buying funnel representative ness	Representative with more emphasis on the final stages, “decision” and “purchase”	Representative with the general funnel shape.	Not representative with emphasis on the “research” stage	Somewhat representative with very little investment in the latter stages, “decision” and “purchase”
***Cost-based* allocation strategy**	▪ Lowest total absolute deviation ($462.99) ▪ Consistent underage of about 10 percent	▪ Lowest total absolute deviation ($94.80) ▪ Consistent overage of about six percent	▪ Lowest total absolute deviation ($2,977.10) ▪ Consistent overage of about 29 percent	▪ Same total absolute deviation ($96.74) as the *Clicks-based* strategy ▪ Consistent underage in all stages
***Volume-based* allocation strategy**	▪ Higher total absolute deviation ($909.86) than the *Cost-based* strategy, but lower than the *Clicks-based* and *Baseline* strategies ▪ Under-budgeted in all stages, except “purchase”, but including the “decision” stage	▪ Slightly higher total absolute deviation ($99.03) than the *Cost-based* strategy ▪ Provides sufficient budget in all stages including the focal “awareness” and “research” stages, but is under-budgeted by a percentage in the “decision” stage	▪ Slightly higher total absolute deviation ($3,212.30) than the *Cost-based* strategy ▪ Provides adequate budget to all stages except for the “purchase” stage. ▪ The focal segments have excess budget	▪ Higher total absolute deviation ($234.79) than the *Cost-* and *Clicks-based* strategies, but much lower than the *Baseline* ▪ Under-budgeted in all stages, except the “research” stage
***Clicks-based* allocation strategy**	▪ Higher total absolute deviation ($1,768.38) than *Cost-* and *Volume-based* strategies, but much lower than the *Baseline* ▪ Under-budgeted in “awareness” and “research” and over-budgeted in “decision” and “purchase” stages where more budget is required	▪ Slightly higher total absolute deviation ($157.70) than *Cost-* and *Volume-based* strategies, but much lower than the *Baseline* ▪ Under-budgeted in the focal “awareness” stage and sufficiently budgeted in the other three stages	▪ Slightly higher total absolute deviation ($3,518.03) than *Cost-* and *Volume-based* strategies, but much lower than the *Baseline* ▪ Under-budgeted in the focal “research” stage and in the “purchase stage ▪ Overage in the “awareness” and decision” stages	▪ Same total absolute deviation ($96.74) as the *Cost-based* strategy ▪ Under-budgeted in all stages ▪ Has the lowest lost opportunity of one percent in the focal “awareness” stage, but 19 percent underage in the “research” stage

***Baseline* allocation strategy**	▪ Highest total absolute deviation ($3,161.69) ▪ Unable to account for the varying costs across the stages	▪ Highest total absolute deviation ($1,215.24) ▪ Unable to account for the varying costs across the stages. ▪ Significant underage (-149%) in the “awareness” stage	▪ Highest total absolute deviation ($5,135.05) ▪ Unable to account for the varying costs across the stages ▪ Over 80 percent of the allocated budget is unutilized in the “decision” and “purchase” stages	▪ Highest total absolute deviation ($2,093.35) ▪ Unable to account for the varying costs across the stages ▪ Has better budget allocation in the “research” stage than both the *Volume-based* and *Clicks-Based* strategies
Key implications	The inferred strategy of clicks and conversions is best attained using the *Clicks-based* allocation strategy. This strategy allocates more budget to the final two stages “decision” and “purchase” where clicks are the most important	The inferred strategy of awareness is best attained by both the *Cost-based* and *Volume-based* allocation strategies. The *Cost-based* is recommended due to the one percent misallocation in the “decision” stage using the *Volume-based* allocation strategy	The inferred strategy of awareness is best attained by both the *Cost-based* and *Volume-based* allocation strategies. The Volume-based strategy only misallocates the “purchase” stage but with such little contribution this translates to a loss of $117.60	The inferred strategy of clicks is in the early stages of “awareness” and “research” is best served by the *Clicks-based* allocation strategy. While there is 19 percent overage in the “research” stage, the low misallocation (-1%) in the “awareness” stage is more important
Final recommend-ations	In line with the goal of clicks and conversions, the *Clicks-based* strategy is best able to allocate budget to the “decision” and “purchase” stages	With a focus on generating awareness the *Volume-based* allocation strategy is most in line with the inferred marketing strategy and sufficiently allocates budget to the focal segments	To create consumer awareness of its brand, the *Volume-based* strategy provides more than sufficient budget for the focal stages of “awareness” and “research”	With an emphasis on having consumers utilize its reservation system, the *Clicks-based* strategy is best able to allocate budget to the “awareness” stage

Considering the longer-term higher involvement purchase decisions Lenovo's consumers face, it invests more in "decision" and "purchase" keywords, relative to the other cases (companies). Of note is that, as expected, "purchase" keywords, where in this cognitive stage the objective is to convert consumers, there is a larger contribution of clicks compared to the other segments. This would suggest that the keywords are serving their intended purposes. As a result, the *Clicks-based* strategy allocated more budget to these campaigns and has sufficient budget capable of accommodating the higher number of clicks.

As a department store with a vast array of products varying in consumer purchase involvement and intensity, Macy's investment of keywords follows the shape of the buying funnel. Considering that the majority of the products Macy's carries are low involvement purchases, such as clothing and accessories, it is most likely that it has an overall awareness oriented marketing goal. The purpose is to make a consumer aware that it carries the products they are searching for and then subsequently cross-and up-sell once on its website. In confirmation of this is the dominating contribution of the "awareness" keywords to total cost, clicks, and volume. Interestingly, the shape of the traditional buying funnel is violated by the low percent contribution of the "research" keywords to cost, volume, and clicks. While clicks may be the marketing goal, due to "awareness" keywords having a cost contribution (67%) much larger than its clicks contribution (60%), the *Clicks-Based* strategy did not allocate enough budget. It resulted in overage for "awareness" keywords and has excess budget or underage in the three other segments. However, considering the marketing objective of awareness, the *Volume-based* strategy sufficiently allocates budget to the "awareness" keywords (3% overage), and is only under-allocated for "decision" keywords (-1%).

In Prudential's case, the inferred marking objective is to generate "awareness" and convince consumers to buy their insurance products at the "research" stage, likely targeting them via promotional material to convince them to buy their products through offline channels. The highly personal nature of life insurance products raises issues of privacy and targeting. As Goldfarb & Tucker (2011a) found, privacy plays a crucial role in the likelihood of a consumer clicking on an advertisement, and Kim & Sundar (2010) found that relevance plays an important role in reducing negative perceptions of advertisements. Thus, the highly personal nature of life insurance and related products limits the ability of life insurance advertisers to create effective, targeted, and highly relevant advertisements in the "decision" and "purchase" stages, except for generic products. Considering this awareness focus, the allocation of budget using the *Volume-based* strategy provides more than sufficient budget in the lead generation stages ("awareness" and "research") and only exceeds budget in the "purchase" stage. On the other hand, due to the high CPCs of Prudential's "research" keywords, the *Clicks-based* strategy is unable to allocate enough budget and it exceeds the budget with nine percent underage.

As discussed earlier, the Airlines industry in which United Continental Holdings operates is highly competitive. It has the most competitive keywords of all the companies, in all stages, except for "purchase" (see Figure 5-5). Additionally, given that air travel is treated as a commodity by most consumers and that information about or the service quality of fights is assumed to be virtually similar across companies, the focus of United's sponsored search initiative is competing on price and attaining clicks in the "awareness" and "research" stages. Thus, the inferred goal at these stages is to have consumers click and book through United Continental Holding' reservation system. This awareness and research focus is further highlighted by the overwhelming lack of "decision" and "purchase" keywords, which combined represent only four percent of all keywords.

Considering the importance of clicks and navigation to its website, as shown in Figure 6-15, of all the allocation strategies, the *Clicks-based* strategy allocates budget most appropriately to “awareness” but exceeds its budget in the “research” campaign with 19 percent underage, while the *Volume-based* strategy exceeds the budget for “awareness” keywords with 11 percent underage, and has excess budget in the “research” stage with 13 percent overage and the *Cost-based* is consistently under-budgeted in both “awareness” and “research” campaigns.

Overall, as the results indicate, the company (industry it competes in), its inferred marketing strategy, and its product/service offerings inform which allocation strategy will perform the best. Additionally, while the *Cost-based* allocation strategy essentially spreads the average percent *overage/underage* consistently across all campaigns, it may not be the best strategy for all companies. For example, in United Continental Holdings’ case (Figure 6-15), the allocation strategy more closely associated with United’s inferred marketing strategy provides the relevant campaigns with better allocation of budget.

While Jansen and Shuster (2011) found that the buying funnel is sufficient for discriminating between keywords and their performance measures, they suggested that the ‘funnel’ may not be the best shape for describing the search process using their dataset from a large nation-wide retail chain. As my results confirm, the buying funnel model is adaptable to different domains and is capable of meaningfully segmenting keywords. However, as the resulting classification of keywords and the simulation results for cost, volume, and clicks show, the generalization of the ‘funnel’ shape as an imperfect representation of the search process may not be a correct assertion. Counter to that assertion is, in all cases except for Prudential, the investment in keywords follows the funnel shape. Further, both Prudential and United Continental Holdings have percent contributions of cost, volume, and clicks that follow the ‘funnel’ shape. In the two other cases, Lenovo and Macy's, the low contribution of the "research" keywords shows a disparity from the 'funnel' shape. These findings are not surprising considering the growing list of research that demonstrate that depending on product/industry (Lu & Zhao, 2014; Nottorf & Funk, 2013), differences in consumer behavior is expected. Therefore, the buying funnel is a flexible framework, capable of adapting to domain specific keywords.

The results clearly show that sponsored search is largely used as a lead generation tool. In all cases except for Prudential, the keywords classified into the "awareness" stage account for largest percent of budget consumption. Three of the four cases have 50 percent or more of the budget consumed in the "awareness" stage (Lenovo 50%, Macy's 67%, and United Continental Holdings 76%). This implies that advertisers can improve their budget utilization by segmenting their keywords and organizing their campaigns according to their promotional goals. With this observation, the common strategy of allocating equal budgets to campaigns with different promotional goals can easily lead to poor budget utilization and lost opportunities.

Another observation is that, in all cases except Macy's, in the "research" stage the click contribution is much less than the volume contribution. This is not surprising, considering that, as shown in Figure 5-4 and Figure 5-5, the "research" stage is on average the most competitive and the hardest segment to attain a high position, which drives the cost of clicks up and the click-through rate down. However, it is interesting to note that this phenomenon is shown in all cases with "research" keywords, where in this cognitive stage, the customer has decided on which product they want, most likely demonstrates commercial intent, but has not decided on a brand (Jansen & Shuster, 2011).This finding suggests "research" keywords are crucial. In the "research" stage, awareness or volume and exposure to consumers is important and may be even more than clicks. Specifically, this is the final stage where an advertiser can inject its brand into the consumers' consideration set prior to reaching the "decision" stage, where the consumer is considering alternatives and doing comparison shopping, and prior to the "purchase" stage where the consumer knows what and/or where he/she wants to purchase. As further emphasis of the importance of this stage, in Prudential's case, the "research" stage is its focus, as shown by the contribution to the number of keywords (56%). It is also Prudential's most expensive segment (average CPC of $11.07 and cost contribution of 44%), and relative to the number of keywords (56%), it contributes less volume (38%), and significantly less clicks (28%).

7.1 Limitations and Future Research

This study has a number of limitations that should be acknowledged or addressed in future research. First of all, the simulation experiments are not conducted in a real-time setting. This may have implications on the consumption of budget, where on any given day, a high utility keyword may be highly searched and drain a campaign budget in the first hours of the day, meaning no ads will be shown for that campaign for the remainder of the day (Zhang et al., 2013). Future research can assess the proposed allocation strategies in real-time AdWords accounts to examine the outcomes in a more practical, competitive setting. In addition, throughout the simulation experiments, the account level budget is assumed to be fixed. Future research can plan to combine the account level budget setting decision with the campaign allocation decision.

While this study used the buying funnel model as the main theoretical framework, there is a growing body of literature on the segmentation of keywords (Ashkan & Clark, 2013; Broder, 2002; Ghose & Yang, 2010; Jansen et al., 2008; Lu & Zhao, 2014; Nottorf & Fun; Rutz & Bucklin, 2011). Future studies can compare the effectiveness of these other theoretical frameworks for similar purposes. Furthermore, while the study only used four campaigns, each representative of a stage of the buying funnel, this may not reflect the way advertisers currently organize their accounts. As Zhang et al. (2014) show, according to their data, the average advertiser has 15 campaigns and a max of over 2000. In future studies, I hope to gain access to an active AdWords account to test the effectiveness of the proposed strategies and the overlay of the buying funnel framework on an existing account structure. The segmentation of keywords in this thesis was also limited to the application of the descriptions as given by Jansen and Shuster (2011). Although I was able to easily apply the simple guidelines they set out, a more comprehensive set of definitions is needed to effectively extend it to different domains and product/service categories.

The simulation experiments were also conducted using historical third-party data, which limited parameter estimates and the approximation of distributions across segments and at the individual keyword level. In future studies, I plan to have access to an active AdWords account in order to account for daily, weekly, and other cyclical trends that are typically observed in sponsored search. Additionally, because this study was limited to the use of third party data, data for quality scores, account hierarchy and organization, and the advertisements the keywords trigger were unavailable. With this information, a comparison between the existing account hierarchy and organization could be made with an organization based on the buying funnel. The problems of ad attribution and the estimation of value-per-click (VPC) could be added with the additional information about which advertisements were triggered by a keyword, the landing pages they point to, and the context of the lading page (purpose/function). In addition, more accurate classification could be achieved with the additional context of the advertisements (marketing message). Overall, the use of an active AdWords account could significantly improve the capabilities of the proposed model and its practical validation.

While these simple allocation strategies were better able to allocate the campaign level budget than the *Baseline*, more complex strategies and algorithms can improve upon this further.

To this end, the combined use of attribution modeling and an active AdWords account would provide ample opportunity to investigate the effects of additional measures. For instance, knowledge of the purpose (marketing context) of the landing page of a keyword would provide better opportunities for value estimation and the extension of this problem. Additionally, the true valuation of a search user landing on an advertiser's webpage would make the prioritization of marketing objectives in dollar terms more clear. This allows for better value estimation throughout the course of a sponsored search initiative. Furthermore, this allows for extending the current simulation model to incorporate value-based optimization.

Chapter 8 Conclusion

In this thesis, I studied the segmentation of keywords, organization of campaigns, and budget allocation problems. For this purpose, I proposed the use of the buying funnel framework to segment keywords and organize campaigns, in addition to the use of the proposed allocation strategies to align sponsored search initiatives with marketing goals and to improve the allocation of the daily advertising budget across marketing campaigns. I used four different cases sampled from a large collection of companies from many industries to demonstrate the implications of four different allocation strategies, including a baseline strategy. Experimental results show that the proposed allocation strategies consistently outperform the baseline strategy commonly used in practice.

I demonstrated the benefits of segmenting keywords into the buying funnel as a means to identify and track marketing objectives and for prioritizing campaign level budget allocation based on the marketing objectives. Further, I demonstrate that the buying funnel framework is adaptable and flexible enough to be applied to multiple account structures and that it can be applied in conjunction with other frameworks such as the product development life cycle to organize and justify keyword investments.

For future work, I plan to extend the current study in multiple angles. For example, I plan to introduce more complex allocation strategies/algorithms, incorporate constraints, apply other methods of keyword segmentation and campaign organization, augment with value-based optimization and validate the current simulation model expanded samples.

Bibliography

Abhishek, V., Fader, P. S., & Hosanagar, K. (2012) Media exposure through the funnel: A model of multi-stage attribution. Available at SSRN: http://ssrn.com/abstract=2158421 or http://dx.doi.org/10.2139/ssrn.2158421

Abhishek, V., & Hosanagar, K. (2007). Keyword generation for search engine advertising using semantic similarity between terms. In Proceedings of the ninth international conference on Electronic commerce, 89-94.

Agarwal, A., Hosanagar, K., & Smith, M. D. (2011). Location, location, location: An analysis of profitability of position in online advertising markets. *Journal of Marketing Research, 48*(6), 1057-1073.

Agashe, A. A. (2010). *Stochastic revenue optimization in online advertising*, State University of New York at Buffalo.

Aggarwal, G., Feldman, J., & Muthukrishnan, S. (2007). Bidding to the top: VCG and equilibria of position-based auctions. In *Approximation and Online Algorithms,* 15-28. Springer Berlin Heidelberg.

Animesh, A., Ramachandran, V., & Viswanathan, S. (2010). Research note-quality uncertainty and the performance of online sponsored search markets: An empirical investigation. *Information Systems Research, 21*(1), 190-201.

Ashkan, A., & Clarke, C. L. (2013). Impact of query intent and search context on clickthrough behavior in sponsored search. *Knowledge and Information Systems*, *34*(2), 425-452.

Auerbach, J., Galenson, J., & Sundararajan, M. (2008). An empirical analysis of return on investment maximization in sponsored search auctions. In *Proceedings of the 2nd International Workshop on Data Mining and Audience Intelligence for Advertising*, 1-9.

Ayanso, A. & Karimi, A. (2015) The moderating effects of keyword competition on the determinants of ad position in sponsored search advertising. *Decision Support Systems 70*(2015), 42-59.

Ayanso, A., & Mokaya, B. (2013). Efficiency evaluation in search advertising. *Decision Sciences*, *44*(5), 877-913.

Balakrishnana, R. & Kambhampati, S. (2008) Optimal Ad Ranking for Profit Maximization. *Proceedings of the 11th International Workshop on Web and Databases* (*WebDB* 2008), June 13, 2008, Vancouver, Canada

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research, 25*(3), 187-217.

Bloomberg (2015) Bloomberg Industry Market Leaders. As retrieved May 4, 2015 from http://www.bloomberg.com/visual-data/industries/rank/name:market-share

Braun, M., & Moe, W. W. (2013). Online display advertising: modeling the effects of multiple creatives and individual impression histories. *Marketing Science*, *32*(5), 753-767.

Brettel, M., & Spilker-Attig, A. (2010). Online advertising effectiveness: A cross-cultural comparison. *Journal of Research in Interactive Marketing*, *4*(3), 176-196.

Broder, A. (2002). A taxonomy of web search. *In ACM Sigir forum*, *36*(2), 3-10.

Callard, A. (2014). *Paid search pays off in a bigger way for online retailers*. As retrieved August 25, 2014 from http://www.internetretailer.com/2014/07/22/paid-search-pays-bigger-way-online-retailers?bro_mid=48906914&bro_rid=0bc103e80000000000000000000054084df

Caspari, P. *Management dynamics*. Wiley, New York, 2004.

Chaitanya, N., & Narahari, Y. (2012). Optimal equilibrium bidding strategies for budget constrained bidders in sponsored search auctions. *Operational Research*, *12*(3), 317-343.

Chen, L. (2010). Using a two-stage technique to design a keyword suggestion system. *Information Research*, *15*(1).

Chitika (2013) *The value of Google result positioning*. As retrieved August 25, 2014 from http://chitika.com/google-positioning-value.

Chakrabarty, D., Zhou, Y., & Lukose, R. (2008) Budget Constrained Bidding in Keyword Auctions and Online Knapsack Problems. *WWW2007*, May 8–12, 2007, Banff, Canada.

Cholette, S., Özlük, Ö., & Parlar, M. (2012). Optimal keyword bids in search-based advertising with stochastic ad positions. *Journal of Optimization Theory and Applications*, *152*(1), 225-244.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice*, *11*(1), 17-33.

Dayanik, S., & Parlar, M. (2013). Dynamic bidding strategies in search-based advertising. *Annals of Operations Research*, 211(1), 103-136.

Dhar, V. & Ghose, A. (2010) Research commentary—Sponsored search and market efficiency. *Information Systems Research, 21*(4), 760–772.

Edelman, B., Ostrovsky, M., & Schwarz, M. (2007). Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *The American Economic Review*, *97*(1), 242-259.

Edelman, B., & Schwarz, M. (2010). Optimal auction design and equilibrium selection in sponsored search auctions. *The American Economic Review, 100*(2), 597-602.

Engine Ready (2014) *Observations on how ad position impacts CTR, conversion rate and CPA*. As retrieved August 25, 2014 from http://www.engineready.com/sem-resources/industry-studies/ppc-ad-position-study.php

Feldman, J., Muthukrishnan, S., Pal, M., & Stein, C. (2007). Budget optimization in search-based advertising auctions. *In Proceedings of the 8th ACM conference on Electronic commerce*, 40-49.

Feng, J., Bhargava, H., & Pennock, D. (2007). Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms. *INFORMS Journal on Computing 19*(1), 137-148.

Fischer, M., Albers, S., Wagner, N., & Frie, M. (2014). Practice prize winner-dynamic marketing budget allocation across countries, Products, and Marketing Activities. *Marketing Science, 30*(4), 568-585.

Ghose, A., & Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, *55*(10), 1605-1622.

Ghose, A., and Yang, S. (2010) Modeling cross-category purchases in sponsored search advertising. *Working paper, Leonard N. Stern School of Business, New York University*, January 25, 2010.

Goldfarb, A., & Tucker, C. (2011a). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, *30*(3), 389-404.

Goldfarb, A., & Tucker, C. (2011b) Search Engine Advertising: Channel Substitution When Pricing Ads to Context. *Management Sciences*, *57*(3), 458-470.

Google (2015) *Manage a shared budget across campaigns*. As retrieved July 31, 2015, from https://support.google.com/adwords/answer/2517512?hl=en&ref_topic=3123046

Groves, T. (1973). Incentives in teams. *Econometrica: Journal of the Econometric Society*, *41*(4), 617-631.

Hoffman, D., & Novak, T. (1996). Marketing in hypermedia computer-mediated environments: conceptual foundations. *The Journal of Marketing*, 60, 50–68

Howard, J. A., & Sheth, J. N. (1969). *The theory of buyer behavior*. Wiley, New York.

IAB (2013*) IAB internet advertising revenue report: 2012 full year results An industry survey conducted by PwC and sponsored by the Internet Advertising Bureau (IAB).* Available at http://www.iab.net/AdRevenueReport

IAB (2014) *IAB internet advertising revenue report: 2013 full year results An industry survey conducted by PwC and sponsored by the Internet Advertising Bureau (IAB).* Available at http://www.iab.net/AdRevenueReport

Jansen, B. J., Booth, D. L., & Spink, A. (2007). Determining the user intent of web search engine queries. *In Proceedings of the 16th international conference on World Wide Web*, 1149-1150.

Jansen, B. J., Booth, D. L., & Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Information Processing & Management*, *44*(3), 1251-1266.

Jansen, B. J., & Mullen, T. (2008). Sponsored search: An overview of the concept, history, and technology. *International Journal of Electronic Business*, *6*(2), 114-131.

Jansen, B.J., Sobel, K.; & Zhang, M. (2011) The brand effect of key phrases and advertisements in sponsored search. *International Journal of Electronic Commerce, 16*(1), 77–106.

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

Jansen, B. J., & Schuster, S. (2011). Bidding on the buying funnel for sponsored search and keyword advertising. *Journal of Electronic Commerce Research*, *12*(1), 1-18.

Ji, L., Rui, P., & Hansheng, W. (2010). Selection of best keywords: A poisson regression model. *Journal of Interactive Advertising*, *11*(1), 27-35.

Kim, A., Balachander, S., & Kannan, K. (2012). On the optimal number of advertising slots in a generalized second-price auction. *Marketing Letters*, *23*(3), 851-868.

Kim, N. Y., & Sundar, S. S. (2010). Relevance to the rescue: Can “smart ads” reduce negative response to online ad clutter?. *Journalism & Mass Communication Quarterly*, *87*(2), 346-362.

Kiritchenko, S., & Jiline, M. (2008). Keyword optimization in sponsored search via feature selection. *The Journal of Machine Learning Research: Workshop and Conference Proceedings*, *4*, 1-13.

Kitts & Leblanc (2004) Optimal bidding on keyword auctions. *Electronic Markets, 14*(3), 186-201

Laffey, D. (2007). Paid search: The innovation that changed the Web. *Business Horizons*, *50*(3), 211-218.

Lahaie, S. (2006). An analysis of alternative slot auction designs for sponsored search. *In Proceedings of the 7th ACM Conference on Electronic Commerce*, 218-227.

Lancaster, G., & Withey, F. CIM Coursebook 06/07 *Marketing Fundamentals*. Butterworth-Heinemann, Oxford, 2006

Lavidge, R. J., & Steiner, G. A. (1961). A model for predictive measurements of advertising effectiveness. *The Journal of Marketing*, *25*(6), 59-62.

Lee, L., and Ariely, D. (2006) Shopping goals, goal concreteness, and conditional promotions. *Journal of Consumer Research. 33*(1) 60–70

Li, J., Pan, R., & Wang, H. (2010) Selection of best keywords: A poisson regression model. *Journal of Interactive Advertising 11*(1) 27-35

Liu, D., & Chen, J. (2006). Designing online auctions with past performance information. *Decision Support Systems, 42*(3), 1307-1320.

Liu, D., Chen, J., & Whinston, A. B. (2010). Ex ante information and the design of keyword auctions. *Information Systems Research, 21*(1), 133-153.

Lu, X. & Zhao, X. (2014) Differential effects of keyword selection in search engine advertising on direct and indirect sales. *Journal of Management Information Systems 30*(4) 299-325

Meyerson, M., & Scarborough, M. E., *Mastering online marketing: 12 keys to transform your website into a sales powerhouse*. Entrepreneur Press, Newburgh, NY, 2007.

Muthukrishnan, S., Pál, M., & Svitkina, Z. (2010). Stochastic models for budget optimization in search-based advertising. *Algorithmica, 58*(4), 1022-1044.

Naldi, M., D’Acquisto, G., & Italiano, G. F. (2010). The value of location in keyword auctions. *Electronic Commerce Research and Applications, 9*(2), 160-170.

Nottorf, F., & Funk, B. (2013). A cross-industry analysis of the spillover effect in paid search advertising. *Electronic Markets, 23*(3), 205-216.

Özlük, Ö., & Cholette, S. (2007). Allocating expenditures across keywords in search advertising. *Journal of Revenue and Pricing Management*, *6(4)*, 347-356.

Ramos, A., & Cota, S. (2008). *Search engine marketing*. McGraw-Hill, Inc.

Ree, P. (2000) Arena: Simulating e-commerce agent strategies. *Thesis Massachusetts Institute of Technology*

Regelson, M. & Fain, D. C. (2006) Predicting Click-Through Rate Using Keyword Clusters. *EC'06, SSA2* June 11, 2006, Ann Arbor, MI.

Rusmevichientong, P., & Williamson, D. P. (2006). An adaptive algorithm for selecting profitable keywords for search-based advertising services. *In Proceedings of the 7th ACM Conference on Electronic Commerce*, 260-269.

Rutz, O. J., & Bucklin, R. E. (2011). From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research*, *48*(1), 87-102.

Rutz, O. J., Bucklin, R. E., & Sonnier, G. P. (2012). A latent instrumental variables approach to modeling keyword conversion in paid search advertising. *Journal of Marketing Research*, *49*(3), 306-319.

Seda, C. *Search engine advertising: Buying your way to the top to increase sales*. New Riders, Boston, 2004. Thousand Oaks, CA, USA

SpyFu (2015) *Search marketing top lists*. As retrieved September 16, 2014 from http://www.spyfu.com/outreach/top-lists

SpyFu (2015) *Download Your Competitors' Most Profitable Keywords and Ads For Paid and Organic Search.* www.spyfu.com/

Statista (2015) *Worldwide market share of leading search engines.* from January 2010 to July 2014. As retrieved July 29, 2015 from http://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Varian, H. R. (2007). Position auctions. *International Journal of Industrial Organization*, *25*(6), 1163-1178.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, *16*(1), 8-37.

Webb, M. (2006). *Sales and marketing the six sigma way*. Kaplan Publishing.

Weber, T. A., & Zheng, Z. (2007). A model of search intermediaries and paid referrals. *Information Systems Research, 18*(4), 414-436.

Xie, Y., & O'Hallaron, D. R. (2002) Locality in search engine queries and its implications for caching. *In Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, vol. 3, New York,

NY, June 23–27, 2002, 1238–1247.

Yang, S., & Ghose, A. (2010). Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence?. *Marketing Science*, *29*(4), 602-623.

Yang, Y., Qin, R., Jansen, B. J., Zhang, J., & Zeng, D. (2014). Budget planning for coupled campaigns in sponsored search auctions. *International Journal of Electronic Commerce, 18*(3), 39-66.

Yang, Y., Yang, Y., & Zeng, D. (2013). Keyword optimization in sponsored search advertising: A multilevel computational framework. *Proceedings of the 23rd Workshop on Information Technology and Systems 2013.*

Yang, Y., Zhang, J., Qin, R., Li, J., Wang, F. Y., & Qi, W. (2012). A budget optimization framework for search advertisements across markets. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 42*(5), 1141-1151.

Yao, S., & Mela, C. F. (2009). *Sponsored search auctions: Research opportunities in marketing,* Now Publishers Inc.

Young, R. A., Weiss, A. M., & Stewart, D. W. (2006) *Marketing champions: Practical strategies for improving marketing's power, influence, and business impact*. Wiley, New York

Zhang, J., Yang, Y., Li, X., Qin, R., & Zeng, D. (2014a). Dynamic dual adjustment of daily budgets and bids in sponsored search auctions. *Decision Support Systems*, *57*, 105-114.

Zhang, M. & Feng, J. (2011) Cyclical bid adjustments in search-engine advertising. *Management Science*, *57*(9), 1703–1719.

Zhang, W., Zhang, Y., Gao, B., Yu, Y., Yuan, X., & Liu, T. (2012). Joint optimization of bid and budget allocation in sponsored search. *In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1177-1185.

Zhang, Y., Zhang, W., Gao, B., Yuan, X., & Liu, T. (2014b) Bid keyword suggestion in sponsored search based on competitiveness and relevance. *Information Processing and Management 50*(2014) 508-523

Zhou, Y., Chakrabarty, D., & Lukose, R. (2008). Budget constrained bidding in keyword auctions and online knapsack problems. *Internet and Network Economics,* Anonymous Springer, 566-576.

Zhu, Y., & Wilbur, K. C. (2011). Hybrid advertising auctions. *Marketing Science, 30*(2), 249-273.

Appendix A Sample Data, Variables, and Descriptive Statistics

Table A-1 Industries and companies selected

INDUSTRY	COMPANY NAME	URL
Airlines	United Continental Holdings	United.com
	Delta Air Lnes	delta.com
	Deutsche Luftansa AG	lufthansa.com
	Air France-KLM	airfrance.com
	American Airlines Group Inc.	aa.com
Apparel, Footwear, Acc Design	Nike Inc.	nike.com
	LVMH Moet Hennessy Louis Vuitton	louisvuitton.com
	Adidas AG	adidas .com
	Cie Financial Richmont SA	dunhill.com
	Kering	gucci.com
Automobile OEM	Toyota Motor Corp	toyota.com
	Volkswagen AG	vw.com
	General Motors Co	gm.com
	Ford Motor Co	ford.com
	Honda Motor Co Ltd	honda.com
Banking	Bank of America Corp	bankofamerica.com
	Wells Fargo & Co	wellsfargo.com
	JPMorgan Chase & Co	jpmorgan.com
	US Bancorp/MN	usbank.com
	capitalone.com	capitalone.com
Beverages	Coca-Cola Co/The	coca-colacompany.com
	Nestle SA	nestle.com
	PepsiCo Inc	pepsi.com
Computer Hardware	Lenovo Group Ltd	lenovo.com
	Hewlett-Packard Co	hp.com
	Acer Inc	acer.com
	Asustek Computer Inc	asus.com
	Apple Inc	apple.com
Department Stores	Macy's Inc	macys.com
	Sears Holdings Corp	sears.com
	Kohl's Corp	kohls.com
	Nordstrom Inc	nordstrom.com
	JC Penny Co Inc	jcpenney.com
IT Services	International Business Machines	ibm.com
	Hewlett-Packard Co	hp.com
	Accenture PLC	accenture.com
	Cap Gemini SA	capgemini.com
	Computer Sciences Corp	csc.com

Table A-1 Industries and companies selected (cont'd)

INDUSTRY	COMPANY NAME	URL
Internet Media	Google Inc	google.com
	Facebook Inc	facebook.com
	Yahoo! Inc	yahoo.com
	Groupon Inc	groupon.com
Life Insurance	Prudential Financial Inc	prudential.com
	AXA SA	axa.com
	Metlife Inc	metlife.com
	UnitedHealth Group	unitedhealthgroup.com
Lodging	Accor SA	novotel.com
	Hilton Worldwide Holdings Inc	hilton.com
	InterContinental Hotels Group	intercontinental.com
	Marriot International Inc/DE	marriott.com
	Wyndham Worldwide Corp	wyndham.com
Mass Merchants	Wal-Mart Stores Inc	walmart.com
	Costco Wholesale Corp	costco.com
	Target Corp	target.com
	Dollar General Corp	dollargeneral.com
	Sears Holdings Corp	sears.com
Restaurants	Starbucks Corp	starbucks.com
	McDonald's Corp	mcdonalds.com
	Darden Restaurants Inc	olivegarden.com
	Bloomin' Brands Inc	outback.com
	Chipotle Mexican Grill Inc	chipotle.com
Retail Discretionary	Home Depot Inc/The	homedepot.com
	Amazon.com	amazon.com
	Lowe's Cos Inc	lowes.com
	Best Buy Co Inc	bestbuy.com
	TJX Cos Inc/The	tjx.com
Software	Microsoft Corp	microsoft.com
	Internationa Business Machines	ibm.com
	Oracle Group	oracle.com
	SAP SE	sap.com
	Symantec Corp	symantec.com
Telecom Carriers	AT&T Inc	att.com
	Verizon Communications Inc	verizonwireless.com
	Nippon Telegraph & Telephone Company	ntt.com
	Deutshe Telekom AG	t-mobile.com
	Vodafone	vodaphone.com

Table A-2 Term definitions according to spyfu.com

Item from Spyfu.com CSV file	Description According to Spyfu.com
term	The keyword term bid on by a domain
position	At the time that we searched, it is the order where this domain's ad appeared on this keyword.
exact_local_daily_search_volume	Average number of times each month that Google users search this keyword, based (by default) on exact match figures in the US. "Local" is specific to US nationwide searches.
exact_global_daily_search_volume	Average number of times each month that Google users search this keyword, based (by default) on exact match figures in the US. "Global" figures include world-wide traffic from Google searches in all countries.
full_url	The exact URL (link) where Google directs you for this result.
broad_cost_per_click	The average amount advertisers pay Google anytime someone clicks their own ad for this keyword. The default is "broad match" costs, but exact match and phrase match are available. By experimenting in the Google Traffic estimator, we find the cost per click that comes closest to our target Avg. Ad Position of 2.0.
phrase_cost_per_click	The average amount advertisers pay Google anytime someone clicks their own ad for this keyword. The default is "broad match" costs, but exact match and phrase match are available. By experimenting in the Google Traffic estimator, we find the cost per click that comes closest to our target Avg. Ad Position of 2.0.
exact_cost_per_click	The average amount advertisers pay Google anytime someone clicks their own ad for this keyword. The default is "broad match" costs, but exact match and phrase match are available. By experimenting in the Google Traffic estimator, we find the cost per click that comes closest to our target Avg. Ad Position of 2.0.
broad_clicks_per_day	The number of clicks your ad may receive each day for this keyword. Each advertiser's expected clicks will vary. By experimenting in the Google Traffic estimator, we find the clicks per day that come closest to our target Avg. Ad Position of 2.0.
phrase_clicks_per_day	The number of clicks your ad may receive each day for this keyword. Each advertiser's expected clicks will vary. By experimenting in the Google Traffic estimator, we find the clicks per day that come closest to our target Avg. Ad Position of 2.0.
exact_clicks_per_day	The number of clicks your ad may receive each day for this keyword. Each advertiser's expected clicks will vary. By experimenting in the Google Traffic estimator, we find the clicks per day that come closest to our target Avg. Ad Position of 2.0.
broad_cost_per_day	This amount is what Google estimates you would spend, on average, to advertise on this keyword each day if you were to aim for an average ad position of 2.0. We multiply the CPC for this keyword by the number of expected clicks.
phrase_cost_per_day	This amount is what Google estimates you would spend, on average, to advertise on this keyword each day if you were to aim for an average ad position of 2.0. We multiply the CPC for this keyword by the number of expected clicks.
exact_cost_per_day	This amount is what Google estimates you would spend, on average, to advertise on this keyword each day if you were to aim for an average ad position of 2.0. We multiply the CPC for this keyword by the number of expected clicks.
max_search_date_id	In Ad History, this coded date tells us which month to assign a particular ad in a domain's activity history. For more, see "Run Date." Search Position: See "Ad Position/Search Position." Run Date: We refresh our data monthly, and "run date" is how we tag that entire month. Something with a run date of 2013/06/01 was collected sometime in June and published at the beginning of July. (Also listed as Search Date)
weighting_rank	This column header is exclusive to Excel/CSV downloads and does not offer insight outside of its context. It serves an internal purpose, helping to keep the default order of paid keywords so that a domain's ten best keywords are at the top.
ad_history_overview	SpyFu's Ad History lets you see all of the keywords that a domain has bought over the last 6+ years, along with the evolution of ad copy they used for each one. Clicking the history button from this grid takes you to the domain's history for this specific keyword.
advertisers	This shows how many unique advertisers have appeared on this keyword in the last 3 months via AdWords.

Table A-3 Descriptive statistics by company

LENOVO	VOLUME	POSITION	ADVERTISERS	MIN CPC	EXP. CPC	MAX CPC
Mean	18.71	5.19	11.58	1.03	1.73	2.91
Median	5.00	4.00	10.00	0.70	1.29	2.08
Mode	1.00	1.00	19.00	0.01	0.01	0.99
Standard Deviation	38.42	4.31	7.36	1.11	1.66	2.99
Kurtosis	24.73	1.67	-1.64	3.62	5.63	27.25
Skewness	4.41	1.38	-0.02	1.76	1.96	3.75
Minimum	1.00	1.00	1.00	0.01	0.01	0.02
Maximum	330.00	21.00	22.00	6.48	12.00	36.54
MACY'S	VOLUME	POSITION	ADVERTISERS	MIN CPC	EXP. CPC	MAX CPC
Mean	11.93	6.70	14.71	0.58	0.97	1.50
Median	3.00	5.00	17.00	0.46	0.79	1.16
Mode	1.00	2.00	19.00	0.02	0.01	1.23
Standard Deviation	28.23	4.96	5.68	0.53	0.85	1.38
Kurtosis	55.40	0.06	-0.39	4.04	35.05	22.33
Skewness	6.51	1.00	-0.99	1.64	3.86	3.64
Minimum	1.00	1.00	1.00	0.01	0.01	0.02
Maximum	330.00	22.00	22.00	3.92	12.11	15.59
PRUDENTIAL	VOLUME	POSITION	ADVERTISERS	MIN CPC	EXP. CPC	MAX CPC
Mean	7.93	8.20	16.03	5.33	8.49	12.87
Median	2.33	7.00	19.00	2.96	6.03	9.23
Mode	1.00	1.00	19.00	0.01	0.01	0.10
Standard Deviation	18.44	5.91	6.29	6.45	9.00	12.96
Kurtosis	84.47	-0.76	-0.47	3.10	11.57	13.82
Skewness	7.80	0.63	-1.01	1.75	2.33	2.76
Minimum	1.00	1.00	1.00	0.01	0.01	0.03
Maximum	270.00	22.00	22.00	35.74	93.96	122.53
UNITED	VOLUME	POSITION	ADVERTISERS	MIN CPC	EXP. CPC	MAX CPC
Mean	7.51	7.53	17.09	0.95	1.55	2.67
Median	3.00	6.00	19.00	0.77	1.31	1.96
Mode	1.00	1.00	19.00	0.01	0.01	1.58
Standard Deviation	11.59	5.60	5.46	0.88	1.28	2.87
Kurtosis	9.35	-0.48	0.51	9.85	15.07	52.09
Skewness	2.95	0.75	-1.29	2.19	2.84	5.49
Minimum	1.00	1.00	1.00	0.01	0.01	0.02
Maximum	63.00	22.00	22.00	8.83	12.53	42.72

Table A-4 Descriptive statistics by campaign for the four companies

	C1 AWARENESS						C2 RESEARCH						C3 DECISION						C4 PURCHASE					
	VOL	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**
LENOVO																								
Mean	21.88	6.18	14.91	$1.26	$2.19	$3.64	9.09	6.37	13.95	$1.48	$2.34	$3.81	22.66	3.61	6.94	$0.66	$1.09	$1.93	18.43	3.16	6.05	$0.35	$0.60	$1.09
Median	5.00	5.00	18.00	$0.95	$1.91	$3.12	3.00	6.00	17.00	$1.19	$1.80	$2.81	7.00	2.00	4.00	$0.47	$0.90	$1.30	7.00	2.00	4.00	$0.18	$0.47	$0.88
Mode	1.00	6.00	19.00	$0.01	$0.01	$2.01	1.00	6.00	19.00	$0.01	$2.67	$1.68	1.00	1.00	2.00	$0.02	$0.18	$2.22	1.00	1.00	2.00	$0.02	$0.10	$0.75
Standard Deviation	50.57	4.34	6.43	$1.21	$1.76	$2.99	16.42	4.08	6.42	$1.24	$1.88	$3.47	33.10	3.98	6.25	$0.68	$0.98	$2.56	26.53	3.53	5.48	$0.43	$0.50	$0.96
Kurtosis	17.38	1.13	-0.80	$2.70	$4.06	$39.60	13.54	1.66	-1.45	$1.27	$4.18	$9.70	7.21	4.84	0.04	$6.65	$9.19	$48.48	4.68	6.53	0.83	$2.77	$0.24	$11.82
Skewness	4.00	1.20	-0.82	$1.52	$1.58	$4.34	3.44	1.21	-0.37	$1.23	$1.74	$2.49	2.45	2.18	1.17	$2.06	$2.33	$5.84	2.23	2.54	1.40	$1.76	$0.93	$2.64
Minimum	1.00	1.00	1.00	$0.01	$0.01	$0.03	1.00	1.00	1.00	$0.01	$0.01	$0.10	1.00	1.00	1.00	$0.01	$0.01	$0.06	1.00	1.00	1.00	$0.01	$0.01	$0.02
Maximum	330.00	20.00	22.00	$6.48	$12.00	$36.54	97.00	20.00	22.00	$5.72	$11.55	$25.63	180.00	21.00	22.00	$4.55	$6.88	$26.71	120.00	18.00	22.00	$2.10	$2.33	$7.21
MACY'S																								
	VOL	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**
Mean	15.44	6.78	15.03	$0.63	$1.08	$1.71	7.59	7.72	15.55	$0.61	$0.96	$1.47	9.67	5.45	13.32	$0.47	$0.79	$1.20	7.49	6.58	13.86	$0.44	$0.72	$1.05
Median	4.00	6.00	17.00	$0.51	$0.91	$1.31	3.00	6.00	17.00	$0.51	$0.78	$1.11	4.00	4.00	16.00	$0.34	$0.62	$0.92	3.00	5.00	16.00	$0.36	$0.63	$0.88
Mode	1.00	2.00	19.00	$0.02	$0.01	$1.53	2.00	4.00	19.00	$0.02	$0.64	$0.86	2.00	3.00	19.00	$0.02	$0.29	$0.67	1.00	3.00	16.00	$0.01	$0.83	$1.69
Standard Deviation	37.20	4.85	5.49	$0.56	$0.97	$1.61	12.98	5.30	5.04	$0.54	$0.74	$1.16	14.78	4.54	6.41	$0.49	$0.71	$1.01	12.36	5.10	5.86	$0.36	$0.45	$0.82
Kurtosis	33.69	-0.07	-0.19	$4.27	$37.59	$20.57	17.23	-0.51	0.88	$1.25	$2.24	$8.37	10.64	1.57	-1.21	$6.45	$17.12	$5.59	9.30	0.74	-0.97	$3.37	$1.42	$20.85
Skewness	5.30	0.91	-1.07	$1.63	$4.27	$3.70	3.84	0.74	-1.37	$1.19	$1.35	$2.25	3.11	1.51	-0.60	$2.09	$3.10	$2.09	3.01	1.28	-0.68	$1.52	$1.07	$3.71
Minimum	1.00	1.00	1.00	$0.01	$0.01	$0.02	1.00	1.00	1.00	$0.01	$0.01	$0.02	1.00	1.00	1.00	$0.01	$0.01	$0.02	1.00	1.00	2.00	$0.01	$0.01	$0.03
Maximum	330.00	20.00	22.00	$3.92	$12.11	$15.59	80.00	22.00	22.00	$2.64	$4.30	$8.76	80.00	20.00	22.00	$3.18	$5.99	$5.99	63.00	21.00	22.00	$1.91	$2.24	$6.17
PRUDENTIAL																								
	VOL	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**
Mean	13.74	6.85	14.77	$1.98	$4.35	$7.94	5.12	9.23	17.37	$7.22	$11.07	$16.18	6.97	6.41	12.17	$3.67	$5.29	$7.66	4.57	7.49	14.31	$7.13	$9.54	$13.53
Median	3.67	6.00	18.00	$0.81	$2.89	$5.33	2.33	8.00	19.00	$4.92	$8.33	$12.54	4.17	5.00	13.50	$2.71	$4.55	$6.58	3.00	6.00	18.00	$5.87	$7.65	$13.48
Mode	1.00	1.00	19.00	$0.01	$0.01	$0.10	1.00	5.00	19.00	$0.10	$0.10	$8.73	1.00	1.00	19.00	$0.01	$0.10	$5.04	1.00	3.00	22.00	$0.01	$0.01	#N/A
Standard Deviation	30.15	5.34	6.41	$2.51	$4.80	$9.53	7.58	6.09	5.54	$7.47	$10.37	$14.53	8.82	5.19	7.60	$3.31	$3.88	$5.06	5.65	5.94	7.40	$5.46	$6.64	$7.80
Kurtosis	32.27	-0.19	-1.03	$1.84	$3.64	$18.35	14.34	-1.02	0.71	$1.17	$9.15	$11.70	7.70	0.10	-1.67	$0.49	-$0.55	$1.35	12.82	-0.42	-1.37	-$0.35	-$0.40	-$0.14
Skewness	5.05	0.90	-0.68	$1.58	$1.74	$3.39	3.42	0.44	-1.44	$1.27	$2.01	$2.53	2.61	0.95	-0.13	$1.06	$0.60	$1.02	3.20	0.90	-0.57	$0.68	$0.60	$0.59
Minimum	1.00	1.00	1.00	$0.01	$0.01	$0.03	1.00	1.00	1.00	$0.01	$0.01	$0.10	1.00	1.00	1.00	$0.01	$0.01	$0.07	1.00	1.00	1.00	$0.01	$0.01	$2.19
Maximum	270.00	22.00	22.00	$11.00	$28.68	$84.73	53.00	22.00	22.00	$35.74	$93.96	$122.53	43.00	21.00	22.00	$12.79	$14.02	$25.82	33.00	20.00	22.00	$20.71	$27.15	$33.12
UNITED																								
	VOL	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**	**VOL**	**POS**	**ADVERTS**	**MIN CPC**	**EXP. CPC**	**MAX CPC**
Mean	8.50	6.57	15.68	$0.79	$1.36	$2.45	6.58	8.85	19.33	$1.12	$1.76	$2.91	2.93	9.30	17.00	$1.32	$1.93	$3.32	3.75	5.67	11.17	$1.85	$2.67	$3.33
Median	3.00	5.00	19.00	$0.63	$1.18	$1.83	3.00	8.00	20.00	$0.95	$1.46	$2.24	2.00	8.00	19.50	$1.06	$1.59	$2.26	4.50	6.00	10.00	$0.83	$1.10	$1.16
Mode	1.00	1.00	19.00	$0.01	$0.01	$0.99	1.00	10.00	22.00	$0.01	$0.01	$1.67	1.00	7.00	22.00	$1.09	#N/A	#N/A	1.00	6.00	10.00	#N/A	#N/A	#N/A
Standard Deviation	12.89	5.41	5.92	$0.73	$1.13	$2.99	9.96	5.63	3.53	$0.90	$1.26	$2.57	2.72	5.75	5.83	$1.12	$1.39	$3.34	2.63	3.77	6.71	$2.68	$3.92	$4.37
Kurtosis	7.12	-0.09	-0.52	$3.97	$17.66	$75.11	12.07	-0.73	6.60	$3.27	$8.22	$14.35	5.63	-0.65	-0.68	$1.03	-$0.88	$3.99	-1.55	-1.52	-1.24	$3.89	$3.27	$0.95
Skewness	2.66	0.97	-0.87	$1.60	$3.03	$7.07	3.20	0.52	-2.49	$1.46	$2.05	$3.19	2.18	0.56	-1.04	$1.23	$0.57	$1.94	0.15	-0.01	0.19	$2.05	$1.99	$1.47
Minimum	1.00	1.00	1.00	$0.01	$0.01	$0.03	1.00	1.00	4.00	$0.01	$0.01	$0.11	1.00	1.00	5.00	$0.04	$0.04	$0.20	1.00	1.00	1.00	$0.02	$0.02	$0.02
Maximum	63.00	22.00	22.00	$4.79	$10.96	$42.72	63.00	22.00	22.00	$5.53	$10.64	$20.55	13.00	20.00	22.00	$4.46	$4.47	$14.60	8.00	11.00	22.00	$8.83	$12.53	$12.56

Table A-5 Average performance across all numeric variables from spyfu.com files

INDUSTRY	COMPANY NAME	URL	MONTHLY BUDGET	# OF KW	PAID CLKS	POSITION	LOCAL SEARCH	GLOBALSEARCH	BROAD CPC	PHRASE CPC	EXACT CPC	BROAD CLKS/DAY	PHRASE CLKS/DAY	EXACT CLKS/DAY	BROAD $/DAY	PHRASE $/DAY	EXACT $/DAY	ADVERTISERS
Airlines	United Continental Holdings	United.com	$61,500.00	21348	22000	7.25	6.52	14.23	$1.70	$1.68	$1.72	1.54	0.35	0.19	$1.92	$0.44	$0.27	16.71
	Delta Air Lnes	delta.com	$57,500.00	7963	12600	6.73	38.45	56.94	$3.19	$3.53	$3.77	2.29	1.15	0.38	$6.53	$2.68	$0.67	14.00
	Deutsche Luftansa AG	lufthansa.com/	$37,000.00	20371	12800	9.39	9.72	31.00	$1.59	$1.47	$1.38	1.51	0.38	0.23	$1.89	$0.45	$0.28	16.74
	Air France-KLM	airfrance.com	$2,040.00	1223	2470	8.63	4.01	17.94	$1.76	$1.43	$1.50	1.88	0.25	0.11	$3.47	$0.44	$0.21	13.57
	American Airlines Group Inc.	aa.com	$73,200.00	28207	31300	7.62	11.19	27.51	$1.40	$1.45	$1.45	2.01	0.78	0.26	$1.67	$0.61	$0.28	15.44
		AVERAGE	**$46,248.00**	**15822.40**	**16234.00**	**7.92**	**13.98**	**29.52**	**$1.93**	**$1.91**	**$1.96**	**1.84**	**0.58**	**0.23**	**$3.10**	**$0.93**	**$0.34**	**15.29**
Apparel, Footwear, Acc Design	Nike Inc.	nike.com	$124,000.00	49927	252000	7.20	19.41	50.38	$1.00	$1.00	$1.14	2.73	1.43	0.40	$2.06	$0.86	$0.30	13.75
	LVMH Moet Hennessy Louis Vuitton	louisvuitton.com/	$6,360.00	1992	7500	4.50	36.32	101.01	$1.29	$1.16	$1.20	2.33	1.23	0.51	$1.56	$0.37	$0.24	13.09
	Adidas AG	adidas .com	$44,000.00	17474	67600	6.40	13.16	40.41	$1.06	$0.99	$1.08	1.79	0.99	0.31	$1.42	$0.60	$0.25	12.93
	Cie Financial Richmont SA	dunhill.com	$1,420.00	286	847	2.41	3.46	18.13	$1.05	$0.87	$0.93	0.39	0.34	0.11	$0.20	$0.13	$0.19	8.05
	Kering	gucci.com/	$21,500.00	7596	40200	5.16	9.07	25.15	$1.01	$1.00	$1.04	1.76	0.71	0.25	$2.02	$0.40	$0.18	14.65
		AVERAGE	**$39,456.00**	**15455.00**	**73629.40**	**5.13**	**16.28**	**47.01**	**$1.08**	**$1.00**	**$1.08**	**1.80**	**0.94**	**0.31**	**$1.45**	**$0.47**	**$0.23**	**12.49**
Automobile OEM	Toyota Motor Corp	toyota.com	$294,000.00	29306	562000	6.18	18.10	45.78	$2.57	$2.31	$2.50	3.46	1.67	0.40	$7.83	$3.45	$1.10	14.41
	Volkswagen AG	vw.com	$183,000.00	39754	255000	4.13	7.99	32.51	$0.89	$0.90	$1.00	1.47	0.90	0.19	$3.47	$1.79	$0.45	8.57
	General Motors Co	gm.com	$16,300.00	1594	14700	2.34	23.50	69.35	$3.51	$3.63	$3.76	1.27	0.40	0.14	$3.69	$1.71	$0.68	5.94
	Ford Motor Co	ford.com	$833,000.00	99175	1160000	6.79	11.68	29.42	$1.62	$1.70	$1.84	3.81	1.56	0.28	$8.06	$3.22	$0.57	12.95
	Honda Motor Co Ltd	honda.com	$256,000.00	35005	462000	7.84	14.13	36.54	$2.21	$2.22	$2.37	3.53	1.40	0.34	$6.59	$2.30	$0.72	14.42
		AVERAGE	**$316,460.00**	**40966.80**	**490740.00**	**5.46**	**15.08**	**42.72**	**$2.16**	**$2.15**	**$2.29**	**2.71**	**1.18**	**0.27**	**$5.93**	**$2.49**	**$0.70**	**11.26**
Banking	Bank of America Corp	bankofamerica.com	$927,000.00	118358	363000	6.43	9.18	12.93	$2.73	$2.92	$2.98	1.10	0.48	0.11	$5.79	$2.34	$0.47	13.12
	Wells Fargo & Co	wellsfargo.com	$1,280,000.00	102922	263000	8.17	11.06	14.64	$4.36	$4.57	$4.61	1.41	0.51	0.12	$7.27	$2.82	$0.63	15.96
	JPMorgan Chase & Co	jpmorgan.com	$16,300.00	1364	11400	3.90	3.12	8.02	$3.53	$3.68	$3.45	0.61	0.22	0.07	$1.71	$0.72	$0.26	11.92
	US Bancorp/MN	usbank.com	$492,000.00	29781	149000	7.58	6.74	9.51	$4.83	$5.18	$5.37	1.90	0.69	0.10	$13.09	$4.71	$0.85	16.75
	capitalone.com	capitalone.com	$1,020,000.00	55168	417000	7.33	13.14	19.86	$5.05	$5.46	$5.64	1.93	0.70	0.19	$11.47	$3.90	$0.82	17.39
		AVERAGE	**$747,060.00**	**61518.60**	**240680.00**	**6.68**	**8.65**	**12.99**	**$4.10**	**$4.36**	**$4.41**	**1.39**	**0.52**	**0.12**	**$7.87**	**$2.90**	**$0.60**	**15.03**
Beverages	Coca-Cola Co/The	coca-colacompany.com	$1,720.00	740.00	1570.00	2.12	9.81	20.29	$1.39	$1.24	$1.24	0.94	0.23	0.16	$0.99	$0.20	$0.15	5.13
	Nestle SA	nestle.com	$467.00	104.00	133.00	6.61	16.97	58.06	$1.64	$1.70	$1.97	2.59	0.79	0.17	$2.78	$0.70	$0.32	11.00
	PepsiCo Inc	pepsi.com	$403.00	32.00	198.00	2.24	27.80	75.33	$0.98	$1.00	$0.95	2.54	1.20	0.26	$0.95	$0.73	$0.11	5.36
		AVERAGE	**$863.33**	**292.00**	**633.67**	**3.66**	**18.19**	**51.22**	**1.34**	**1.31**	**1.39**	**2.03**	**0.74**	**0.20**	**1.57**	**0.54**	**0.19**	**7.16**
Computer Hardware	Lenovo Group Ltd	lenovo.com	$32,800.00	3962	42100	5.43	21.13	87.69	$1.80	$1.65	$1.79	2.77	1.65	0.30	$4.32	$2.31	$0.41	11.99
	Hewlett-Packard Co	hp.com	$769,000.00	92707	468000	3.95	4.93	17.22	$3.41	$3.44	$3.34	1.10	0.48	0.11	$3.95	$1.65	$0.39	11.54
	Acer Inc	acer.com	$45,000.00	6083	42500	5.69	8.76	32.96	$2.06	$2.02	$2.07	3.22	2.51	0.22	$7.42	$6.26	$0.39	12.95
	Asustek Computer Inc	asus.com	$5,730.00	19083	52000	7.21	9.55	35.93	$1.66	$1.53	$1.58	2.00	1.06	0.21	$3.76	$1.95	$0.39	13.53
	Apple Inc	apple.com	$277,000.00	46923	781000	3.58	47.77	157.47	$1.83	$1.78	$1.78	4.36	2.71	0.58	$4.98	$2.89	$0.65	8.79
		AVERAGE	**$225,906.00**	**33751.60**	**277120.00**	**5.17**	**18.43**	**66.25**	**$2.15**	**$2.08**	**$2.11**	**2.69**	**1.68**	**0.28**	**$4.89**	**$3.01**	**$0.45**	**11.76**
Department Stores	Macy's Inc	macys.com	$2,330,000.00	678336	3400000	6.50	11.15	22.74	$0.98	$0.96	$1.03	2.19	1.11	0.32	$1.84	$0.78	$0.27	14.63
	Sears Holdings Corp	sears.com	$2,780,000.00	738496	2960000	6.86	9.71	20.54	$1.10	$1.06	$1.14	2.26	1.17	0.29	$2.35	$1.02	$0.29	15.02
	Kohl's Corp	kohls.com	$2,820,000.00	910274	3550000	6.70	9.59	19.36	$0.92	$0.93	$1.01	2.21	1.17	0.31	$1.80	$0.79	$0.24	15.02
	Nordstrom Inc	nordstrom.com	$912,000.00	289836	839000	8.07	13.17	30.13	$0.99	$0.98	$1.05	2.54	1.40	0.38	$2.13	$0.80	$0.29	14.38
	JC Penny Co Inc	jcpenney.com	$2,180,000.00	573597	3540000	7.51	10.51	20.32	$0.98	$0.97	$1.05	2.41	1.28	0.32	$1.96	$0.84	$0.27	15.53
		AVERAGE	**$2,204,400.00**	**638107.80**	**2857800.00**	**7.13**	**10.83**	**22.62**	**$0.99**	**$0.98**	**$1.05**	**2.32**	**1.23**	**0.32**	**$2.02**	**$0.85**	**$0.27**	**14.92**
IT Services	International Business Machines	ibm.com	$1,660,000.00	126747	385000	5.91	3.66	12.69	$6.95	$7.29	$6.91	1.08	0.42	0.10	$7.39	$3.04	$0.91	12.94
	Hewlett-Packard Co	hp.com	$769,000.00	92707	468000	3.95	4.93	17.22	$3.41	$3.44	$3.34	1.10	0.48	0.11	$3.95	$1.65	$0.39	11.54
	Accenture PLC	accenture.com	$523,000.00	24619	102000	3.77	6.16	20.59	$6.88	$7.22	$7.00	1.30	0.50	0.10	$8.18	$3.33	$0.91	11.85
	Cap Gemini SA	capgemini.com	$147,000.00	12403	23600	8.22	2.84	12.48	$8.08	$8.46	$7.79	1.18	0.42	0.09	$9.01	$3.62	$1.03	13.27
	Computer Sciences Corp	csc.com	$4,950.00	748	1900	7.33	2.86	9.19	$2.70	$3.40	$2.99	1.96	0.52	0.08	$3.98	$1.49	$0.35	12.98
		AVERAGE	**$620,790.00**	**51444.80**	**196100.00**	**5.84**	**4.09**	**14.43**	**$5.60**	**$5.96**	**$5.61**	**1.32**	**0.47**	**0.10**	**$6.50**	**$2.63**	**$0.72**	**12.52**

Table A-5 Average performance across all numeric variables from spyfu.com files (cont'd)

INDUSTRY	COMPANY NAME	URL	MONTHLY BUDGET	# OF KW	PAID CLKS	POSITION	LOCAL SEARCH	GLOBAL SEARCH	BROAD CPC	PHRASE CPC	EXACT CPC	BROAD CLKS/DAY	PHRASE CLKS/DAY	EXACT CLKS/DAY	BROAD $/DAY	PHRASE $/DAY	EXACT $/DAY	ADVERTISERS
Internet Media	Google Inc	google.com	$7,790,000.00	979223	5070000	6.05	23.18	99.06	$3.17	$3.23	$3.19	1.54	0.63	0.14	$4.62	$1.76	$0.42	12.79
	Facebook Inc	facebook.com	$638,000.00	10706	323000	7.03	321.58	1735.63	$2.55	$2.59	$2.63	3.25	1.21	0.23	$6.42	$2.39	$0.48	13.06
	Yahoo! Inc	yahoo.com	$1,100,000.00	1072510	889000	8.76	7.32	18.17	$1.30	$1.26	$1.31	1.30	0.41	0.18	$1.31	$0.43	$0.19	14.28
	Groupon Inc	groupon.com	$1,560,000.00	393133	829000	8.10	9.02	16.28	$2.08	$2.13	$2.24	1.91	0.70	0.22	$3.93	$1.31	$0.42	13.71
		AVERAGE	**$2,772,000.00**	**613893.00**	**1777750.00**	**7.49**	**90.27**	**467.28**	**$2.27**	**$2.30**	**$2.34**	**2.00**	**0.74**	**0.19**	**$4.07**	**$1.47**	**$0.38**	**13.46**
Life Insurance	Prudential Financial Inc	prudential.com	$562,000.00	21766	68300	9.01	4.03	8.28	$6.18	$6.88	$7.30	1.45	0.56	0.10	$12.94	$5.71	$1.02	16.50
	AXA SA	axa.com	$121,000.00	9655	7900	8.47	8.79	15.54	$8.66	$9.94	$10.64	1.67	0.81	0.17	$11.66	$5.34	$1.51	17.22
	Metlife Inc	metlife.com	$2,180,000.00	261870	3640000	6.55	0.75	1.85	$1.89	$2.18	$2.26	0.19	0.09	0.02	$2.72	$1.50	$0.24	20.27
	UnitedHealth Group	unitedhealthgroup.com	$29,100.00	3623	13100	5.68	7.00	11.51	$2.89	$2.95	$2.89	1.34	0.55	0.15	$3.56	$1.20	$0.36	12.37
		AVERAGE	**$723,025.00**	**74228.50**	**932325.00**	**7.43**	**5.14**	**9.30**	**$4.91**	**$5.48**	**$5.77**	**1.16**	**0.50**	**0.11**	**$7.72**	**$3.44**	**$0.78**	**16.59**
Lodging	Accor SA	novotel.com	$17,600.00	4063	3900	9.16	3.34	20.41	$2.37	$2.31	$2.32	1.68	0.51	0.14	$7.09	$2.25	$0.54	16.70
	Hilton Worldwide Holdings Inc	hilton.com	$486,000.00	171226	593000	4.87	7.43	13.06	$1.80	$1.66	$1.78	1.02	0.30	0.16	$1.77	$0.48	$0.30	12.44
	InterContinental Hotels Group	intercontinental.com	$68,200.00	24552	29400	8.05	4.66	14.94	$2.45	$2.19	$2.33	1.00	0.29	0.13	$2.55	$0.74	$0.33	15.97
	Marriot International Inc/DE	marriott.com	$245,000.00	68792	247000	6.93	12.71	20.41	$1.89	$1.93	$2.09	2.06	0.75	0.30	$3.48	$1.06	$0.50	13.89
	Wyndham Worldwide Corp	wyndham.com	$12,500.00	3321	6530	5.29	16.23	21.46	$1.75	$1.73	$1.91	2.88	0.75	0.35	$6.52	$1.61	$0.68	11.52
		AVERAGE	**$165,860.00**	**54390.80**	**175966.00**	**6.86**	**8.87**	**18.05**	**$2.05**	**$1.96**	**$2.09**	**1.73**	**0.52**	**0.22**	**$4.28**	**$1.23**	**$0.47**	**14.10**
Mass Merchants	Wal-Mart Stores Inc	walmart.com	$5,440,000.00	1872147	5580000	6.47	15.65	38.49	$1.10	$1.06	$1.11	2.82	1.55	0.41	$2.97	$1.31	$0.35	12.79
	Costco Wholesale Corp	costco.com	$525.00	407	934	5.45	18.22	26.31	$3.00	$3.36	$3.79	3.95	1.42	0.35	$20.49	$8.43	$2.16	10.83
	Target Corp	target.com	$5,160,000.00	2111723	5540000	6.73	13.89	34.04	$1.01	$0.97	$1.03	3.24	1.67	0.44	$2.97	$1.23	$0.35	13.29
	Dollar General Corp	dollargeneral.com	$20,200.00	4259	16700	7.29	19.18	27.27	$1.09	$1.07	$1.15	3.00	1.45	0.41	$2.38	$1.04	$0.36	15.05
	Sears Holdings Corp	sears.com	$2,780,000.00	738496	2960000	6.81	10.72	22.62	$1.12	$1.07	$1.15	2.54	1.32	0.31	$2.68	$1.16	$0.32	15.05
		AVERAGE	**$2,680,145.00**	**945406.40**	**2819526.80**	**6.55**	**15.53**	**29.75**	**$1.46**	**$1.51**	**$1.65**	**3.11**	**1.48**	**0.38**	**$6.30**	**$2.63**	**$0.71**	**13.40**
Restaurants	Starbucks Corp	starbucks.com	$41,400.00	8047	49000	5.47	28.90	44.17	$1.42	$1.41	$1.54	2.06	1.01	0.38	$2.56	$0.92	$0.33	12.28
	McDonald's Corp	mcdonalds.com	$11,500.00	4169	23800	3.59	58.29	138.92	$1.35	$1.23	$1.28	2.56	1.35	0.35	$4.48	$1.76	$0.34	9.09
	Darden Restaurants Inc	olivegarden.com	$19,700.00	3172	47100	4.62	50.84	66.07	$1.21	$1.32	$1.38	4.07	2.40	0.54	$4.76	$1.57	$0.33	8.75
	Bloomin' Brands Inc	outback.com	$29,200.00	2409	16000	3.35	50.59	62.42	$1.36	$1.42	$1.48	4.20	5.03	0.71	$5.52	$4.82	$0.35	8.09
	Chipotle Mexican Grill Inc	chipotle.com	$46,400.00	6151	38400	5.39	21.43	29.40	$1.95	$1.86	$1.91	2.46	1.05	0.17	$3.20	$1.08	$0.22	11.17
		AVERAGE	**$29,640.00**	**4789.60**	**34860.00**	**4.48**	**42.01**	**68.19**	**$1.46**	**$1.45**	**$1.52**	**3.07**	**2.17**	**0.43**	**$4.10**	**$2.03**	**$0.31**	**9.88**
Retail Discretionary	Home Depot Inc/The	homedepot.com	$3,010,000.00	406760	2520000	5.48	8.60	14.54	$1.59	$1.67	$1.82	2.31	1.06	0.25	$4.05	$1.62	$0.38	14.17
	Amazon.com	amazon.com	$14,500,000.00	4972407	24400000	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Lowe's Cos Inc	lowes.com	$3,290,000.00	583339	3150000	5.63	7.89	14.57	$1.40	$1.46	$1.58	2.25	1.04	0.25	$3.27	$1.28	$0.35	14.11
	Best Buy Co Inc	bestbuy.com	$1,140,000.00	306290	1780000	4.40	13.14	38.31	$1.21	$1.16	$1.18	1.52	0.80	0.24	$1.75	$0.74	$0.22	11.66
	TJX Cos Inc/The	tjx.com	$213,000.00	64951	295000	7.83	14.07	26.15	$0.88	$0.90	$0.96	3.29	1.70	0.44	$2.17	$0.79	$0.30	15.51
		AVERAGE	**$4,430,600.00**	**1266749.40**	**6429000.00**	**5.83**	**10.92**	**23.40**	**$1.27**	**$1.30**	**$1.38**	**2.34**	**1.15**	**0.30**	**$2.81**	**$1.11**	**$0.31**	**13.86**
Software	Microsoft Corp	microsoft.com	$2,070,000.00	193835	1200000	5.59	8.84	47.13	$5.88	$5.98	$5.67	1.49	0.71	0.13	$6.60	$2.59	$0.66	12.78
	Internationa Business Machines	ibm.com	$1,660,000.00	126747	385000	5.91	3.66	12.69	$6.95	$7.29	$6.91	1.08	0.42	0.10	$7.39	$3.04	$0.91	12.94
	Oracle Group	oracle.com	$438,000.00	28724	159000	6.50	5.86	21.60	$7.00	$7.44	$7.29	1.70	0.72	0.12	$12.71	$5.52	$1.24	12.85
	SAP SE	sap.com	$517,000.00	30169	93000	5.73	3.76	14.89	$7.38	$7.71	$7.22	1.25	0.47	0.09	$9.08	$3.50	$0.87	13.53
	Symantec Corp	symantec.com	$309,000.00	18074	60100	6.05	2.85	10.18	$8.60	$9.29	$9.11	0.98	0.43	0.08	$7.65	$3.87	$0.89	13.43
		AVERAGE	**$998,800.00**	**79509.80**	**379420.00**	**5.96**	**4.99**	**21.30**	**$7.16**	**$7.54**	**$7.24**	**1.30**	**0.55**	**0.10**	**$8.69**	**$3.70**	**$0.91**	**13.11**
Telecom Carriers	AT&T Inc	att.com	$2,170,000.00	152103	947000	4.62	13.06	45.62	$4.62	$4.97	$4.94	1.91	0.84	0.22	$8.21	$3.13	$0.79	12.27
	Verizon Communications Inc	verizonwireless.com	$567,000.00	65504	346000	5.35	26.85	69.76	$3.18	$3.32	$3.36	2.41	1.36	0.33	$6.47	$3.27	$0.63	12.57
	Nippon Telegraph & Telephone Company	ntt.com	$32,400.00	2952	6490	6.56	6.03	38.68	$6.45	$6.82	$6.13	1.01	0.41	0.09	$8.66	$4.15	$1.13	12.06
	Deutshe Telekom AG	t-mobile.com	$314,000.00	19248	272000	3.76	62.34	202.41	$2.16	$2.22	$2.19	3.11	2.38	0.86	$5.54	$2.48	$2.28	10.05
	Vodafone	vodaphone.com	$11,700.00	244	2900	6.91	6.56	23.03	$17.05	$23.21	$23.50	4.49	1.54	0.32	$59.61	$19.96	$5.52	18.57
		AVERAGE	**$619,020.00**	**48010.20**	**314878.00**	**5.44**	**22.97**	**75.90**	**$6.69**	**$8.11**	**$8.02**	**2.59**	**1.31**	**0.37**	**$17.70**	**$6.60**	**$2.07**	**13.10**

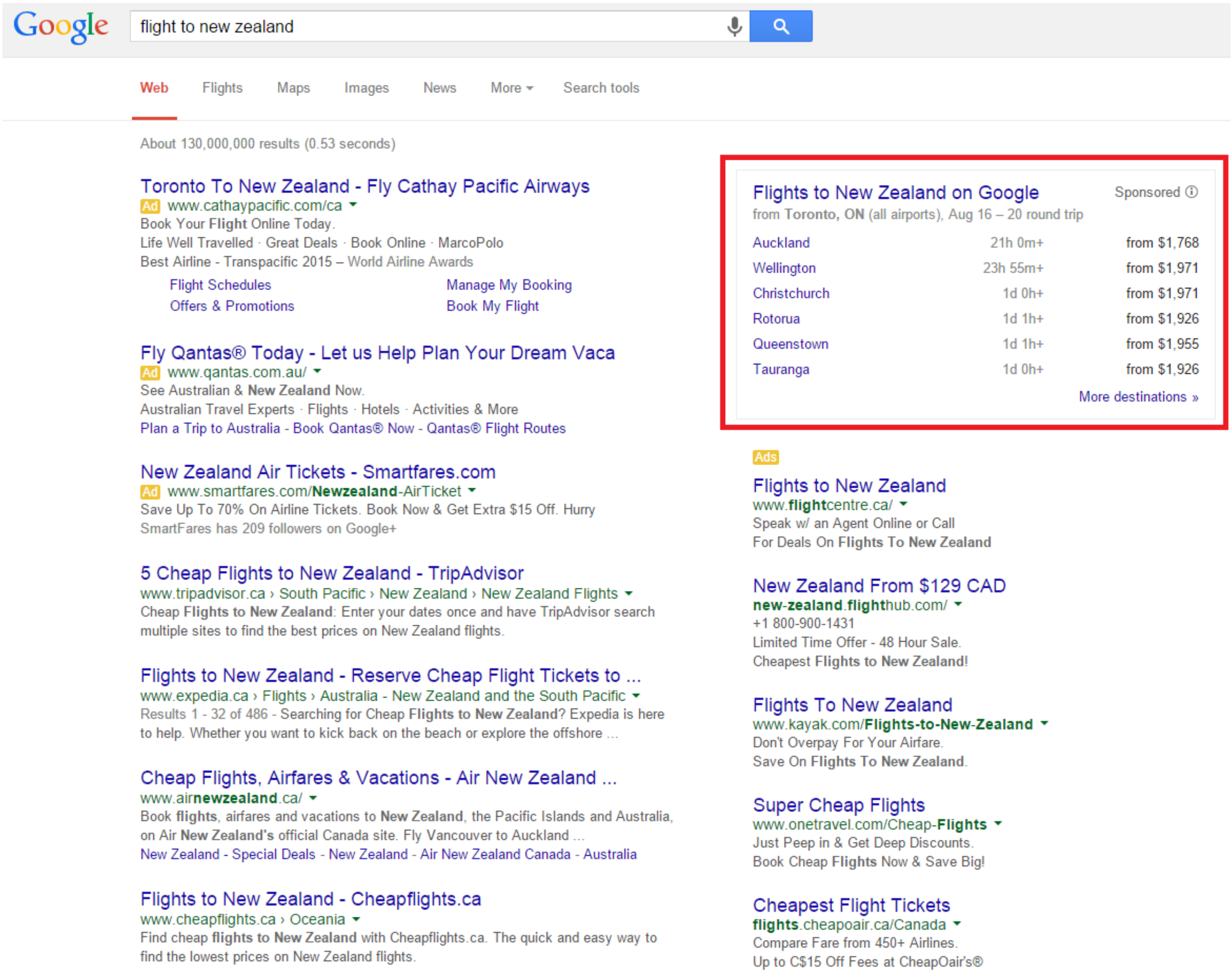

Figure A-1 Google's search results for the term "flight to new zealand"

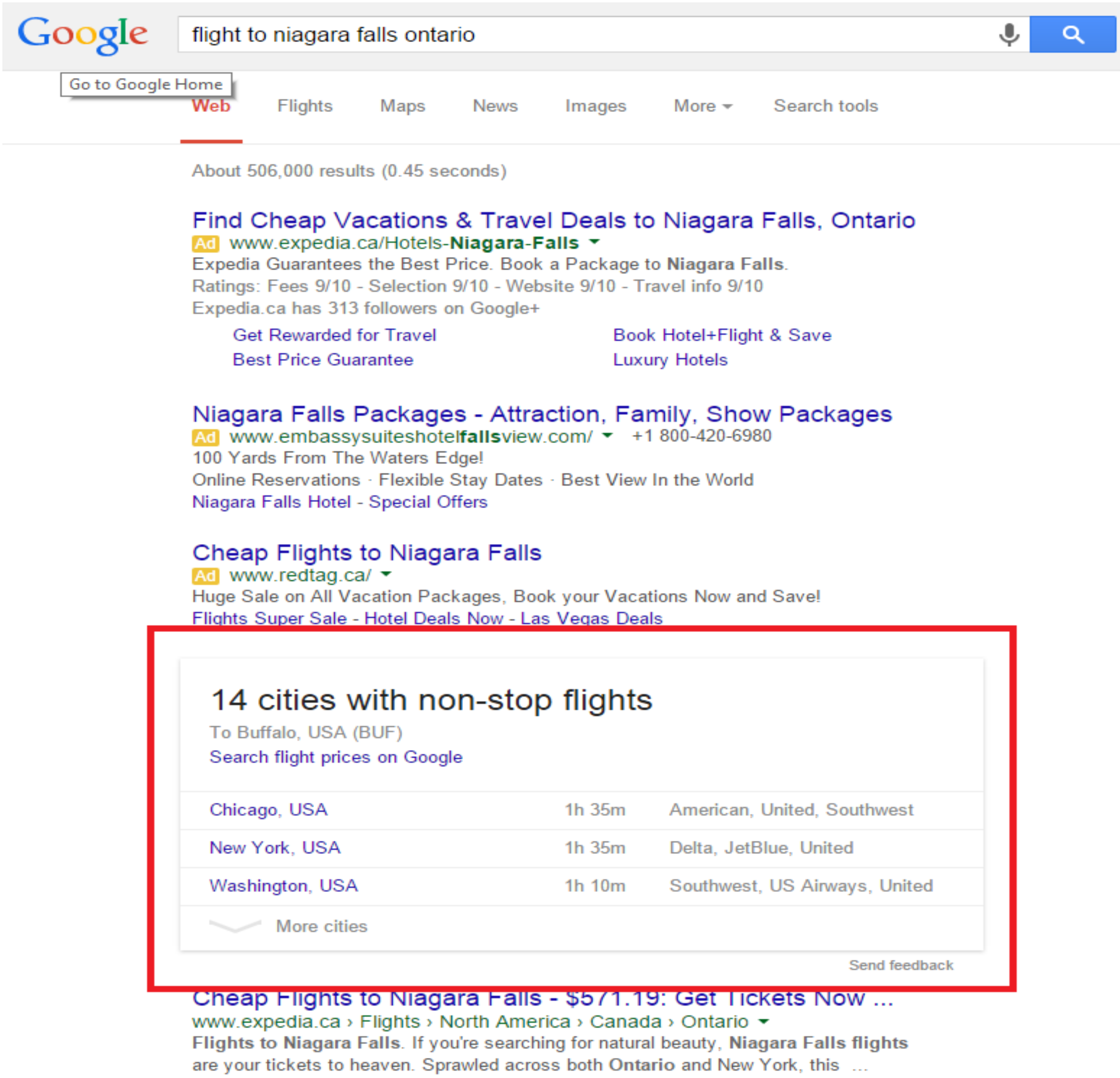

Figure A-2 Google's search results for the term "flight to niagara falls ontario"

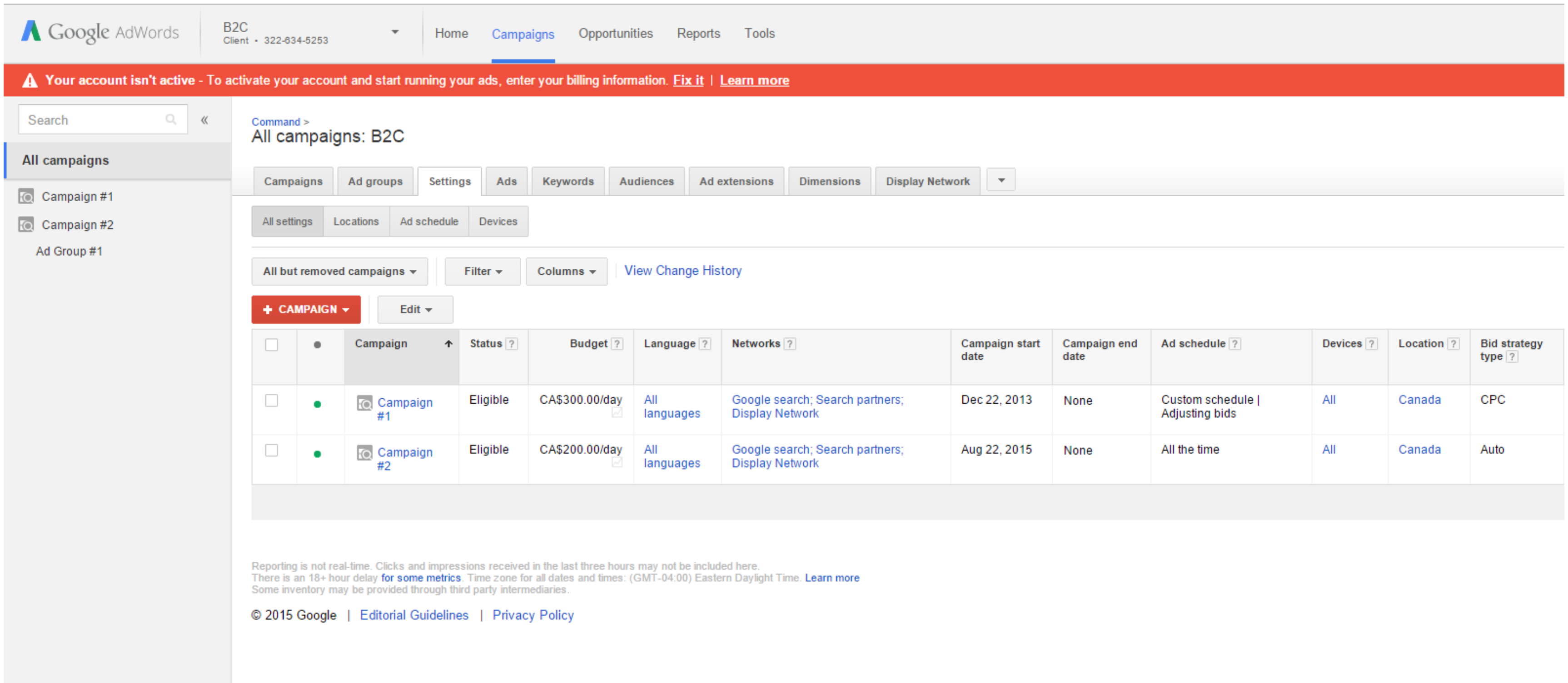

Figure A-3 Budget setting example in an AdWords account

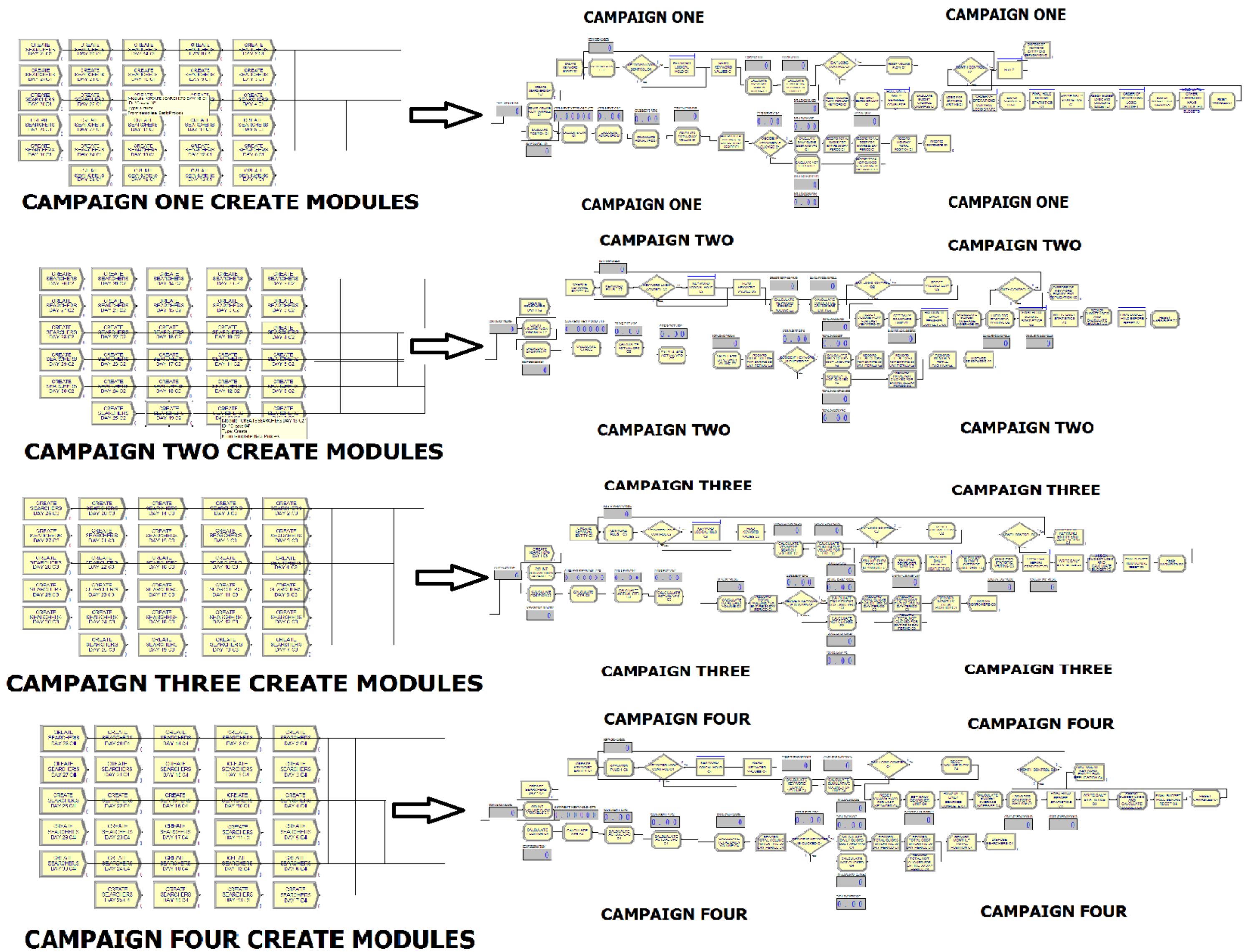

Figure A-4 The simulation model in Arena

Appendix B Statistical Results

Table B-1 Lenovo's paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	563.773617	3000	208.1185846	3.7997081
	BASELINE	3162.490418	3000	223.7141683	4.0844432
Pair 2	VOLUME	937.052155	3000	237.0204206	4.3273810
	BASELINE	3162.490418	3000	223.7141683	4.0844432
Pair 3	CLICKS	1774.786289	3000	293.2919086	5.3547531
	BASELINE	3162.490418	3000	223.7141683	4.0844432

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - BASELINE	-2598.7168018	177.4829420	3.2403804	-2605.0703948	-2592.3632088	-801.979	2999	.000
Pair 2	VOLUME - BASELINE	-2225.4382630	126.1951245	2.3039972	-2229.9558377	-2220.9206882	-965.903	2999	.000
Pair 3	CLICKS - BASELINE	-1387.7041290	295.6303537	5.3974471	-1398.2872021	-1377.1210558	-257.104	2999	.000

Table B-2 Macy's paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	126.654855	3000	57.9899389	1.0587466
	BASELINE	1215.555309	3000	60.5844881	1.1061164
Pair 2	VOLUME	139.938483	3000	51.0900424	.9327723
	BASELINE	1215.555309	3000	60.5844881	1.1061164
Pair 3	CLCIKS	185.596566	3000	60.0531336	1.0964152
	BASELINE	1215.555309	3000	60.5844881	1.1061164

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - BASELINE	-1088.9004532	104.6036251	1.9097922	-1092.6450884	-1085.1558181	-570.167	2999	.000
Pair 2	VOLUME - BASELINE	-1075.6168258	89.4299984	1.6327609	-1078.8182704	-1072.4153811	-658.772	2999	.000
Pair 3	CLCIKS - BASELINE	-1029.9587427	63.5116918	1.1595595	-1032.2323552	-1027.6851301	-888.233	2999	.000

Table B-3 Prudential's paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	2987.809144	3000	395.6768926	7.2240387
	BASELINE	5149.745939	3000	371.4453078	6.7816325
Pair 2	VOLUME	3225.030230	3000	394.8856517	7.2095926
	BASELINE	5149.745939	3000	371.4453078	6.7816325
Pair 3	CLCIKS	3638.497594	3000	491.4641984	8.9728676
	BASELINE	5149.745939	3000	371.4453078	6.7816325

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - BASELINE	-2161.9367959	736.3978223	13.4447233	-2188.2986086	-2135.5749832	-160.802	2999	.000
Pair 2	VOLUME - BASELINE	-1924.7157099	746.9800687	13.6379278	-1951.4563493	-1897.9750704	-141.130	2999	.000
Pair 3	CLCIKS - BASELINE	-1511.2483452	685.8556340	12.5219534	-1535.8008320	-1486.6958585	-120.688	2999	.000

Table B-4 United Continental Holdings' paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	166.201656	3000	87.5421271	1.5982933
	BASELINE	2099.244377	3000	99.6758427	1.8198236
Pair 2	VOLUME	247.340077	3000	104.4977689	1.9078595
	BASELINE	2099.244377	3000	99.6758427	1.8198236
Pair 3	CLCIKS	170.546420	3000	81.0068616	1.4789762
	BASELINE	2099.244377	3000	99.6758427	1.8198236

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - BASELINE	-1933.0427210	90.7584468	1.6570150	-1936.2917219	-1929.7937201	-1166.581	2999	.000
Pair 2	VOLUME - BASELINE	-1851.9043002	56.0287747	1.0229408	-1853.9100368	-1849.8985636	-1810.373	2999	.000
Pair 3	CLCIKS - BASELINE	-1928.6979573	124.1169367	2.2660549	-1933.1411364	-1924.2547781	-851.126	2999	.000

Table B-5 Lenovo's additional paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	563.773617	3000	208.1185846	3.7997081
	VOLUME	937.052155	3000	237.0204206	4.3273810
Pair 2	COST	563.773617	3000	208.1185846	3.7997081
	CLICKS	1774.786289	3000	293.2919086	5.3547531
Pair 3	VOLUME	937.052155	3000	237.0204206	4.3273810
	CLICKS	1774.786289	3000	293.2919086	5.3547531

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - VOLUME	-373.2785388	155.2008833	2.8335675	-378.8344714	-367.7226063	-131.734	2999	.000
Pair 2	COST - CLICKS	-1211.0126728	279.9949621	5.1119852	-1221.0360250	-1200.9893206	-236.897	2999	.000
Pair 3	VOLUME - CLICKS	-837.7341340	270.0263269	4.9299837	-847.4006257	-828.0676423	-169.926	2999	.000

Table B-6 Macy's additional paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	126.654855	3000	57.9899389	1.0587466
	VOLUME	139.938483	3000	51.0900424	.9327723
Pair 2	COST	126.654855	3000	57.9899389	1.0587466
	CLICKS	185.596566	3000	60.0531336	1.0964152
Pair 3	VOLUME	139.938483	3000	51.0900424	.9327723
	CLICKS	185.596566	3000	60.0531336	1.0964152

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - VOLUME	-13.2836275	40.4637072	.7387628	-14.7321606	-11.8350943	-17.981	2999	.000
Pair 2	COST - CLICKS	-58.9417106	79.3846982	1.4493597	-61.7835502	-56.0998709	-40.667	2999	.000
Pair 3	VOLUME - CLICKS	-45.6580831	63.7865051	1.1645769	-47.9415335	-43.3746327	-39.206	2999	.000

Table B-7 Prudential's additional paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	2987.809144	3000	395.6768926	7.2240387
	VOLUME	3225.030230	3000	394.8856517	7.2095926
Pair 2	COST	2987.809144	3000	395.6768926	7.2240387
	CLICKS	3638.497594	3000	491.4641984	8.9728676
Pair 3	VOLUME	3225.030230	3000	394.8856517	7.2095926
	CLICKS	3638.497594	3000	491.4641984	8.9728676

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - VOLUME	-237.2210860	134.3011545	2.4519924	-242.0288432	-232.4133289	-96.746	2999	.000
Pair 2	COST - CLICKS	-650.6884507	512.5036162	9.3569930	-669.0352245	-632.3416768	-69.540	2999	.000
Pair 3	VOLUME - CLICKS	-413.4673646	502.3796662	9.1721559	-431.4517180	-395.4830112	-45.079	2999	.000

Table B-8 United Continental Holdings' additional paired t-test results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COST	166.201656	3000	87.5421271	1.5982933
	VOLUME	247.340077	3000	104.4977689	1.9078595
Pair 2	COST	166.201656	3000	87.5421271	1.5982933
	CLICKS	170.546420	3000	81.0068616	1.4789762
Pair 3	VOLUME	247.340077	3000	104.4977689	1.9078595
	CLICKS	170.546420	3000	81.0068616	1.4789762

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COST - VOLUME	-81.1384208	81.6987882	1.4916090	-84.0631010	-78.2137406	-54.397	2999	.000
Pair 2	COST - CLICKS	-4.3447638	65.1609333	1.1896704	-6.6774164	-2.0121111	-3.652	2999	.000
Pair 3	VOLUME - CLICKS	76.7936571	119.0145772	2.1728990	72.5331339	81.0541802	35.342	2999	.000