# Measuring Causal Position Effects in Search Advertising: A Regression Discontinuity Approach


Sridhar Narayanan — Stanford University

Kirthi Kalyanam — Santa Clara University*


July 28, 2011

Preliminary draft. Please do not circulate or cite without permission.


**Abstract**

In this paper, we investigate the causal effect of position in search engine advertising listings on outcomes such as click-through rates and sales orders. Since the position is typically determined through an auction, there are significant selection issues in measuring position effects. Correlational results are likely to be biased due to the selection in position induced by strategic bidding by advertisers. Additionally, experimentation is rendered difficult in this situation by competitors' bidding behavior, which induces selection biases that cannot be eliminated by randomizing the bids for the focal advertiser. We show that a regression discontinuity approach is a feasible approach to measure causal effects in this important context, where other approaches to obtaining causal effects are typically infeasible. and apply it to a dataset of 23.7 million daily observations containing information on bids, search advertising results and linked outcomes. Our data set is unique in that it contains information not only for the firm but also its major competitors who are advertising in the same category. We find in our empirical application that causal position effects are significantly underestimated if the selection of position is ignored. The data set also contains information on two advertising targeting options offered by Google: Exact and Broad match and our causal estimates provide insights into the value of this type of semantic targeting. We find important differences in the effects of position for exact vs. broad match keywords, with exact match keywords showing strong position effects at the top most position, and broad match keywords having strong position effects only lower down.We are also able to study weekday and weekend effects and find that position effects are lower on the weekend than on weekdays.



*Emails: sridhar.narayanan@stanford.edu and kkalyanam@gmail.com.


We are grateful to an anonymous data provider for sharing the data with us and for useful discussions. All errors are our own

# 1 Introduction

Search advertising, which refers to paid listings on search engines such as Google, Bing and Yahoo, has emerged in the last few years to be an important and growing part of the advertising market. These advertisements can typically be identified as such on the page either by their location, or highlights or both. The order in which these paid listings are served is typically determined through a keyword auction, with advertisers placing bids on the amount paid to the search engine for each click by the consumer. Advertisers bid to get specific positions in these listings, with higher positions costing more than lower positions. It is therefore crucial to understand what the effect of position in search advertising listings is on outcomes such as click-through rates and sales. The measurement of causal position effects is challenging due to the fact that position is not randomly determined, but is rather the outcome of strategic actions by competing advertisers. Correlational inferences of position effects are thus potentially misleading due to selection biases. Further, experimentation is rendered difficult in this context, since randomization of a focal advertiser's bids in the absence of randomization of competitors' bids is insufficient to get valid causal effects. In this paper, we present a regression discontinuity approach to identifying causal position effects, and apply it to a unique dataset of search advertising results and website outcomes that has information on the bids of the focal advertiser as well as its major competitors.

Search advertising has been the focus of a significant stream of literature in multiple fields including marketing, economics and information systems. The effect of position in search advertising has been specifically of interest in this literature. Position in the search advertising listings is the main decision variable for the advertiser, given the limited ability to vary the content of the advertisement itself. Position could affect consumer click-through and purchase behavior through multiple mechanisms, including signaling (Nelson, 1974; Kihlstrom and Riordan, 1984), consumer expectations about the advertisements being ordered on the basis of relevance (Varian, 2007), sequential search (Weitzman, 1979) and behavioral mechanisms such as attention (Hotchkiss, Alston, and Edwards, 2005; Guan and Cutrell, 2007). One or more of these mechanisms could simultaneously be at play, leading to position effects of search advertising. Several empirical studies have documented the relationship between position and behavioral outcomes such as click-through rates, conversion rates and sales (Agarwal, Hosanagar, and Smith, 2007; Ghose and Yang, 2009; Kalyanam, Borle, and Boatwright, 2010; Yang and Ghose, 2010; Rutz and Trusov, 2011).

Measuring causal effects in this context is challenging due to the lack of experimental variation in position in search advertising listings. This is because position is determined through an online auction, with competing advertisers biddings for their advertisements to appear in the listings. This leads to the position being endogenous. Past studies have tried to address this issue either by conducting experiments in which bids for the focal firm are randomized, in order to establish robustness of the results (Agarwal, Hosanagar, and Smith, 2007), or by accounting for the potential endogeneity of position through a parametric approach (Ghose and Yang, 2009; Rutz and Trusov, 2011; Kalyanam, Borle, and Boatwright, 2010). Experimentation is difficult in this context, since randomization of bids of the focal advertiser is insufficient to achieve randomization of position. This is because position is a function not just of the bids of the focal advertiser, but those of competing firms as well. Thus, while randomization of bids might eliminate the issues of selection for the focal advertiser, the selection induced by competitors' strategic bidding behavior is not eliminated, and one cannot make *causal* inferences based on such an experiment. Instruments are difficult if not impossible to find in this context, since demand side factors as well as cost side variables that are correlated with position are likely to be inputs into firms' strategic bidding decisions and are hence not valid instruments. Parametric approaches are also likely problematic in this context, since there is a set of highly complex processes through which position is determined, and the use of an incorrect specification would lead to unpredictable biases in estimates of position effects. Thus, while there is an extant literature on position effects in search advertising, there is still a gap in the literature in finding causal position effects. Advertisers are also interested in finding a robust and easily implementable approach to finding causal estimates of positions.

We present a regression discontinuity approach to finding causal position effects in search advertising. The regression discontinuity (RD) design was first developed in the program evaluation literature (Thistlethwaite and Campbell, 1960; Cook and Campbell, 1979; Shadish, Cook, and Campbell, 2001) and its econometric properties formalized by Hahn, Todd, and van der Klaauw

(2001). It has been applied to the measurement of causal treatment effects in a variety of domains (see Imbens and Lemieux (2008); Lee and Lemeiux (2010); van der Klaauw (2008) for recent reviews of the literature). A recent literature has applied RD to measuring promotional effects (Busse, Silva-Risso, and Zettelmeyer, 2006; Busse, Simester, and Zettelmeyer, 2010; Nair, Hartmann, and Narayanan, 2011). RD measures causal treatment effects in situations where treatment is based on whether an underlying score variable crosses a threshold. With the treatment being the only discontinuity at the threshold, a discontinuous jump in the outcome of interest at the threshold is the treatment effect. Thus, RD measures the treatment effect as the difference between the limiting values of the outcome on the two sides of the threshold.

In the case of search engine advertising, the position is the outcome of an auction conducted by the search engine. In the typical auction, for instance that of Google, the advertisers are ranked on a score called *AdRank*, which is a function of the advertisers' bids and a measure given by the search engine that is termed *Quality Score* (Varian, 2007).[1] This leads to a viable RD design to measure the causal effect of a movement from one position to the adjacent one. Considering the higher position as the treatment, the score is the difference in the *AdRanks* for the bidders in the higher and lower positions. If this score crosses 0, there is treatment, otherwise not. Thus, the RD estimator of the effect of position finds the limiting values of the outcome of interest (e.g. click through rates or sales) on the two sides of this threshold of 0.

While the search engine observes the *AdRanks* of all the bidders, the bidders themselves only observe their own *AdRanks*. They observe their own bids, and the search engine reports the *Quality Score* to them *ex-post* and they can hence construct their own *AdRanks*, but they do not observe the bids or *Quality Scores* of their competitors. Since the score for the RD is the difference between competing bidders *AdRanks*, they cannot construct the score. This ensures the quasi randomization required for the RD design. At the same time, it also poses a challenge to the empirical researcher, who only typically observes the *AdRank* for one advertiser and hence cannot construct the score either. However, we have obtained a unique dataset that allows us to observe Bids and *AdRanks*

[1]Other search engines such as Bing have a similar mechanism to decide the position of the advertisement. Our empirical application uses data for advertisements at Google, which is also the largest search engine in terms of market share. Hence, the rest of the discussion will focus primarily on Google.

for competing bidders. Specifically, our data consist of advertising position and performance data from an online retailer of a particular category of consumer durables that recently acquired three of its major competitors. All of these firms were major advertisers on the Google search engine, and we have a large number of observations where pairs of firms were in adjacent positions. We have historical information from these firms for a period when they operated as independent firms, with independent advertising strategies. Thus, for a large number of observations, we have *AdRanks* and performance measures for advertisers in adjacent positions. We are thus able to implement a valid RD design to measure the treatment effects. This situation is similar to the type of data that would be available to a search engine, which can report causal position effects to the advertiser.

We estimate the effect of position on two main outcomes of interest - click through rates and sales orders (i.e. whether the consumer who clicked on the search advertisement purchased at that or a subsequent occasion). We control for the keyword, advertiser, day of week and advertisement match-type[2] to ensure that our effects are not contaminated by cross-sectional selection biases. We find that position positively affects click-through rates, with higher positions getting greater clicks. However, these effects are not linear, with the top most position seeing a significant effect, the next two being insignificantly different from each other, and again significant effects when moving below the top three positions. Further, we find that the correlational results significantly underestimate the effect of position, suggesting a negative selection bias in the case of these data. The effect of position on sales orders is positive and highly significant when moving from position 6 to 5, but insignificant otherwise. We also investigate the differences in these effects between two different types of advertisements - an exact match-type where the advertisement is served when the consumer types in the exact keyword phrase that the advertiser has bid on, and a broad match-type, where the advertisement is served for any search phrase that contains the keyword phrase the advertiser has bid on. We find that while the position effects for the broad match-type mirror the pooled results, exact match type shows much stronger effects with respect to position 1 but are insignificant for other positions. Finally, we compare the effects for weekdays and weekends, and find that click

[2]There are two match-types - exact, where the ads are served when the consumer's search phrase exactly matches the advertisement keyword phrase; and broad, where the ad is served as long as the keyword phrase is present in the consumer's search phrase, though they do not necessarily match exactly.

through rates are significantly lower for weekends than weekdays and position effects are weaker. Importantly, such a comparison would be missed by the correlational estimates.

This paper makes several contributions to the literature. First, this paper tries to make causal inferences on position effects, which have been hard to make in the literature so far due to limitations in the data and in the empirical strategies used. Second, it demonstrates the nature of the selection bias that can result in these contexts, both in a cross-sectional sense and even within a specific advertisement. Third, it documents important differences in these effects for advertisements of broad vs. exact match types, and weekdays and weekends, which have not been documented before. Fourth, it is a novel application of regression discontinuity to an important context where it has not been considered before, and where other ways of obtaining causal effects are typically infeasible.

The rest of the paper is organized as follows. We give a background to search advertising in general and position effects in particular in section 2. In section 3, we discuss the selection of position in search advertising contexts, and extant approaches to deal with the issue. In section 4, we discuss a regression discontinuity approach to deal with the issue of selection of position and to find causal effects. We discuss the data in section 5 and the results of our empirical analysis in section6. Finally, we conclude in section 7.

# 2 Background on search advertising

## 2.1 Overview of search advertising

Search advertising involves placing text ads on the top or side of the search results page on search engines. An example of the results of a search for the phrase "golf clubs" on Google, which is the most popular search engine, is shown in figure 1. Search advertising is a large and rapidly growing market. For instance, Google reported revenues of almost \$8.5 billion for the quarter ending December 31, 2011, with a growth of 26% over the same period in the previous year. The revenues from Google's sites, primarily the search engine, accounted for two-thirds of these revenues.[3] According to the Internet Advertising Bureau, \$12 billion was spent in the United States alone on search advertising

[3]These data were obtained from Google's earnings report for Q4 2010, available at http://investor.google.com/earnings/2010/Q4_google_earnings.html (last accessed on July 20, 2011).

in 2010. Search advertising is the largest component of the online advertising market, with 46% of all online advertising revenues in 2010. Despite the fact that it is a relatively new medium for advertising, it already accounted for over 9% of total advertising spending (at about $131 billion for 2010) and grew at a faster rate than the industry as a whole (12% vs. 6.5% in 2010).[4]

Several features of search advertising have made it a very popular online advertising format. Search ads are triggered by specific keywords (search phrases). For example consider an advertiser who is selling health insurance for families. Some of the search phrases related to health insurance could include "health insurance", "family health insurance", "discount health insurance" and "California health insurance". The advertiser can specify that an ad will be shown only for the phrase "family health insurance". Further, these ads can be geography specific, with potentially different ads being served in different locations. This enables an advertiser to obtain a high level of targeting if desired.

Search advertising is sold on a performance basis, with advertisers bidding on keyword phrases. The search engine conducts an automated online auction for each keyword phrase on a regular basis, with the set of ads and their order being decided by the outcome of the auction. Advertisers only pay the search engine if a user clicks on an ad and the payment is on a per click basis (hence the commonly used term - PPC or pay per click for search advertising). By contrast, online display advertising is sold on the basis of impressions, so the advertiser pays even if there is no behavioral response. In search advertising advertisers are able to connect the online ad to the specific online order it generated by matching cookies. The combination of targeting, pay for clicks and sales tracking make the sales impact of search advertising highly measurable. This creates strong feedback loops as advertisers track performance in real time and rapidly adjust their spending.

Advertisers are interested in measuring the causal effect of position in the context of search advertising, since it is one of the key levers available to the advertiser. Unlike many other forms of advertising, there is limited scope in adjusting the ad copy, since there are only a couple of lines of text available for the advertisement, without the possibility of images or other types of media. Advertisers are interested in the causal effects of position on outcomes such as click through rates

[4] Internet Advertising Bureau's report on internet advertising can be accessed at http://www.iab.net/media/file/IAB_Full_year_2010_0413_Final.pdf (last accessed on July 15, 2011).

(CTRs), conversion rates (the proportion of clicks that resulted in a sales order), orders and average sales value per click.

Before we move on to position effects, we discuss the auction mechanism by which search engines such as Google decide positions of advertisers. Advertisers bid on keywords, with the bid consisting of the amount that the advertiser would pay the search engine every time a consumer clicked on the advertiser's search ad. Since the search engine gets paid on a per click basis, the search engine's revenue would be maximized if the winning bidder has higher product of bid and clicks. Thus, Google ranks bidder not on their bids, but on a score called *AdRank*, which is the product of bid and a score called *Quality Score (QS)* assigned by Google. While the exact procedure by which Google assigns a $QS$ to a particular ad is not publicly revealed, it is known that it is primarily a function of expected click through rates (which Google knows through historical information combined with limited experimentation), adjusted up or down a little bit by factors such as the quality of the landing page of the advertiser. The positions of the search ads of the winning bidders is then based in descending order of their *AdRanks.* The winning bidder pays an amount that is just above what would be needed to win that bid. Thus, the cost per click of the winning bidder in position $i$ is given by

$$CPC_i = \frac{Bid_{i+1} \times QS_{i+1}}{QS_i} + \varepsilon \tag{1}$$

where $\varepsilon$ denotes a very small number.

## 2.2 Position effects

One of the most important issues in search advertising is the position of the ad on the page. Since the position of an ad is the outcome of an auction, higher positions cost more for the advertiser, everything else remaining equal, and hence would be justified only if they generate higher returns for the advertiser. As already mentioned, position is one of the key levers available to the advertiser, and the only one with any significant cost implication. Measurement of causal position effects are thus of critical importance to the advertiser.

A variety of mechanisms can lead to positions affecting outcomes such as clicks and sales. One mechanism could be that of signaling (Nelson, 1974; Kihlstrom and Riordan, 1984). In this mechanism, which is most relevant for experience goods, advertisers with higher quality goods spend greater amounts on advertising in equilibrium, and consumers take advertising expenses as a signal of product quality. Since it is well known that advertisers have to spend more money to obtain higher positions in the search advertising results, consumers might infer higher positions as a signal of higher quality.

A second mechanism might relate to consumers' learned experience about the relationship between position and the relevance of the advertisement. Since the auction mechanism of search engines such as Google inherently score ads with higher relevance higher (Varian, 2007). Over a period of time, consumers might have learned that ads that have higher positions are more likely to be relevant to them. Since consumer incurs a cost (in terms of time and effort) each time they click on a link, they might be motivated to click on the higher links first given their higher expected return from clicking higher links. Such a mechanism is consistent with a sequential search process followed by the consumer (Weitzman, 1979), where they start with the ad in the highest position and move down the list until they find the information they need. Chen and He (2006) show, using an analytical model, that it is viable equilibrium for advertisers with higher relevance to be positioned higher and consumers to be more likely to click on higher positions.

A third mechanism that could drive position effects is that of attention. Several studies have pointed to the fact that consumers pay attention only to certain parts of the screen. Using eye-tracking experiments, these studies show that consumers pay the greatest attention to the top four positions. Such an effect is particularly pronounced on Google and is often called the Google Golden triangle (Hotchkiss, Alston, and Edwards, 2005; Guan and Cutrell, 2007). The reasons for such an effect may be due to spillovers from attention effects for organic (unpaid) search results. The organic search results are sorted on relevance to consumers, and hence consumers may focus their attention first on the top positions in the organic search results. Since search advertising results are above or by the side of organic search results, consumers' attention might be focused on those ads that are closest to the organic results they are focused on. Thus, in addition to the economic mechanisms such as signaling and relevance, there might be behavioral mechanisms why higher positions get greater clicks and position effects show local inflections.

### 2.3 Matching options on Google

Google, which brands its search advertising product as *Adwords* provides targeting options to advertisers. When bidding on keywords, advertisers can specify the match-type of the ad. Some matching options available currently to advertisers on Google are broad match and exact match, with broad match being the default option[5]. An ad that is classified as a broad match is shown as long as one of the words in the ad phrase is in the search phrase entered by the consumer. An example of a broad match keyword phrase and the kinds of ads that might be shown is in table 1. As this example shows, in broad match the ad is eligible to be shown when any of the keywords for the ad appears in the search query. They can be in any order, singular or plural forms, synonyms and other variations. By contrast, if the advertiser specifies an exact match, the ad is served to the consumer only if the keywords are contained in the consumer's search phrase exactly. It does not allow for variations including order, singular vs. plural or synonyms. Table 2 illustrates an exact match situation, pointing to ads that will be served and that would not be served. Note that all the ads that would not be served in the exact match example in table 2 would have been served if the ad were a broad match type as in table 1.

Google's *Adwords* website[6] highlights several benefits of broad match. The claim is that it generates increased traffic and conversions, with a third of all clicks and conversions on Google being for broad match keywords. A reference is made to the fact that consumer search behavior is unpredictable, and hence it may be difficult to anticipate the exact keywords consumers may be searching for at a particular point in time. Broad match keywords, which by nature accommodate variation in the keywords consumers are searching for, can thus allow ads to be served in many situations where the advertiser may have failed to anticipate the exact keyword match consumers

[5]The discusson in this section is based primarily on the following two sources: http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=6100, (last accessed on 27 June 2011), and http://adwords.blogspot.com/2008/11/reach-more-customers-with-broad-match.html, )(last accessed on 27 June 2011)

[6]http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=6137, accessed 2011/6/27

are searching for. Third, Google claims to have an automatic mechanism by which global traffic trends for search phrases are analyzed and the ad is served only for the higher performing phrases, with the lower performers automatically discarded. Another benefit is that for broad match, the organic listing for the advertiser may be lower on average than in the exact match case, thus potentially increasing the incremental impact of the search ad.

On the flip side, broad match advertisements are potentially more expensive, due to a larger number of potential advertisers for whom a particular search phrase may be relevant and may be matched with the keywords they have bid on. For instance, a broad match on the search phrase "tennis shoes" would be relevant to any advertiser with products related to tennis (e.g. tennis rackets) and shoes (e.g. biking shoes). This intensified competition may make it more expensive for an advertiser to obtain a given position in the search advertising results.

Thus, it is of interest for advertisers to know the differences between broad and exact match in the causal effect of search advertising. Given our focus on position effects, we focus on finding if position effects differ between broad and exact match.

### 2.4 Weekend effects

Retail environments often see a significant difference in purchase behavior between weekdays and weekends. Such effects, and particularly their relationship with retail pricing has received some attention in the literature (Warner and Barsky, 1995). The argument for lower prices in the weekends, which are periods of higher demand is explained on the basis of lower search and transportation costs relative to weekdays, leading to more intensive search and hence lower prices offered by competing retailers in equilibrium. Scholten, Livingston, and Chen (2009) argue that since online retail environments significantly reduce search costs across the board both on weekdays and weekends, the price differential between weekdays and weekends should be reduced, and find empirical evidence for this.

While this literature talks about weekend effects in pricing, the differences in search costs between weekdays and weekends has some bearing on advertising effects as well. In particular, if there is any difference in search costs between weekdays and weekends, it should affect position effects of advertising. Recall that one rationale for positions effects in the first place is that consumers might sequentially search through the search advertising listings, starting at the high positions, which have higher expected returns for them, and stopping when the expected benefit from further search is lower than the expected cost. If search costs are lower in the weekends due to greater time available to the consumer, it would imply that consumers continue to search for longer periods, running further down the advertising listings on weekends than on weekdays. By this rationale, position effects should be weaker over the weekends than on weekdays.

By the same rationale of lower search costs over the weekends, consumers might have more time to search through organic listings during the weekends than on weekdays. Furthermore, in the case of product categories that are also sold offline in brick and mortar stores, they may have greater ability to search offline for the goods they are looking for on the weekends. Added to this is the fact that consumers who wish to shop offline over the weekends may pre-shop online before the weekend. The implication of these effects is that consumers might depend less on search advertising results on weekends than on weekdays. This may result in lower click through rates for search advertisements. At the same time, retailers' prices (both online and offline), and indeed advertising strategies could be systematically different on the weekends . Thus, raw comparisons of outcomes on weekends and weekdays could lead to erroneous conclusions on the effect of search advertising.

To sum up, it would be interesting to investigate if position effects systematically differ on the weekends relative to weekdays. The implications of the search cost explanation is that position effects should be weaker on the weekends and this can be tested in the data.

# 3 Selection issues

## 3.1 Selection on observables

As we have discussed in the previous section, measuring causal position effects, and investigating if they are systematically different for different match-types and different days of the week is of critical importance to the retailer. However, there are likely significant selection biases in the correlational effects, and we discuss them in this section.

First, we discuss the selection biases that may result if we compare outcomes for different positions by pooling observations across keywords, advertisers, match-types, days etc. Consider the case where we observe positions and outcomes for a set of keywords. It is likely that there are significant differences in click through rates or sales across different keywords. For instance, an advertiser who primarily sells tennis shoes but only a few biking shoes would likely get greater clicks for ads related to tennis shoes than biking shoes. At the same time, the ads for tennis shoes for this advertiser are likely to be in higher positions than for biking shoes, both because the expected click through rates (and hence Quality Scores) are higher for these ads, and potentially because the advertiser has greater advertising budgets for ads for tennis shoes, leading to higher bids. These two effects both raise the advertiser's *AdRanks* for keywords related to tennis shoes. A cross-sectional analysis across keywords would pick up these systematic differences between keywords as a spurious position effect.

Similarly, there could be selection biases when pooling observations across broad and exact match types (fewer clicks and lower positions for broad match relative to exact match), the advertiser (a bigger retailer might have higher clicks and position, leading to spurious position effects even when the true causal effect is zero) and day of week. To sum up, any analysis that pools across keywords, advertisers, match-types and days of the week are likely to give spurious effects of position. A solution to these selection issues on observables is to conduct a within keyword, within advertiser, within match-type and within day of week analysis of the position effects, which is feasible if we have panel data. If we repeatedly observe ads for the same advertiser, keyword, match-type and day of week, we can include fixed effects (or equivalently use the differences between the outcomes and their average values for a given keyword, match-type, advertiser and day of week combination) to control for selection on observables.

## 3.2 Selection on unobservables

In the previous sub-section, we discussed how we might obtain spurious effects if we pooled observations across ads where both position and outcomes may be systematically different due to underlying causes other than the position effects. However, while elimination of such selection biases using panel data may be useful, it does not guarantee elimination of all selection biases. This is because there is potential for selection on unobservables as well.

Consider a situation where we repeatedly observe outcomes and positions for ads for a particular advertiser, a particular keyword, a particular match-type and a particular day of week, say for a period of a year. Thus, we would have 52 observations of outcomes and positions corresponding to the 52 weeks in the year. Suppose the product sees seasonal variations in sales, with certain weeks having high sales due to exogenous reasons (e.g. holiday weeks). A strategic advertiser may focus their advertising efforts during those weeks of peak sales, and may therefore have higher positions during those weeks than the rest of the year. A correlational analysis would pick this correlation between position and outcome as a position effect, when in reality these are driven by exogenous seasonal variation. One could come up with innumerable other examples of unobservables that might cause correlational estimates to be confounded due to selection in position.

Selection may also be induced typical processes used by advertisers to set their bids. One mechanism that is often used by advertisers sets a fixed advertising to sales ratio for deciding advertising budgets. In the search engine context, this mechanism involves a continuous feedback loop from performance measures to bidding behavior. As sales per click increases, advertisers might automatically increase advertising budgets, which in turn increases their bid amounts and hence ensures higher positions for their ads. Similarly, as sales drop, advertising budgets and eventually position also fall. Such a mechanism would induce a positive bias in position effects, as higher position might be induced by increasing sales rather than the reverse.

A negative bias is also feasible due to potential rules used by advertisers in setting their bids. Consider an advertiser who has periodical sales, with higher propensity of consumers to visit their sites even without search advertising during that period (through other forms of advertising or marketing communication, such as catalogs for instance). The advertiser may in this instance reduce their search advertising budgets if they believe that they would have got the clicks that they obtain through search advertising anyway, and without incurring the expense that search advertising entails. Thus, they may generate high clicks and sales, even though their strategy is to spend less (and hence obtain lower positions) on search advertising during this period. This mechanism would induce a negative bias on estimates of position effects.

Another potential cause for selection biases is competition. Since search advertising positions are determined through a competitive bidding process, the bidding behavior of competitors could also induce biases in correlational estimates of position effects. Consider a competing bidder, who offers similar products and services as the focal advertiser, with data on the competing bidder unavailable to the latter. Due to mechanisms similar to those described above, competing bidders may place high or low bids when their sales are high. Since the competing bidder offers similar products as the focal advertiser, higher sales for the competing bidder, for instance due to a price promotion, may lower the sales for the focal advertiser. At the same time, the competing bidder may place a low bid on the keyword auction, thus pushing the focal advertiser higher in position. This negative correlation between position and sales for the focal advertiser is due to the unobserved strategic bidding behavior by the competitor.

To sum up, there are significant selection issues that may render correlational estimates of positions highly unreliable, with unpredictable signs and magnitude of the biases induced by selection on unobservables. In the subsequent sub-sections, we discuss the extant approaches to deal with this issue and present our solution.

### 3.3 Extant approaches to deal with selection

As mentioned in the introduction, position effects have been studied in the literature. An early study of the effect of position was Agarwal, Hosanagar, and Smith (2007), which concluded that click through rates decrease monotonically as one moves down the search advertising listings, but conversion rates go up and then down. This study controls for heterogeneity across keywords, but not across match-types, days etc. Further, it does not control for selection on unobservables. Instead, it reports a robustness check using an experimental design with randomly varying bids for a small number of observations. While such randomization can deal with potential selection biases induced by the bidding behavior of the focal firm, it cannot eliminate selection biases induced by strategic bidding behavior of competitors. This discussion also demonstrates why experimentation is difficult in this context, since it would require randomization of bids of all bidders - the focal bidder and its competitors.

A second approach has been to control for selection by modeling the process by which positions are determined using a parametric specification, and jointly estimating both the outcome and position equations (Ghose and Yang, 2009; Kalyanam, Borle, and Boatwright, 2010; Yang and Ghose, 2010). The selection in positions is explicitly modeled by estimating correlations between the errors of the two equations. This approach crucially depends on the validity of the parametric specification of the position equation. It might be hard to come up with a parametric specification for the position equation given that position is determined through an auction. Typically, a parametric specification is assumed for the position as a function of lagged variables for the focal advertiser, with no information on competitors. Such a specification is unlikely to be able to control for the selection issues induced by competitors' strategic bidding behavior. Furthermore, we have seen that there are a set of complex processes at work even within the focal bidder, with potentially different mechanisms operating at different times inducing biases of opposite signs. Such complexities would be hard to capture using a parametric specification. Added to this is the fact that such an approach is demanding computationally and requires that the researcher has access to appropriate exclusion restrictions that are necessary for parameter identification.

A third approach, adopted by Rutz and Trusov (2011) is to instrument for the position. Since an instrumental variable is hard to come by, this study uses the latent instrumental variables approach of Ebbes, Wedel, and Bockenholt (2005) to account for the potential endogeneity of position. The method relies on two crucial assumptions - normality of the outcome equation and departures from normality for the position equation. While the latter is not problematic, the former may be. As we will see in our empirical application, outcomes such as click through rates and sales are highly non-normal. For instance, periodical sales and promotional events, if unobserved in the data, would induce the distribution of the outcome variables to be skewed and even potentially multi-modal. Thus, it is unlikely that this approach, which is also inherently a parametric one, is viable in most contexts.

To sum up, the extant literature has either ignored the endogeneity/selection issues altogether, or taken parametric approaches to control for endogenous positions, both of which are problematic.

Further, this is a situation where experimentation, which is the typically advocated approach for obtaining causal effects, is usually infeasible. , . The realities of search engine advertising create significant challenges to uncovering causal effects using the types of approaches advocated in the literature. There are also significant challenges to their practical implementation.

# 4 Applying regression discontinuity to finding position effects

## 4.1 Regression discontinuity

Regression discontinuity (RD) designs can be employed to measure treatment effects when treatment is based on whether an underlying continuous score variable crosses a threshold. Under the condition that there is no other source of discontinuity, the treatment effect induces a discontinuity in the outcome of interest at the threshold. Thus, the limiting values of the outcome on the two sides of the threshold are unequal and the difference between these two directional limits measures the treatment effect.

Formally, let $y$ denote the outcome of interest, $x$ the treatment and $z$ the score variable, with $\bar{z}$ being the threshold above which there is treatment. Further define the two limiting values of the outcome variable as follows

$$y^{+} = \underset{\lambda \to 0}{Lim} \mathbb{E}\left[ y | z = \bar{z} + \lambda \right] \tag{2}$$

$$y^{-} = \underset{\lambda \to 0}{Lim} \mathbb{E}\left[ y | z = \bar{z} - \lambda \right] \tag{3}$$

Then the local average treatment effect is simply given by

$$d = y^{+} - y^{-} \tag{4}$$

Practical implementation of RD involves finding these limiting values non-parametrically using a local regression, often simply a local linear regression within a pre-specified bandwidth $\lambda$ of the threshold $\bar{z}$ and then assessing sensitivity to the bandwidth.

## 4.2 RD in the search advertising context

As described earlier in section 2.1, the position in search advertising listings is determined by an auction, with bidders ranked on a variable called *AdRank*, which in turn is the product of the bid and the *Quality Score* assigned by Google to the bidder for each specific keyword phrase for a particular match-type. The application of RD to this context relies on knowledge of the *AdRank* of competing bidders for a given position. Specifically, if bidder $A$ gets position $i$ in the auction and bidder $B$ gets position $i+1$, it must be the case that

$$AdRank_i > AdRank_{i+1} \tag{5}$$

or in other words

$$\Delta AdRank_i \equiv (AdRank_i - AdRank_{i+1}) > 0 \tag{6}$$

The score for the RD design is this difference in *AdRanks* and the threshold for the treatment (i.e. the higher of the two positions) is 0. The RD design measures the treatment effect by comparing outcomes for situations when $\Delta AdRank_i$ is just above zero and when it is just below zero. Thus, it compares situations when the advertiser just barely won the bid to situations when the advertiser just barely lost the bid.

For an RD design to be valid, it should be the case that the only source of discontinuity is the treatment. One consequence of this condition is that RD is invalidated if there is selection at the threshold. If it is the case that an advertiser can select his bid so as to have an *AdRank* just above the threshold, the RD design would be invalid. However, what comes to our assistance in establishing the validity of RD is that *AdRanks* are unobserved *ex ante* by the advertiser. Their own *AdRank*s are observed *ex post*, since Google reports the *Quality Score* on a daily basis at the end of the day, and the advertiser observes his own bid. However, *AdRanks* of competitors are not observed even *ex post*. Thus, the advertiser cannot strategically self-select to be on one side of the cutoff. Thus, occasions when the advertiser just barely won the bid and when he barely lost the bid can be considered equivalent in terms of underlying propensities for click throughs, sales etc.

Any difference between the limiting values of the outcomes on the two sides of the threshold can be entirely attributed to the position.

Advertisers observe their own *AdRanks* (albeit *ex post*) but not that of other advertisers. Typically, only the search engine observes the *AdRanks* for all advertisers. Therefore, the RD design could be applied by the search engine, but not by advertisers, or by researchers who have access to data only from one firm. Unfortunately, search engines like Google are typically unwilling to share data with researchers, partly due to the term of agreement with their advertisers. However, we have access to a dataset where we observe *AdRanks* for four firms in the same category. One of these firms acquired the three other firms in this set, and hence we have access to data from all firms, including from a period where they operated and advertised independently. We describe the data in more detail in section 5.

As discussed in section 3.1, an analysis of position effects across observables which may induce selection of position would lead to the estimation of spurious effects. A regression discontinuity design such as the one described above can account for such selection, but would lead to noise in the estimates. Controlling for observable variables such as advertiser, keyword, day of week, match-type etc. that can induce selection would lead to more precise estimates of the position effects. In a regression framework, this would involve including fixed effects for advertiser, keyword, match-type and day of week. The most general specification would include a fixed effect for every combination of these variables. An equivalent estimator is a differenced specification where the mean differenced outcome (i.e. with the mean of outcome for each unique combination of these observable variables subtracted from the outcomes corresponding to that combination of variables). The position effect, which compares these differenced outcomes across positions is thus a within estimator. This idea can be extended easily to the RD design, by comparing the limiting values of the mean differenced outcome variable on the two sides of the threshold. This is the estimator we use in our empirical application.

## 5 Data Description

To implement an RD design to measuring position effects in search advertising, we require observations for competing advertisers, whose advertisements appear in adjacent positions for at least a subset of the observations. We could then use these observations where the ads are at adjacent locations to find the RD estimate of the causal position effects. This can be done pairwise for every set of adjacent locations for which we have observations to obtain the local average position effect for each position. We need to observe the position of the ad, its *AdRank* (or bid and *Quality Score*), and performance measures such as click through rates and sales for the competing bidders. The dataset we use in this empirical application has exactly these features.

Our data consist of information about search advertising for a large online retailer of a particular category of consumer durables[7]. This firm, which is over 50 years old started as a single location retailer, expanding over the years to a nationwide chain of over 160 stores both through organic growth and through acquisition of other retailers. Since the category involves a very large number of products, running into the thousands, a brick and mortar retail strategy was dominated in terms of its economics by a direct marketing strategy. Thus, over the years, its strategy evolved to stocking a relatively small selection of entry-level, low-margin products with relatively high sales rates in the physical stores, with the very large number of slower moving, high margin products being sold largely through the direct marketing channel. Recently, the firm acquired three other large online retailers. Two of the four firms are somewhat more broadly focused, while the two others are more narrowly focused on specific sub-categories. However, each of them has significant overlaps with the other in terms of products sold. For a significant period of time after the acquisition, the firms continued to operate independently, with independent online advertising strategies. Our data have observations on search advertising on Google for these four firms, and crucially for the period where they operated as independent advertisers.

We have a total number of about 23.7 million daily observations over a period of 9 months in the database, of which about 10.5 million observations involve cases where two or more advertisers

[7] We are unable to disclose the name of the firm or details of the category due to confidentiality concerns on the part of the firm.

among the set of 4 firms bid on the same keyword. Since the keywords are often not in adjacent positions, we filter out observations where the observations are not adjacent. We also drop observations where we don't have bids and *Quality Scores* for both of the adjacent advertisements. Since the position is a daily average (with Google varying the position by geography and due to some degree of experimentation), we also drop observations where the average positions are more than 0.1 positions away from the whole number. We are thus left with a total of 330,336 observations where we observe advertisements in adjacent positions, spanning 22471 unique keyword phrase/match-type combinations. An overwhelming majority (79%) of the 22471 keywords are of the broad match-type, and the rest are of the exact match-type. There are a total of 18875 unique keywords in this analysis dataset, with most exact match-type keywords also advertised as broad match type, but clearly not necessarily vice versa.

Table 4 has the list of variables in the dataset (including variables we have constructed such as click through rates, conversion rates and sales per click) and the summary statistics for these variables. We report these statistics for broad match and exact match keywords, in addition to the overall summaries. Observations are only recorded on days that have at least one impression, i.e. when at least one consumer searched for the keyword phrase. Through a tracking of cookies on consumer's computers, each impression is linked to a potential click, order, sales value, margin etc. As per standard industry practice, a sales order is attributed to the last impression, click etc., with previous impressions not getting credit for these sales.

On average, there are about 46 impression per keyword phrase per day, but the dispersion in the number of impression is large, with a standard deviation of almost 226. On average, broad match keywords receive greater impressions than exact match keywords. The number of clicks are however higher for exact keywords than for broad match keywords. Virtually all performance metrics, such as clicks, click through rates, orders, conversions etc. are higher for the exact match keyword the firms advertise on than for broad match keywords. Note that the broad and exact keywords are not necessarily comparable, since the firms might be bidding on different kinds of keywords in the broad and exact cases.

To summarize, we have obtained a unique dataset, consisting of information at a daily level on keywords, type of match, bids, quality score and key performance metrics for the advertisement. To the best of our knowledge, this is the very first time that a dataset has been assembled which includes competitive information and allows us to get to causal effects which might otherwise not be feasible to obtain.

## 6 Results

We next present the results of our empirical analysis. We conducted an analysis of the effect of position on two key metrics of interest to advertisers - click through rates (henceforth CTR) and the number of sales orders (henceforth orders). The reasons to select these two metrics is that they are the most important metrics from the point of view of the advertiser. The CTR measures the proportion of consumers who were served the ad who clicked on it and arrived at the advertiser's website. Since the advertiser's control on the consumer's experience only begins once the consumer arrives at the website, CTR is of critical importance to the advertiser in measuring the effectiveness of the advertisement in terms of driving 'volume' of traffic. We could conduct an analysis on raw clicks instead, but it does not make any material difference to the results, and CTR is the more commonly reported metric.

The second measure we consider is the number of sales orders corresponding to that keyword. This is again a key metric for the firm since it generates revenues only when a consumer places an order. We attempted an analysis on measures like conversion rates, sales value and sales per click, but do not report these estimates since almost all the estimates were statistically insignificant. This is partly driven by the fact that the category in focus sees very infrequent purchases, reducing the statistical significance of results.

As discussed earlier, we include fixed effects for the advertiser, keyword, match-type and day of week, to control for selection on these observables. This is achieved by differencing the mean outcomes (click through rates and orders) for advertiser, keyword, match-type and day combination from each outcome for that combination. We compare pairs of positions at a time, using local linear regression to find the RD estimates and their standard errors. An important consideration for RD estimators is bandwidth selection, which refers to the size of the neighborhood around the threshold for which to conduct the analysis. We report RD estimates for a bandwidth of 5% of a standard deviation of the score (*AdRank*) for each pair of positions. We check for robustness of estimates to bandwidth choice by finding the estimates when the bandwidth is 10% of a standard deviation and 2.5% of a standard deviation respectively.

### 6.1 Effect of position on click through rates

First, we consider the effect of position in search engine advertising on click through rates. The pooled results of all advertisements in the analysis sample, with fixed effects for advertiser, keyword, match-type and day of week are reported in table 5. We report both correlational estimates (comparisons of means across each pair of positions) and the RD estimates, with a bandwidth set at 5% of a standard deviation of the score. We report the baseline click through rates for each position, which is the click through rate for the lower position in the pair. We report these baseline numbers separately for the correlational and RD estimates, with the RD baseline representing the observations within the bandwidth. One point to note is that these comparisons should only be conducted on a pairwise basis. For instance, the observations in position 2 that are used for analyzing the shift from position 2 to 1 are not the same as the observations used to compare 3 to 2. Hence, it will not be the case that the baseline for position 2 is the sum of the baseline for position 3 and the effect of moving from position 3 to 2.

The correlational estimates would suggest that there is a significant effect only at position 1. The remaining effects are statistically insignificant. However, when we look at the RD estimates, we see significant effects across multiple positions. The effects at position 1, 3, 5 and 6 are significant. As seen in figure 1, the topmost position is often above the organic search results and hence distinctive relative to the other ads. Thus, the effect at position 1 is to be expected. There is no significant position effect between positions 3 and 2. However, there is a significant and positive effect when moving from position 4 to position 3. Such an effect is consistent with the Google golden triangle effect, which has been postulated to be due to attention effects and documented in eye tracking studies (Hotchkiss, Alston, and Edwards, 2005; Guan and Cutrell, 2007) as well as in using advertising and sales data in Kalyanam, Borle, and Boatwright (2010). Further, there seem to be significant effects when moving from positions 6 to 5 and 7 to 6. These positions are typically below the page fold and often require consumers to scroll down (whether position 6 or 5 appears below the fold depends on the size of the browser window, the number of ads that appear above the organic results, etc.). This result is also intuitive.

The differences between the correlational and RD estimates are important, since they indicate the nature of the selection in positions. The fact that correlational estimates are insignificant where RD estimates are suggest that the selection bias is negative in the case of positions 3, 5 and 6, washing out the causal effects of these positions. This can result from advertisers or their competitors' strategic behavior, as indicated earlier. Further, the effect of selection differs significantly by position, with the magnitude of the difference between the correlational and RD estimates ranging between 0.0003 and 0.0010. This suggests a heavy burden for parametric controls for selection such as the ones used in the extant literature (Ghose and Yang, 2009; Yang and Ghose, 2010; Rutz and Trusov, 2011) .

The causal position effects are not just statistically significant, but have huge economic significance as well. For instance, the causal effect at position 1 as a proportion of the baseline click through rate is 18.8% of the baseline click through rates. They are 11.1%, 14.9% and 21.2% respectively as positions 3, 5 and 6, again large effects. Thus, it seems like in this category at least, if the objective of search advertising is to drive up clicks, it is effective at these positions and by a large magnitude.

## 6.2 Effect of position on sales orders

We next investigate if the position in search advertising results causally affects the number of sales orders that are generated, and report the RD estimates in table 6. A note of caution here is that data is sparse for orders, given the nature of the category and statistically insignificant estimates may reflect this sparsity.

We find that the correlational effects are once again misleading. They suggest that there are positive effects on sales only at the top position. By contrast, the RD estimates suggest that the only significant effect is at position 5, with no significant effects above that. This suggests that the nature of the mechanisms that may cause position to affect sales, such as quality signaling really play out only once goes below the top 5 positions. In terms of economic significance, these effects are even stronger than for click through rates, with sales orders jumping up by over 200% relative to the baseline.

### 6.3 Broad vs. exact match types

We have earlier discussed why we might expect differences in effects between broad and exact match types. We report the RD estimates for broad and exact match types for click through rates in table 7. The comparisons of these two types of match types reveals an interesting asymmetry in effects. For broad match types, there are significant effects at position 3, 5 and 6 only but not at position 1. For exact match types, on the other hand, the only significant effect is at position 1. This effect is consistent with the idea that for broad match types, there is heterogeneity in relevance of the ad for a particular consumer, and lower head to head competition between ads in the top positions. However, there may be attention effects such as the Google golden triangle effect still at play. On the other hand, for exact match types, there is likely to be more direct competition between advertisements. Here the top position plays a strong effect. Another explanation for the strong effect of position 1 in exact match type is that consumers might use the search engine merely to reach a website that they had anyway decided to go to, and which due to its high relevance may appear at the top of the search listings as well.

Table 8 reports the broad and exact match type effects for sales orders. The broad match type results reflect the pooled results, while the exact match type has no significant effects (there are insufficient data for analysis beyond the 4th position in the case of exact match).

### 6.4 Weekend effects

We investigate the weekend effects in this context. The results for the position effects separated by weekday and weekend are reported in Tables 9 and 10 respectively for click through rates and number of orders. The weekday results for CTR are largely similar to the pooled results, with a significant effect at position 1, 3 and 5. The weekend effects are less significant in general, partly reflecting the smaller number of observations, but also show differences in the position effects. The only marginally significant results (i.e. at 90% significance level) are at positions 4 and 6, which typically are below the usual zone that consumers pay the most attention to. The absence of significant position effects may reflect the differences in search costs of consumers between weekdays and weekends. If consumers search costs are lower on weekends, they are more likely to search lower down the advertising results before stopping, giving rise to the effects we estimate. Thus, these results are consistent with the story in Warner and Barsky (1995). In terms of sales orders, there are no major directional differences between weekdays and weekends, with significant effects largely at lower positions like the 4th and 5th positions.

### 6.5 Robustness to Bandwidth choice

Finally, we investigate the choice of bandwidth, which is an important aspect of the RD design. We chose an arbitrary bandwidth of 5% of a standard deviation of the score. Bandwidth choice entails a tradeoff between bias and efficiency. A large bandwidth will lead to more biased estimates but with better efficiency (lower standard errors), while a smaller bandwidth will have the opposite effect. We check for the robustness of our empirical results to bandwidth choice by repeating the analysis for the pooled results with a lower bandwidth of 2.5% of a standard deviation. The comparisons of our results (at 5% bandwidth) with those at the lower bandwidth are reported in tables 11 and 12. These comparisons illustrate the bias vs. efficiency tradeoffs described above. However, the main point is that the results are largely similar with the lower bandwidth, although the level of significance of many estimates reduces. This gives us confidence about our estimates. While not reported here, we tested other bandwidths, including larger ones than 5% of a standard deviation and find that our results are robust to bandwidth choice over a relatively wide range.

## 7 Conclusion

In this paper, we investigate the important question of the causal effect of position in search advertising on outcomes such as website visits and sales. It is challenging to obtain causal estimates using the approaches suggested in the literature. While acknowledging the potential for significant selection biases, the literature has typically taken a parametric approach to control for selection. We present a novel regression discontinuity based approach to uncovering causal effects in this context. The importance of this approach is particularly high in this context due to the difficulty of experimentation and the infeasibility of other approaches such as instrumental variable methods.

We obtain a unique dataset of advertising by a durable goods retailer on the Google search engine. The application of regression discontinuity requires that the researcher observe the *AdRanks* of competing retailers, which is the score used by Google to decide position. Typically, only Google observes the *AdRanks* of competing advertisers and hence we would not be able to apply RD to measuring causal position effects using data from only one advertiser who observes only his own *AdRank*. However, our data also consist of *AdRanks* of firms ,that competed with this firm in Google's search advertising listings, and this allows us to set up an RD design.

We find that there are significant position effects, and that these would be understated by correlational analyses. The selection bias in this context happen to be negative and hence wipe out the causal position effects. Further, we find that the position effects are of great economic significance, increasing the click through rates by about a fifth in positions where they are significant. We find important differences in these effects between broad and exact match keywords, and that exact match delivers significantly higher click through rates. Also position effects are much stronger, with exact match having significant effects only at the very top position, while broad match has significant effects only lower down. Finally, we find important weekend effects in this context. Position effects are weaker on the weekend and this result is consistent with the idea that consumers' search costs are lower during the weekends.

The results of our empirical analysis would be of great interest to managers who are setting firms' online advertising strategies. Further, they should be of interest to the academic audience since we point to significant selection issues in this context and point to a viable way to correct for them. The methodological innovation should be of interest to search engines as well, who might be interested in viable alternatives to experimentation, which tends to be difficult and expensive in this context, in addition to being subject to contractual limitations. We hope this study leads researchers to pay greater attention to concerns about selection and endogeneity, which have received less than the necessary attention in the field of online advertising specifically and web analytics in general.

## References


Agarwal, A., K. Hosanagar, and M. D. Smith (2007): "Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets," *(forthcoming) Journal of Marketing Research.*

Busse, M., J. Silva-Risso, and F. Zettelmeyer (2006): "$1000 Cash Back: The Pass-through of Auto Manufacturer Promotions," *American Economic Review*, 96(4), 1253–1270.

Busse, M., D. Simester, and F. Zettelmeyer (2010): "The Best Price You'll Ever Get: The 2005 Employee Discount Pricing Promotions in the U.S. Automobile Industry," *Marketing Science*, 29(2), 268–290.

Chen, Y., and C. He (2006): "Paid Placement: Advertising and Search on the Internet," Working Paper, University of Colorado at Boulder.

Cook, T. D., and D. T. Campbell (1979): *Quasi-Experimentation.* Houghton Miffin.

Ebbes, P., M. Wedel, and U. Bockenholt (2005): "Solving and Testing for Regressor Error (in)Dependence When No Instrumental Variables are Available: With New Evidence for the Effect of Education on Income," *Quantitative Marketing and Economics*, 3(4), 365–392.

Ghose, A., and S. Yang (2009): "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets," *Management Science*, 55(10), 1605–1622.

Guan, Z., and E. Cutrell (2007): "An Eye Tracking Study of the Effect of Target Rank on Web Search," *Proceedings of the ACM CHI'07 Conference*, pp. 417–420.

Hahn, J., P. Todd, and W. van der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design," *Econometrica*, 69, 201–209.

Hotchkiss, G., S. Alston, and G. Edwards (2005): *Eye Tracking Study: An In Depth Look at Interactions with Google using Eye Tracking Methodology.* Enquiro Search Solutions Inc.

Imbens, G. W., and T. Lemieux (2008): "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2), 615–635.

Kalyanam, K., S. Borle, and P. Boatwright (2010): "The Impact of Ad Position in Search Engine Marketing: The Effect of Brand Name and Price," Working Paper, Santa Clara University.

Kihlstrom, R. E., and M. H. Riordan (1984): "Advertising as a Signal," *Journal of Political Economy*, 92(3), 427–450.

Lee, D. S., and T. Lemeiux (2010): "Regression Discontinuity Designs in Economics," NBER Working Paper 14723, National Bureau of Economic Research.

Nair, H. S., W. R. Hartmann, and S. Narayanan (2011): "Identifying Causal Marketing Mix Effects using a Regression Discontinuity Design," *Marketing Science (forthcoming).*

Nelson, P. (1974): "Advertising as Information," *Journal of Political Economy*, 82(4), 729–754.

Rutz, O. J., and M. Trusov (2011): "Zooming In on Paid Search Ads - A Consumer-Level Model Calibrated oin Aggregated Data," *Marketing Science (forthcoming).*

Scholten, P. A., J. A. Livingston, and J. Chen (2009): "Do countercyclical-weekend effects persist in online retail markets?," *Electronic Commerce Research and Applications*, 8, 174–181.

Shadish, W. R., T. D. Cook, and D. T. Campbell (2001): *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Wadsworth Publishing.

Thistlethwaite, D. L., and D. T. Campbell (1960): "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment," *Journal of Educational Psychology*, 51(6), 209–317.

van der Klaauw, W. (2008): "Regression Discontinuity Analysis: A Survey of Recent Developments in Economics," *Labour: Review of Labour Economics and Industrial Relations*, 22(2), 219–245.

Varian, H. R. (2007): "Position Auctions," *International Journal of Industrial Organization*, 25, 1163–1178.

Warner, E. J., and R. B. Barsky (1995): "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays," *Quarterly Journal of Economics*, 110(2), 321–352.

Weitzman, M. L. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47(3), 641–654.

Yang, S., and A. Ghose (2010): "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence," *Marketing Science*, 29(4), 602–623.

Table 1: Example of a broad match keyword phrase

| Keyword phrase entered by the consumer | Ads may be shown for |
|---|---|
| Tennis Shoes | Tennis |
| | Shoes |
| | Buy Tennis Shoes |
| | Tennis Shoes Photos |
| | Running Shoes |
| | Tennis Sneakers |

Table 2: Example of an exact match keyword phrase

| Keyword phrase entered by the consumer | Ads may be shown for | Ads will not be shown for |
|---|---|---|
| Tennis Shoes | Tennis Shoes | Running Shoes |
| | Buy Tennis Shoes | Tennis Shoe |
| | Tennis Shoes Photo | Shoes Tennis |

Table 4: Summary statistics of the data

| Variable | All keywords | | Broad match | | Exact match | |
|---|---|---|---|---|---|---|
| | Mean | Std. Dev | Mean | Std. Dev | Mean | Std. Dev |
| Impressions | 45.8977 | 225.9401 | 48.8384 | 239.5032 | 35.2865 | 166.6019 |
| Clicks | 0.5471 | 2.5304 | 0.4497 | 1.7811 | 0.8883 | 4.1941 |
| Click through rate (%) (Clicks/Impressions) | 1.9132 | 6.7079 | 1.3726 | 5.5256 | 3.8151 | 9.5531 |
| Number of orders | 0.0046 | 0.0724 | 0.0033 | 0.0593 | 0.0093 | 0.1062 |
| Conversion rate (% of 75593 non-zero clicks that resulted in orders) | 0.7468 | 7.3578 | 0.6291 | 6.8207 | 1.0343 | 8.5251 |
| Sales ($) | 0.4887 | 17.0041 | 0.3514 | 14.7753 | 0.9730 | 23.2148 |
| Average Sales per (non-zero) click ($) | 0.7435 | 21.0691 | 0.6360 | 20.1133 | 1.0081 | 23.2670 |
| Gross margin ($) | 0.1958 | 6.7226 | 0.1341 | 5.4559 | 0.4131 | 9.9757 |
| Bid ($ per click) | 0.3969 | 0.8219 | 0.3381 | 0.8676 | 0.6037 | 0.5928 |
| Quality Score | 5.9791 | 1.2496 | 6.0160 | 1.2464 | 5.8518 | 1.2515 |
| AdRank | 2.3523 | 5.0899 | 2.0164 | 5.3770 | 3.5340 | 3.6961 |

Table 5: Position effects on click through rates

| Position | Correlational Estimates (CTR %) | | | RD Estimates (CTR %) | | |
|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 2.1372 | 0.3633 | 0.0000 | 2.3404 | 0.4415 | 0.0106 |
| 3 to 2 | 1.3737 | 0.0163 | 0.4922 | 1.2802 | 0.0774 | 0.2059 |
| 4 to 3 | 1.1026 | -0.0124 | 0.5799 | 1.0304 | 0.1143 | 0.0142 |
| 5 to 4 | 0.8832 | 0.0186 | 0.4539 | 0.8620 | 0.0589 | 0.2078 |
| 6 to 5 | 0.7537 | 0.0085 | 0.7976 | 0.7635 | 0.1135 | 0.0236 |
| 7 to 6 | 0.5791 | 0.0626 | 0.2232 | 0.7161 | 0.1521 | 0.0378 |
| 8 to 7 | 0.4991 | -0.0339 | 0.6305 | 0.4913 | -0.0082 | 0.9459 |

Table 6: Position effects on number of sales orders

| Position | Correlational Estimates (Orders) | | | RD Estimates (Orders) | | |
|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 0.0044 | 0.0013 | 0.0048 | 0.0044 | 0.0005 | 0.7783 |
| 3 to 2 | 0.0030 | -0.0000 | 0.9992 | 0.0026 | 0.0005 | 0.4993 |
| 4 to 3 | 0.0031 | 0.0001 | 0.8138 | 0.0033 | -0.0005 | 0.4137 |
| 5 to 4 | 0.0019 | 0.0004 | 0.3785 | 0.0016 | 0.0009 | 0.2385 |
| 6 to 5 | 0.0011 | 0.0001 | 0.8570 | 0.0009 | 0.0019 | 0.0108 |

Table 7: Position effects on click through rates: broad vs. exact match

| Position | Pooled Estimates (CTR %) | | | Broad Match (CTR %) | | | Exact Match (CTR %) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 2.3404 | 0.4415 | 0.0106 | 1.7547 | 0.2238 | 0.1899 | 3.2451 | 0.7823 | 0.0400 |
| 3 to 2 | 1.2802 | 0.0774 | 0.2059 | 1.1486 | 0.0687 | 0.2796 | 2.2061 | 0.1171 | 0.5873 |
| 4 to 3 | 1.0304 | 0.1143 | 0.0142 | 0.9924 | 0.1075 | 0.0219 | 1.3951 | 0.1446 | 0.4527 |
| 5 to 4 | 0.8620 | 0.0589 | 0.2078 | 0.8486 | 0.0525 | 0.2736 | 1.1598 | -0.1532 | 0.4660 |
| 6 to 5 | 0.7635 | 0.1135 | 0.0236 | 0.7773 | 0.1061 | 0.0415 | 0.8027 | 0.2051 | 0.4380 |
| 7 to 6 | 0.7161 | 0.1521 | 0.0378 | 0.7675 | 0.1409 | 0.0606 | 0.3241 | -0.3060 | 0.6988 |
| 8 to 7 | 0.4913 | -0.0082 | 0.9459 | 0.5691 | 0.0150 | 0.9105 | 0.6944 | -0.1394 | 0.8101 |

Table 8: RD Estimates of Position effects on number of sales orders: broad vs. exact match

| Position | Pooled Estimates (Orders) | | | Broad Match (Orders) | | | Exact Match (Orders) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 0.0044 | 0.0005 | 0.7783 | 0.0006 | -0.0002 | 0.7435 | 0.0097 | -0.0018 | 0.6505 |
| 3 to 2 | 0.0026 | 0.0005 | 0.4993 | 0.0024 | 0.0000 | 0.9585 | 0.0037 | 0.0033 | 0.1707 |
| 4 to 3 | 0.0033 | -0.0005 | 0.4137 | 0.0031 | -0.0007 | 0.2691 | 0.0057 | -0.0005 | 0.8989 |
| 5 to 4 | 0.0016 | 0.0009 | 0.2385 | 0.0016 | 0.0012 | 0.1372 | | | |
| 6 to 5 | 0.0009 | 0.0019 | 0.0108 | 0.0005 | 0.0021 | 0.0109 | | | |

Table 9: RD Estimates of Position effects on click through rates: weekday vs. weekend

| Position | Pooled Estimates (CTR %) | | | Weekday (CTR %) | | | Weekend (CTR %) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 2.3404 | 0.4415 | 0.0106 | 2.4797 | 0.4395 | 0.0333 | 1.9884 | 0.4658 | 0.1405 |
| 3 to 2 | 1.2802 | 0.0774 | 0.2059 | 1.2447 | 0.1091 | 0.1193 | 1.4106 | -0.0066 | 0.9581 |
| 4 to 3 | 1.0304 | 0.1143 | 0.0142 | 1.0130 | 0.1283 | 0.0194 | 1.0859 | 0.0423 | 0.6138 |
| 5 to 4 | 0.8620 | 0.0589 | 0.2078 | 0.8211 | 0.0256 | 0.6448 | 0.9743 | 0.1637 | 0.0603 |
| 6 to 5 | 0.7635 | 0.1135 | 0.0236 | 0.7499 | 0.1448 | 0.0120 | 0.7968 | 0.0407 | 0.6899 |
| 7 to 6 | 0.7161 | 0.1521 | 0.0378 | 0.5821 | 0.1056 | 0.2009 | 1.0061 | 0.2730 | 0.0874 |
| 8 to 7 | 0.4913 | -0.0082 | 0.9459 | 0.5368 | -0.0238 | 0.8738 | 0.5444 | 0.1123 | 0.5896 |

Table 10: RD Estimates of Position effects on number of sales orders: weekday vs. weekend

| Position | Pooled Estimates (Orders) | | | Weekday (Orders) | | | Weekend (Orders) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 0.0044 | 0.0005 | 0.7783 | 0.0055 | -0.0004 | 0.8460 | 0.0015 | 0.0027 | 0.3106 |
| 3 to 2 | 0.0026 | 0.0005 | 0.4993 | 0.0017 | 0.0006 | 0.4377 | 0.0048 | 0.0002 | 0.8887 |
| 4 to 3 | 0.0033 | -0.0005 | 0.4137 | 0.0031 | -0.0004 | 0.5516 | 0.0038 | -0.0006 | 0.7112 |
| 5 to 4 | 0.0016 | 0.0009 | 0.2385 | 0.0022 | 0.0002 | 0.7729 | 0.0000 | 0.0033 | 0.0594 |
| 6 to 5 | 0.0009 | 0.0019 | 0.0108 | 0.0006 | 0.0008 | 0.1029 | 0.0018 | 0.0053 | 0.0454 |

Table 11: Robustness check - bandwidth selection (CTR)

| Position | Bandwidth = $0.05\sigma^2$ (CTR %) | | | Bandwidth = $0.025\sigma^2$(CTR %) | | |
|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 2.3404 | 0.4415 | 0.0106 | 2.3792 | 0.3108 | 0.1362 |
| 3 to 2 | 1.2802 | 0.0774 | 0.2059 | 1.3240 | 0.0974 | 0.2625 |
| 4 to 3 | 1.0304 | 0.1143 | 0.0142 | 0.9811 | 0.1371 | 0.0339 |
| 5 to 4 | 0.8620 | 0.0589 | 0.2078 | 0.9428 | 0.0404 | 0.5503 |
| 6 to 5 | 0.7635 | 0.1135 | 0.0236 | 0.7573 | 0.0622 | 0.3011 |
| 7 to 6 | 0.7161 | 0.1521 | 0.0378 | 0.5287 | 0.1837 | 0.4750 |
| 8 to 7 | 0.4913 | -0.0082 | 0.9459 | 0.4750 | 0.0794 | 0.6576 |

Table 12: Robustness check - bandwidth selection (orders)

| Position | Bandwidth = $0.05\sigma^2$ (Orders) | | | Bandwidth = $0.025\sigma^2$ (Orders) | | |
|---|---|---|---|---|---|---|
| | Baseline | Estimate | p-value | Baseline | Estimate | p-value |
| 2 to 1 | 0.0044 | 0.0005 | 0.7783 | 0.0057 | -0.0015 | 0.5145 |
| 3 to 2 | 0.0026 | 0.0005 | 0.4993 | 0.0024 | -0.0003 | 0.7732 |
| 4 to 3 | 0.0033 | -0.0005 | 0.4137 | 0.0032 | -0.0008 | 0.3449 |
| 5 to 4 | 0.0016 | 0.0009 | 0.2385 | 0.0026 | 0.0006 | 0.6521 |
| 6 to 5 | 0.0009 | 0.0019 | 0.0108 | 0.0000 | 0.0012 | 0.1317 |

Figure 1: Example of search advertising results