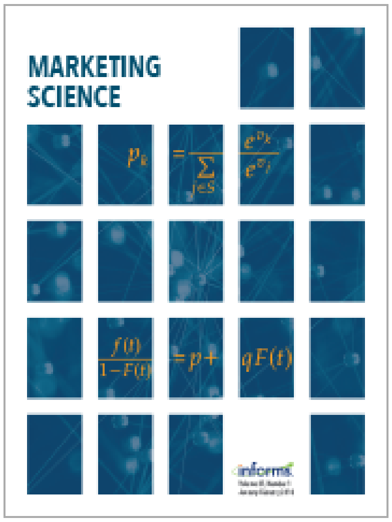

MARKETING
SCIENCE

MARKETING SCIENCE
Vol. 34, No. 3, May–June 2015, pp. 388–407
ISSN 0732-2399 (print) | ISSN 1526-548X (online)
http://dx.doi.org/10.1287/mksc.2014.0893

Position Effects in Search Advertising and their Moderators: A Regression Discontinuity Approach

Sridhar Narayanan
Stanford University, Stanford, California 94305, sridhar.narayanan@stanford.edu

Kirthi Kalyanam
Santa Clara University, Santa Clara, California 95053, kkalyanam@scu.edu

We investigate the causal effect of position in search engine advertising listings on outcomes such as click-through rates and sales orders. Because positions are determined through an auction, there are significant selection issues in measuring position effects. A simple mean comparison of outcomes at two positions is likely to be biased due to these selection issues. 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. Econometric approaches to address the selection are rendered infeasible due to the difficulty of finding suitable instruments in this context. We show that a regression discontinuity (RD) approach is feasible to measure causal effects in this important context. We apply the approach to a large and unique data set of daily observations containing information on a focal advertiser as well as its major competitors. Our RD estimates demonstrate that there are significant selection biases in the more naive estimates. While a mean comparison of outcomes across positions would indicate very large position effects, we find that our RD estimates of these effects are much smaller, and exist only in some of the positions. We further investigate moderators of these effects. Position effects are stronger when the advertiser is smaller, and when the consumer has low prior experience with the keyword for the advertiser. They are weaker when the keyword phrase has specific brand or product information, when the ad copy is more specific as in exact matching options, and on weekends compared to weekdays.

Keywords: search advertising; online advertising; position effects; advertising effects; causal effects; regression discontinuity
History: Received: December 17, 2011; accepted: September 30, 2014; Preyas Desai served as the editor-in-chief and Sanjog Misra served as associate editor for this article. Published online in *Articles in Advance* March 19, 2015.

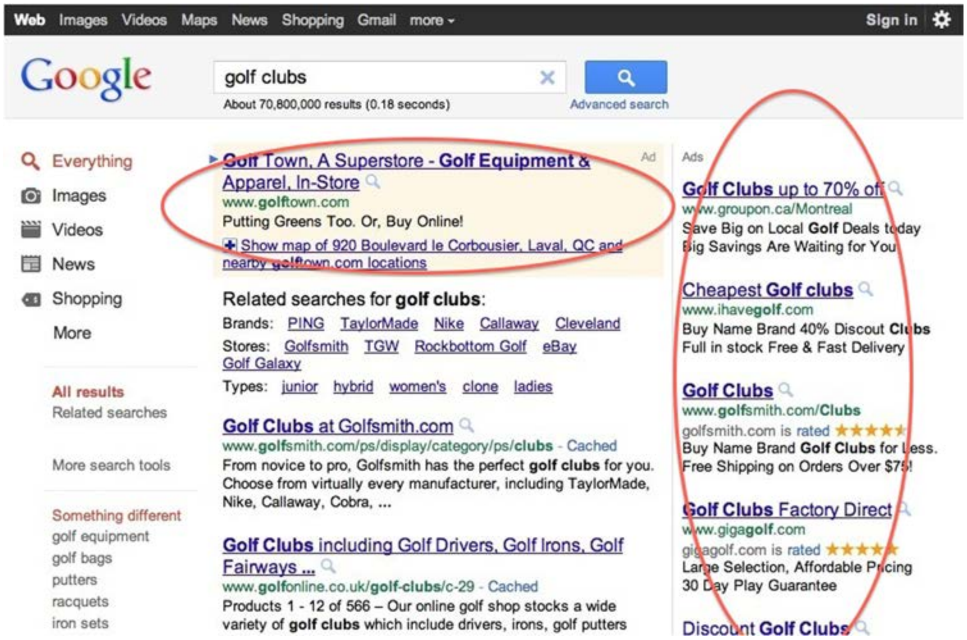
1. Introduction

Search advertising, which refers to paid listings on search engines such as Google, Bing, and Yahoo, has emerged since the mid-2000s to be an important and growing part of the advertising market. An example is shown in Figure 1 of the results of a search for the phrase "golf clubs" on Google, the most popular search engine. The order in which these paid listings are served is determined through a keyword auction. Advertisers place bids to get specific positions in these listings; higher positions cost more than lower positions. It is therefore crucial for advertisers to understand the effect of position in search advertising listings on outcomes such as click-through rates (CTRs) and sales.

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 on 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, and because it is the only variable with any significant cost implications. 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 et al. 2005, Guan and Cutrell 2007). One or more of these mechanisms could be at play simultaneously, leading to position effects of search advertising. Several empirical studies have documented the relationship between position and behavioral outcomes such as CTRs, conversion rates, and sales (Agarwal et al. 2011, Ghose and Yang 2009, Kalyanam et al. 2010, Yang and Ghose 2010, Rutz and Trusov 2011).

In measuring the effects of search advertising, one has to address some significant endogeneity issues. Because search ads are shown based on the keywords that a consumer types into the search engine, they are targeted at those who potentially have a higher propensity to view the website or Web pages that

Figure 1 (Color online) Example of Search Advertising Results



Note. This research was conducted on July 21, 2011 in Montréal.
Source. Google and the Google logo are registered trademarks of Google Inc., used with permission.

the ad leads to, or to purchase the product being advertised. The correlation between ad views and clicks/page views or purchases may thus overstate the causal effect of viewing the ad. Experimentation is a feasible way to uncover the causal effect. For instance, Blake et al. (2015) report the results of a series of field experiments run by eBay. These results showed that the causal effect of search ads is much smaller than correlational measures might indicate, that these are negligible for brand keywords, where organic search listings may act as a good substitute for paid ads, and that it is most effective for consumers who have not had prior experience with the product.

In the case of measuring position effects, there is a related but not identical endogeneity issue. Essentially, measuring causal effects of position is challenging due to the inability to generate experimental variation in position in search advertising listings. This is because position is determined through an online auction, with competing advertisers bidding for their advertisements to appear in the listings. This leads to position being endogenous. Past studies have tried to address this issue either by conducting experiments in which bids for the focal firm are randomized (Agarwal et al. 2011) or by accounting for the potential endogeneity of position through a simultaneous equations approach that incorporates a parametric selection equation (Ghose and Yang 2009, Rutz and Trusov 2011, Kalyanam et al. 2010). Experimentation is difficult in this context; 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 also of those of competing firms. Imagine, for instance, a competing advertiser bidding for higher positions on days when it expects higher sales due to a sales promotion. The sales promotion at this competing advertiser might lower sales at the focal advertiser. On such days, the focal advertiser's position would likely be lower due to the higher bids of the competitors. Even in the absence of a true position effect, such lower sales associated with lower position may be picked up spuriously as an effect of position on sales. Thus, while randomization of bids might eliminate the selection biases induced by the bidding behavior of the focal advertiser, the selection biases induced by competitors' strategic bidding behavior are not eliminated. One cannot make causal inferences based on such an experiment. Instruments are difficult if not impossible to find in this context; demand side factors that are correlated with position cannot typically be excluded from consumer outcome variables, and it is hard to find cost-side instruments that vary with position. Parametric selection equations are also likely to be problematic in this context. Because position is determined through a set of highly complex processes (Jerath et al. 2011), the nature of the selection effects could vary by position and the use of an incorrect specification would lead to unpredictable biases in estimates of position effects. Furthermore, this approach requires the availability of a valid exclusion restriction with sufficient variation, which is hard to find in practice. In addition, the simultaneous equations approach of specifying a parametric selection equation is computationally burdensome in practice. Thus, while there is an extant literature on position effects in search advertising, there remains a gap in the literature in finding causal position effects and its moderators. Advertisers are also interested in finding a robust and easily implemented approach to finding causal estimates of positions.

We present an RD approach to finding causal position effects in search advertising. The RD design, a quasi-experimental approach, was first developed in the program evaluation literature (Thistlethwaite and Campbell 1960, Cook and Campbell 1979, Shadish et al. 2001) and its econometric properties formalized by Hahn et al. (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). Some recent literature has applied RD to measuring promotional effects (Busse et al. 2006, 2010; Nair et al. 2011). RD measures causal treatment effects in situations where treatment is based on whether an underlying forcing variable crosses a threshold. With the treatment being the only discontinuity at this 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 forcing variable is the difference in the *AdRank* for the bidders in the higher and lower positions. If this crosses 0, there is treatment; otherwise there is not. Thus, the RD estimator of the effect of position finds the limiting values of the outcome of interest (e.g., CTRs or sales) on the two sides of this threshold of 0. This application satisfies the conditions for a valid RD design laid out in Nair et al. (2011) and thus we obtain valid causal effects of position.

While the search engine observes the *AdRank* of all of the bidders, the bidders themselves only observe their own *AdRank*. They observe their own bids, and the search engine reports the *Quality Score* to them *ex-post*. Hence they can construct their own *AdRank*, but they do not observe the bids or *Quality Scores* of their competitors. Because the forcing variable for the RD design is the difference between competing bidders *AdRank*, they cannot construct the forcing variable even ex-post. In addition, the typical modified second-price auction mechanism for position auctions eliminates the incentives for advertisers to second-guess what their competitors are bidding. This ensures the local randomization required for the RD design; this nonobservability of competitors' *AdRank* implies that advertisers cannot precisely select into a particular position. At the same time, this poses a challenge to the empirical researcher wishing to use RD in this context, who obtains data from one advertiser and only typically observes the *AdRank* for that advertiser, thus not allowing the forcing variable to be observed. However, we have obtained a unique data set that contains information on bids and *AdRank* and performance information for a focal advertiser and its main 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 in our data 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 *AdRank* and performance measures for advertisers in adjacent positions, and we can 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: CTRs and sales orders (i.e., whether the consumer who clicked on the search advertisement purchased at that or a subsequent occasion). We find that position positively affects CTRs, with higher positions getting greater clicks. However, these effects are highly localized, with significant effects only at certain positions and no significant effects at others. We find that position effects are largely insignificant when it comes to sales orders. The only exception is at the typical position where consumers have to scroll down to see the next advertisement. We also document several interesting findings about moderators for position effects. In general, we find stronger position effects for smaller advertisers, for keyword phrases where the advertiser generally has lower prior experience with the advertiser, for keyword phrases that are less specific about the product or brand, and where the advertiser allows the search engine to display the ad even where the advertised keyword phrase is not an exact match to the keyword phrase the consumer searches for (referred to

[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.

as "broad match" as opposed to "exact match"). We also document interesting differences in the position effects between weekdays and weekends; the differences may be driven by different search costs.

This paper makes several contributions to the literature. First, it makes causal inferences on position effects, which have been hard to make in the literature so far due to limitations in the data and the empirical strategies used. Second, this paper demonstrates the nature of the selection bias that can result in these contexts, not only due to observed factors such as the advertiser, keyword, etc., but also due to other unobserved factors that drive the strategic bidding behavior of firms. Third, it documents moderators to these effects across types of advertisers, type of keyword, the advertisement match-type, and across weekday versus weekend. (To our knowledge, this has not been documented before.) Finally, we present a novel application of RD to an important context where (to our knowledge) it has not been considered before, and where other ways of obtaining causal effects are typically infeasible. This approach could be applied by search engines and advertisers to find position effects with data that they already have, using relatively simple econometric methods and without the cost and effort involved in experimentation.

The rest of the paper is organized as follows. We give some background on search advertising in general and position effects in particular in §2. In §3, we discuss the selection of position in search advertising contexts, and extant approaches to address the issue. In §4, we discuss an RD approach to find causal effects. We discuss the data in §5 and the results of our empirical analysis in §6. Section 7 provides summarizing remarks including a discussion of the limitations of this research as well as avenues for future study.

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 is shown in Figure 1 of the results of a search for the phrase "golf clubs" on Google, the most popular search engine. Search advertising is a large and rapidly growing market. For instance, Google reported revenues of almost $14.9 billion for the quarter ending September 30, 2013, with a growth of 12% over the same period in the previous year. The revenues from Google's sites, primarily the search engine, accounted for 68% of these revenues.[2] According to the Internet Advertising Bureau, $16.9 billion was spent in the United States alone on search advertising in 2012. Search advertising is the largest component of the online advertising market, comprising 46% of all online advertising revenues in 2012. Despite the fact that it is a relatively new medium for advertising, search is the third largest medium after TV and Print, and surpassed Radio in 2012.[3]

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." Furthermore, these ads can be geography specific, with potentially different ads being shown across locations. This enables an advertiser to obtain a high level of targeting.

Search advertising is sold on a "pay for performance" basis, with advertisers bidding on keyword phrases. The search engine regularly conducts an automated online auction for each keyword phrase. The set of ads and their order are decided by the outcome of the auction. Advertisers only pay the search engine if a user clicks on an ad. Payment is on a per click basis (hence the commonly used term pay per click (PPC) for search advertising). By contrast, online display advertising is typically sold on the basis of impressions; the advertiser pays even if there is no behavioral response. In search advertising, advertisers can 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.

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. The bid consists of the maximum amount that the advertiser would pay the search engine every time a consumer clicked on the search ad. Because the search engine gets paid on a per click basis, revenue is maximized if the winning bidder has higher product of bids and clicks. Thus, Google ranks bidders not on their bids, but on a score called *AdRank*, which is the product of

[2] These data were obtained from Google's earnings report for Q3 2013, available at https://investor.google.com/earnings/2013/Q3_google_earnings.html (last accessed, October 31, 2013).

[3] Internet Advertising Bureau's report on Internet advertising can be accessed at http://www.iab.net/media/file/IABInternetAdvertisingRevenueReportFY2012POSTED.pdf (last accessed, October 31, 2013).

a bid, and a metric called *Quality Score* assigned by Google. While the exact procedure by which Google assigns a *Quality Score* to a particular ad is not publicly revealed, it is known that it is primarily a function of expected CTRs (which Google knows through historical information combined with limited experimentation), adjusted marginally up or down by factors such as the quality of the landing page of the advertiser. The positions of the search ads of the winning bidders are then in descending order of their *AdRank*. The winning bidder pays an amount just above what would be needed to win that position. Thus, the cost per click of the winning bidder in position i is given by

$$CPC_i = \frac{Bid_{i+1} \times Quality\ Score_{i+1}}{Quality\ Score_i} + \varepsilon, \qquad (1)$$

where ε 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. Because the position of an ad is the outcome of an auction, higher positions cost more for the advertiser (all else remaining equal) and hence would be justified only if they generate higher returns for the advertiser. 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 might be 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. Because 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 signals of quality.

A second mechanism might relate to consumers' learned experience about the relationship between position and the relevance of the advertisement. The auction mechanism of search engines such as Google inherently scores ads with higher relevance higher (Varian 2007). Over a period of time, consumers might have learned that ads in higher positions are more likely to be relevant to them. Because consumers incur 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. 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. Using an analytical model, Chen and He (2011) show that it is a viable equilibrium for advertisers with higher relevance to be positioned higher and consumers to be more likely to click on higher positions. By contrast, Katona and Sarvary (2010) and Jerath et al. (2011) discuss situations under which it may be optimal for firms to not be ranked in order of relevance or quality, and for clicks to also not necessarily be higher for higher placed search ads.

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 a triangular area that contains the top three ad positions above the organic results and the fourth ad position at the top right. Such an effect is particularly pronounced on Google and is often called the Google golden triangle (Hotchkiss et al. 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. Hence, consumers may focus first on the top positions in the organic search results. Because search advertising results are above or beside 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 for position effects. In general, whether there are significant position effects at a particular position is an empirical question.

2.3. Moderators of Position Effects

Whether there are position effects at particular positions is an interesting first order question in itself. However, advertisers may also be interested in learning if there are moderators to these effects, across different types of advertisers, different types of keywords, etc. Additionally, advertisers in search engines such as Google have decisions to make about the nature of targeting. Specifically, they need to decide whether to bid for a keyword to appear as a broad match ad (where the advertisement is shown when the keyword phrase searched for by the consumer is close to but not an exact match to the advertised keyword phrase) or an exact match.

How advertising effects vary by advertiser has been the topic of considerable interest in prior research. For example, Lodish et al. (1995) provide an empirical generalization that advertising elasticities vary widely across brands. Advertising elasticities are higher for durables than for nondurables (Sethuraman and Tellis 1991). They also decrease over the product life cycle and hence are higher for new brands compared to existing brands (Parker and Gatignon 1996). Advertising has been found to be more effective for experience goods than for search goods (Hoch and Ha 1986, Nelson 1974). However, to our knowledge, there has been no research on how position effects in search advertising vary across advertisers. This is an important question to study since some studies in the theoretical literature on position auctions (see for instance Varian 2009) assume that position effects (for example the ratio of CTRs across positions) are independent of the advertiser. To our knowledge, this assumption has not been empirically tested until now. Other theoretical studies (e.g., Jerath et al. 2011) have pointed to mechanisms by which position effects may vary for higher quality and lower quality firms. Because our data set contains information for multiple firms that are selling similar categories in the same time period, we have a unique opportunity to empirically examine how position effects vary across advertisers.

Another aspect of advertising that has received attention in the literature is that prior experience with the product or firm is a substitute for advertising. For example Deighton et al. (1994) show that advertising is effective in attracting consumers who have not recently purchased the brand (low recent experience) but that advertising does little to change the repeat-purchase probabilities of consumers who have just purchased the brand (high recent experience). Ackerberg (2001) reports that advertising's effect on inexperienced consumers is positive and significant, whereas it has a small and insignificant effect on experienced consumers. Narayanan and Manchanda (2009) also find that experience and advertising are substitutes in the context of a learning model. In the context of our data, one of the key components of the search engine assigned *Quality Score* is the historic CTR that the firm obtains on its ads. It also reflects the overall historic CTR of all of the ads and keywords for the advertiser, and the quality of the advertiser's landing page. Because a searcher who has clicked through in the past has actually experienced the firm's online offering, the different components of *Quality Score* together provide a plausible measure of the stock of aggregate prior digital experience that searchers have with the firm. Analyzing how position effects vary for ads with different *Quality Scores* allows us to examine the linkage between position effects and prior experience in the context of search advertising.

Prior research has also focused on the distinction between category and brand terms in search advertising. For example the keyword *Rayban sunglasses* would be classified as a brand phrase because it contains specific brand/product information, whereas the keyword *Sunglasses* would be categorized as a category phrase. Prior research has focused on the sequential use of category searches followed by brand searches. Lamberti (2003) reports that consumers initiated their search process with category terms and used more specific brand terms later in the purchase process. Rutz and Bucklin (2011) report that increased exposure to category advertising terms leads to increased searches of brand terms. So there is some evidence that the use of category terms precedes the use of brand terms. These findings are consistent with the literature on product category expertise. As per this literature, consumers who are early in the purchase process know little about the product category or the underlying attributes in a product or service. They may not know what questions to ask (Sheth et al. 1999 Chapter 14, p. 534) and may have a limited consumption vocabulary (West et al. 1996). This would explain the use of broad category terms early in the search process. In addition to the sequence of use, there are implications for how position effects vary for category terms versus brand terms. Because searchers who use category terms are early in the search process and have a lower level of knowledge and experience with the product, they might rely on the information in position more than searchers who use brand terms. To our knowledge, this distinction in position effects has not been examined in prior empirical work.

Google and other search engines also allow advertisers to set match criteria when they bid on ads for particular keyword phrases. On Google, these match types include broad match, where the ad is displayed when there is an imprecise match between the consumer's search term and the keyword phrase that the advertiser is bidding on, and exact match, where the ad is displayed only in cases where the two match exactly. As a result the headline and ad copy of the exact match ad will match the consumer's query better when compared to a broad match ad. To put it differently, with exact match ads advertisers can provide more precise information to consumers with the headline and ad copy compared to a broad match ad. Because there is more information in exact match ads, we can expect position effects to be less salient compared to those for broad match ads. This is consistent with the prior discussion of high versus low quality keywords, and category versus brand keywords, and how information or experience substitute for position. To our knowledge, this distinction between broad and exact match types has not been examined in the existing literature.

The distinction between weekday and weekend effects is also important in retailing. For example Warner and Barsky (1995), in the context of offline retailing, suggest that consumer search costs are lower on the weekends. If this is true, they are more likely to search lower down the advertising results on a search engine page before stopping. This would imply that position effects are stronger on weekdays compared to weekends. We examine this distinction in our empirical analysis.

3. Selection Issues

3.1. Selection on Observables

As discussed in the previous section, measuring causal position effects is of critical importance to the advertiser. However, there are likely significant selection biases in naive estimates. 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 advertisers, keywords, match types, days, etc. This is a common strategy in empirical work. Consider the case wherein we observe positions and outcomes for a set of keywords. There are likely to be systematic differences in CTRs across different keywords. Because of the auction mechanism itself, keywords with high CTRs in the past would typically have higher positions because they are assigned higher *Quality Scores*. Because these keywords are also likely to get higher clicks in the future, there is a correlation between position and CTRs that is not causal, but is instead driven by the auction mechanism. Similar arguments can be made about spurious effects when pooling across advertisers, match-types, etc. If panel data are available, fixed effects for keywords, advertisers, match type, etc., could eliminate these selection biases.

3.2. Selection on Unobservables

In addition to selection biases for observables, there is potential for selection on unobservables. For example, selection may also be induced by the bidding behavior of advertisers. Advertisers in the search advertising context often use bidding engines to decide bids for keyword phrases. Typically, they address a very large number of keywords. For instance, the advertisers in our empirical application bid on tens of thousands of keywords on any given day. These bidding engines can be programmed to use specific bidding rules, with adjustments made to these rules on a case-by-case basis. For instance, advertisers often set a fixed advertising-to-sales ratio for deciding advertising budgets. In the search engine context, this involves a continuous feedback loop from performance measures to the bidding engine. As sales per click increases, the bidding engine might be programmed to 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 other rules used by advertisers in setting their bids. Consider an advertiser that has periodical sales, with a higher propensity of consumers visiting 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 gotten 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. Because search advertising positions are determined through a competitive bidding process, the competitors' bidding behavior could also induce biases in naive 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. Because of mechanisms similar to those described above, competing bidders may place high or low bids when their sales are high. Because the competing bidder offers products similar to the focal advertiser, higher sales for the competing bidder, for instance due to a price promotion, may lower the sales for the focal advertiser. Even CTRs for the focal advertiser could be affected if the search advertising listing for the competitor mentions that there is a price promotion at that website. At the same time, the competing bidder may place a low bid on the keyword auction through a similar set of mechanisms as those described earlier in this subsection, thus pushing the focal advertiser higher in position. This negative correlation between position and sales for the focal advertiser induced by the price promotion at the competing advertiser's website and the unobserved strategic bidding behavior by the competitor would be picked up as a position effect by a naive analysis. In general, any unobservables that affect positions through the bidding behavior of the competing advertiser may also affect outcomes such as sales and CTRs for the focal advertiser, and this would induce selection biases.

To summarize, there are significant selection issues that may render naive estimates of positions highly unreliable, with unpredictable signs and magnitude of the biases induced by selection on unobservables.

3.3. Extant Approaches to Address Selection

As mentioned in §1, position effects have been studied in the literature. An early study of the position effect was Agarwal et al. (2011), which concluded that CTRs decrease monotonically as one moves down the search advertising listings, but conversion rates increase and then decrease. This study controls for heterogeneity across keywords, but not across match types, days, etc. Furthermore, 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. For robust results for even the main effect of positions, the experiment would need to be carefully designed to randomize bids across the various keywords, positions, days of the week, match type, etc. A small number of observations would typically be insufficient. Exploring moderators for these effects would greatly increase the number of observations needed. Irrespective of the scale of the experiment, it cannot eliminate selection biases induced by competitors' strategic bidding behavior. For instance, if a competing advertiser bids lower on days when they have sales promotions than on other days, their low bids can drive the focal advertiser's position to be higher, all else remaining equal. Because of the sales promotion at this competing advertiser's website, the clicks and sales at the focal advertiser may be lower. This would lead to a negative correlation between positions and outcome variables such as clicks and sales. This spurious correlation cannot be eliminated by randomizing the bids of the focal advertiser alone.[4] This discussion demonstrates why experimentation is difficult in this context, as it would require randomization of bids of all bidders, i.e., the focal bidder *and* its competitors, and on a large scale. This is typically infeasible for a given advertiser. A search engine could randomize positions of all advertisers, but large scale experimentation is difficult for the search engine, particularly without the concurrence of all of the advertisers. Furthermore, it is expensive for the search engine to randomize the positions on a large scale; experimental pages typically do not generate revenues for it.

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 the outcome and position equations (Ghose and Yang 2009, Kalyanam et al. 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 is hard to generate 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. It would be difficult to control for the selection issues induced by competitors' strategic bidding behavior using such a specification. Furthermore, there is 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. Moreover, such an approach is computationally demanding and requires that the researcher have access to appropriate exclusion restrictions that are necessary for identification of parameters.

A third approach, adopted by Rutz and Trusov (2011), is to instrument for the position. Because a valid instrumental variable with sufficient variation is hard to come by, this study uses the latent instrumental variables approach of Ebbes et al. (2005) to account for the potential endogeneity of position. The method relies on several crucial assumptions, including normality of the outcome equation and departures from normality for the position equation. While the latter is not problematic, the former may be more so. As we will see in our empirical application, outcomes such as CTRs and sales are highly non-normal. For instance, periodic sales and promotional events, if unobserved in the data, would induce a skewed and potentially multimodal distribution of the outcome variables. This would make the approach challenging in many contexts. Furthermore, this approach relies on a single latent instrument variable model, implicitly assuming that selection effects do not vary across position. Finally, this approach relies on the assumption that position effects can be modeled with a parametric specification, whereas the complex multiple mechanisms underlying selection in position (for instance, own strategic bidding, competitive bidding, with the set of competitors being potentially different across positions) imply that the selection bias is likely to be highly local. In this case, a parametric specification would be subject to the potential for specification bias.

To summarize, to our knowledge, the extant literature has either ignored the endogeneity/selection issues, or taken parametric approaches to control for endogenous positions. Ignoring the endogeneity/selection issues would be problematic. Parametric approaches to control for endogenous positions

[4] In the context of advertising on the Google search engine, positions are reported only on a daily basis. Furthermore, data on clicks or sales within a shorter period of time within a day would typically have a lot of zeroes and hence lower variation. Across multiple days of experiments, it would be hard to make the case that competing advertisers do not vary their bids in a manner that induces selection in positions.

would have significant limitations.[5] Furthermore, this is a situation wherein experimentation, the typically advocated approach for obtaining causal effects, is usually infeasible.

4. Applying Regression Discontinuity to Finding Position Effects

4.1. Regression Discontinuity

RD designs can be used to measure treatment effects when treatment is based on whether an underlying continuous forcing variable crosses a threshold. If 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. A necessary condition for the validity of the RD design is that the forcing variable itself is continuous at the threshold (Hahn et al. 2001). This is achieved in the typical marketing context if the agents are uncertain about the score or the threshold (Nair et al. 2011).

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

$$y^+ = \lim_{\lambda \to 0} \mathbb{E}[y \mid z = \bar{z} + \lambda], \tag{2}$$

$$y^- = \lim_{\lambda \to 0} \mathbb{E}[y \mid z = \bar{z} - \lambda]. \tag{3}$$

Then the local average treatment effect is given by

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

Practical implementation of RD involves finding these limiting values nonparametrically using a local regression, often simply a local linear regression within a prespecified bandwidth λ of the threshold $\bar{z}$ and then assessing sensitivity to the bandwidth. More details on estimating causal effects using RD designs, including the difference between sharp and fuzzy RD designs, the selection of nonparametric estimators for y^+ and y^-, the choice of bandwidth λ, and the computation of standard errors can be found in Hahn et al. (2001) and Imbens and Lemieux (2008).

[5] We have not explicitly compared the estimates from our approach to those from these alternative approaches because such comparisons are beyond the scope of this paper. Moreover, they would not be "apples to apples" comparisons due to the differences in the nature of data used in the various approaches. We conduct more formal comparisons of the methodologies in Narayanan and Kalyanam (2014).

4.2. RD in the Search Advertising Context

As described earlier in §2.1, positions in search advertising listings are determined by an auction. Bidders are ranked on a variable called *AdRank*, which in turn is the product of the bid. The *Quality Score* is assigned by Google to the bidder for each specific keyword phrase for a particular match type. The application of RD in 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 forcing variable for the RD design in this difference is *AdRank* 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. This achieves the quasi-experimental design that underlies RD. The latter set of observations acts as a control for the former.

For an RD design to be valid, the only source of discontinuity should be the treatment. One consequence of this condition is that RD is invalidated if there is selection at the threshold. If an advertiser can select his bid so as to have an *AdRank* just above the threshold, the RD design would be invalid. However, in establishing the validity of RD we are aided by the modified second-price auction mechanism used by Google. As per this mechanism, the winner actually pays the amount that ensures that its ex-post *AdRank* is just above that of the losing bidder. Specifically, the cost per click for the advertiser is determined as in Equation 1, which ensures that ex-post, the following is true:

$$\Delta AdRank_i \equiv (AdRank_i - AdRank_{i+1}) > \varepsilon, \tag{7}$$

where ε is a very small number. An important consequence of this modified second price mechanism is that it is approximately optimal[6] for advertisers to set bids so that they reflect what the position is worth to them as opposed to setting bids such that they are just above the threshold for the position. Thus, the second-price auction design eliminates incentives for advertisers to second guess their competitors bids and

[6] See Varian (2007) for a discussion on this.

put in their own bids so as to have an *AdRank* just above that of their competitors.

Furthermore, *AdRank* are unobserved ex-ante by the advertiser. Their own *AdRank* are observed ex-post, as Google reports the *Quality Score* at the end of each day, and the advertiser observes only his own bid ex-ante. However, *AdRank* of competitors are not observed even ex-post. Thus, the advertiser cannot strategically self-select to be just above the cutoff. Occasions when the advertiser 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. The fact that *AdRank* of competitors are unobserved satisfies the conditions for validity of RD laid out in Nair et al. (2011), with the advertiser being uncertain about the score (Δ*AdRank*).

Typically, only the search engine observes the *AdRank* for all advertisers. Therefore, the RD design could be applied by the search engine, but not by advertisers or researchers who have access to data only from one firm. Unfortunately, search engines such as Google are typically unwilling to share data with researchers, partly due to the terms of agreement with their advertisers. However, we have access to a data set where we observe *AdRank* for four firms in the same category. One of these firms acquired the three other firms in this set; 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 §5.

It would also be relevant at this stage to discuss the role of other unobservables in this approach. In our empirical application, we have observations for four firms in the category, which constitute an overwhelming share of sales and search advertising in this market. However, it is possible that there are other advertisers that we do not observe in our data set. This is not problematic in our context. Our analysis is only conducted on those sets of observations where we observe *AdRank* for pairs of firms within our data set. Because our interest is in determining how position affects outcomes, all else remaining constant, we conduct a within firm, within keyword, within match type, and within day-of-week analysis. The *AdRank* data for the firms and competitors is only used to classify which observations fall within the bandwidth for the RD design. Thus, the presence of other firms not in our data set does not affect our analysis.

4.3. Implementing the RD Design to Measure Position Effects

In this subsection, we describe implementation of the RD design to measure the effect of position on CTRs. An analogous procedure can be easily designed to measure position effects on other outcomes such as conversion rates, sales, etc.

Consider the case wherein we wish to find the effect of moving from position $i+1$ to position i on the CTR. Note that the $(i+1)$th position is lower than the ith position. Let y_j refer to the outcome (e.g., CTR) for the advertisement j (which is a unique identifier of a particular keyword phrase, an advertiser, a specific day, and a specific match type). Let $AdRank_j$ refer to the *AdRank* for that ad, and pos_j refer to the position of the ad in the search engine listings. The following steps are involved in implementing the RD design to measure the incremental CTRs of moving from position $i+1$ to position i.

1. Select observations for which we observe *AdRank* for competing bidders in adjacent positions; the forcing variable z_j for the RD design is the difference between the *AdRank* of adjacent advertisers, i.e., Δ*AdRank*. For an advertiser in position i, z_j is the difference between that advertiser's *AdRank* and that of the advertiser in position $i+1$ and has a positive value. For an advertiser in position $i+1$, z_j is the difference between the advertiser's *AdRank* and that of the advertiser in position i and has a negative value.

2. Select the bandwidth λ for the RD. This could be an arbitrary small number, with the researcher checking for robustness of estimates to the choice of bandwidth. In our case, we select the optimal bandwidth using a "leave one out" (LOO) criterion described later in this section.

3. Retain observations with scores within the bandwidth λ. The RD design compares observations for which $0 < z_j < \lambda$ with those for which $-\lambda < z_j < 0$. Thus, retain observations for which $|z_j| < \lambda$.

4. Find the position effect using a local linear regression. One could use a local polynomial regression but a local linear regression works well in this context due to its boundary properties (see for instance Fan and Gijbels 1996, Imbens and Lemieux 2008). The local linear regression is the following regression applied to the subset of the data within the bandwidth:

$$y_j = \alpha + \beta \cdot 1(pos_j = i+1) + \gamma_1 \cdot z_j + \gamma_2 \cdot z_j 1(pos_j = i+1) + f(j{:}\theta) + \varepsilon_i. \tag{8}$$

In this regression, $1(pos_j = i+1)$ is an indicator for whether the ad is in the higher position, and β is the position effect of interest. The γ_1 and γ_2 parameters control for the variation in CTR with changes in the forcing variable, allowing the variation to be different on the two sides of the threshold through an interaction effect. The $f(j{:}\theta)$ term includes a set of fixed effects with θ being the parameter vector.

5. Find the bandwidth $\lambda^\star$ using a "leave one out cross validation" (LOOCV) optimization procedure described in Ludwig and Miller (2007) and Imbens and Lemieux (2008). This involves leaving out one observation y_k at a time, and finding the parameter estimates using the remaining observations. These estimates are then used to find a predicted value $\hat{y}_k$ for that left out observation. Because RD estimates involve finding the limiting values of the outcomes on the two sides of the threshold, we use only observations very close to the threshold (i.e., with $\tilde{\lambda} < \lambda$) for the cross validation. The criterion used to find the optimal bandwidth is the mean squared error of these predictions

$$C_Y = \frac{1}{N} \sum_{\{k: -\tilde{\lambda} < z_k < \tilde{\lambda}\}} (\hat{y}_k - y_k)^2, \tag{9}$$

where N is the number of observations with forcing variable z_k lying between $-\tilde{\lambda}$ and $\tilde{\lambda}$. The optimum bandwidth is then

$$\lambda^* = \arg\min_{\lambda}(C_Y). \tag{10}$$

5. Data Description

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 stores both through organic growth and acquisition of other retailers. Because 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 two others are more narrowly focused on specific subcategories. However, each of them has significant overlaps with the others 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 include 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 28.5 million daily observations over a period of nine months in the database, of which about 13.1 million observations involve cases where two or more advertisers among the set of four firms bid on the same keyword. Because the keywords are often not in adjacent positions, we filter out observations where the observations are not adjacent. We also drop observations where we do not have bids and *Quality Scores* for both of the adjacent advertisements. Because the position reported in the data set is a daily average, we also drop observations where the average positions are more than 0.1 positions away from the nearest integer.[8] We are thus left with a total of 414,310 observations where we observe advertisements in adjacent positions, spanning 22,825 unique keyword phrase/match type combinations. An overwhelming majority (79.4%) of the 22,825 keywords are of the broad match type, and the rest are of the exact match type. There are a total of 19,205 unique keywords in this analysis data set, with most exact match type keywords also advertised as broad match type, but not necessarily vice versa.

Table 1 shows the list of variables in the analysis data set (including variables we have constructed such as CTRs, conversion rates, and sales per click) and the summary statistics for these variables. Observations are only recorded on days that have at least one impression, i.e., when the ad was served at least once. Through a tracking of cookies on consumer's computers, each click is linked to a potential order, sales value, margin, etc. As per standard industry practice, a sales order is attributed to the last click from a search ad within an attribution window with previous clicks not getting credit for these sales.

To summarize, we have obtained a unique data set consisting of information at a daily level on keywords, type of match, bids, quality score, and key performance metrics for the advertisement. To our knowledge, this is the first time that a data set has been assembled that includes information on *AdRank*, clicks, and sales outcomes for competing advertisers.

6. Results

We conducted an analysis of the effect of position on two key metrics of interest to advertisers: CTRs and the number of sales orders (henceforth orders). We selected these two metrics because they are the most important metrics from the advertiser's point of view.

[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.

[8] Google reports only the daily average positions for any ad. Variation in position within a day may be because of many factors including different ads being served in different geographies, and a limited degree of experimentation by Google itself to calibrate *Quality Scores* for all advertisers.

Table 1 Summary Statistics of the Data

Variable	Mean	Std. dev.
Impressions	45.9050	224.6952
Clicks	0.5452	2.5371
Click-through rate (%) (Clicks/Impressions)	1.8932	6.6629
Number of orders	0.0046	0.0733
Conversion rate (% of non-zero clicks that resulted in orders)	0.7530	7.3889
Sales ($)	0.4850	16.7715
Average sales per (non-zero) click ($)	0.7446	20.7283
Average sales per (non-zero) order ($)	107.3475	225.7770
Bid ($ per click)	0.3947	0.8231
Quality Score	5.9811	1.2473
AdRank	2.3402	5.0966

CTR measures the proportion of consumers who saw the ad, clicked on it, and arrived at the advertiser's website. Because the advertiser's control over the consumer's experience 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 of raw clicks instead, but it does not make any material difference to the results, and CTR is the more commonly reported metric in this industry.

The second measure we consider is the number of 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 of measures such as conversion rates, sales value, and sales per click, but do not report these estimates since almost all of them were statistically insignificant. This is partly driven by the fact that the category in question sees very infrequent purchases, thus reducing the statistical significance of results.

6.1. Effect of Position on CTRs

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 2. The table shows the ordinary least squares (OLS) estimates (which in essence are the same as mean comparisons across pairs of adjacent positions), fixed effect estimates (with fixed effects added for advertiser, keyword, match type, and day of week to control for selection in observables) and the RD estimates. The RD estimates are reported at the optimum bandwidth as described in §4.3. We report the baseline CTRs for each position, which is the mean CTR for the lower position in the pair. The baseline is the same for the raw and fixed effects estimates. The baseline for the RD estimates are different and are mean CTRs for the lower position for observations within the optimized bandwidth. Note 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 position 3 to 2. Hence, the baseline for position 2 will not be the sum of the baseline for position 3 and the effect of moving from position 3 to 2.

When we look at the RD estimates we see significant effects across multiple positions. The RD estimates are significant from position 2 to 1. 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,[9] which has been postulated to be due to attention effects and documented in eye tracking studies (Hotchkiss et al. 2005, Guan and Cutrell 2007) as well as using advertising and sales data (Kalyanam et al. 2010). Furthermore, 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.).

For position 2 to 1, the OLS and fixed effects estimates are also significant. The OLS estimate shows an increase in CTR of 2.4413 which is an increase of 133.2%. The fixed effect estimate is 0.4286 which is an increase of 23.4%. The RD estimate shows an increase of 0.4796 which is a 20.6% increase. The key point here is that there are very significant biases in the position 1 effect in the OLS and fixed effects estimates. There is a similarly large bias for the position 3 to 2 effect where the RD estimates are not significant, but the OLS and fixed effects estimates are. For position 4 to 3, the RD estimate is 0.1109, which is an increase of 10.7%. The OLS estimate is much higher than this with an increase of 64.0%. The fixed effects estimate is 5.4%, which is lower than the RD estimate and is significant only at the 90% level. The fixed effect estimates are insignificant for position 6 to 5 and position 7 to 6. The OLS estimate is significant for position 6 to 5 but the percentage increase is much higher than that for the RD estimate.

The differences between the OLS, fixed effects, and RD estimates are important, as they indicate the nature of the selection in positions. The OLS estimates are generally highly positively biased and demonstrate the selection on observables (e.g., keyword,

[9] The Google golden triangle effect refers to the finding that consumers focus on the top organic listing, top advertising listing, and then on results lower down, skipping some positions in between.

Table 2 Position Effects on CTRs

	Naive estimates (CTR %)				RD estimates (CTR %)		
Position	Baseline	Num. of obs.	OLS estimate (*p*-value)	Fixed eff. estimate (*p*-value)	Baseline	Num. of obs.	Estimate (*p*-value)
2 to 1	1.8325	141,492	2.4413 (0.0000)	0.4286 (1.39e−23)	2.3260	4,736	0.4796 (0.0049)
3 to 2	1.1597	192,346	0.9217 (0.0000)	0.0643 (0.0138)	1.2864	14,134	0.0597 (0.2884)
4 to 3	0.8732	124,488	0.5588 (0.0000)	0.0470 (0.0690)	1.0348	14,350	0.1109 (0.0147)
5 to 4	0.7280	63,754	0.2642 (7.82e−09)	−0.0406 (0.1679)	0.8621	9,318	0.0570 (0.2069)
6 to 5	0.6405	26,564	0.1394 (0.0198)	−0.0318 (0.4126)	0.7746	4,532	0.1294 (0.0147)
7 to 6	0.4998	11,258	0.1458 (0.1729)	0.0479 (0.4286)	0.6345	1,864	0.1239 (0.0474)
8 to 7	0.2740	4,960	0.1708 (0.0703)	0.0583 (0.3185)	0.3855	426	−0.0058 (0.9601)

advertiser) as well as unobservables. The fixed effects estimates, which correct for selection on observables such as keyword, advertiser, match type, and day of week, on the other hand, are generally downward biased. This suggests that in this context, the selection on unobservables causes a negative bias. This can result from advertisers or their competitors' strategic behavior, as indicated earlier. Furthermore, the effect of selection differs significantly by position, with the magnitude and even signs of the bias varying with position. This last point has some important implications. Parametric approaches to control for selection must be able to allow for selection effects to vary by position. Instrumental variable approaches need to allow for the possibility that the choice of instrument varies by position.

The causal position effects are not just statistically significant, but have large economic significance as well. For instance, the causal effect at position 1 as a proportion of the baseline CTR is 20.6%. They are 10.7%, 16.7%, and 19.5%, respectively, at positions 3, 6, and 7, and hence of large magnitude even at these positions. Thus, it seems that, in this category at least, if the objective of search advertising is to drive up clicks, there are opportunities at these positions and by a large magnitude. Furthermore, there is little agreement between the OLS, fixed effects, and RD estimates; the OLS estimates generally overstate the true effect and the fixed effects estimates generally understate them. The magnitude of the bias is quite significant and varies by position, indicating that selection effects vary by position.

6.2. Effect of Position on Orders

We next investigate whether ad position in search advertising results causally affects the number of orders that are generated. The effect of position on sales is a combination of the effect on clicks and on conversion from clicks to orders. There are multiple mechanisms by which position might affect conversions and therefore sales. For instance, an ad in a higher position might generate more clicks due to the greater attention that it would attract, but the incremental clicks may be from consumers who are only marginally interested in the category. If that is the case, higher positions might generate more clicks, but with lower conversion rates to sales. On the other hand, a higher position could positively affect conversion rates if position plays a quality signaling role, i.e., in equilibrium consumers credibly believe that higher quality advertisers are in higher positions and firms' bidding strategies align in the same way (Athey and Ellison 2011, Chen and He 2011). The strength of this effect depends on the degree of prior experience the consumer has with the advertiser and the products being advertised, and the extent to which it is a search versus experience good. Jerath et al. (2011) present the argument for why a lower quality firm might have incentives to bid for a higher position and vice versa, and discuss how the search engine itself might have similar incentives to place a lower quality firm in a higher position through the *Quality Scores* it assigns to the advertisers. Thus, the effect of position on conversion rates is complex; different mechanisms drive effects with opposite signs. The existence and sign of the effect of position on conversion from clicks to sales is ambiguous, and consequently the effect of position on orders is an empirical question.

We report the RD estimates for orders in Table 3. We find that the OLS and the fixed effects estimates are once again misleading. They suggest that there are positive incremental effects on sales only when moving to the top position from the next one. By

Table 3 Position Effects on Number of Sales Orders

	Naive estimates (CTR %)				RD estimates (CTR %)		
Position	Baseline	Num. of obs.	OLS estimate (*p*-value)	Fixed eff. estimate (*p*-value)	Baseline	Num. of obs.	Estimate (*p*-value)
2 to 1	0.0045	141,492	0.0049 (4.75e−13)	0.0009 (0.0550)	0.0042	4,338	0.0006 (0.7299)
3 to 2	0.0023	192,346	0.0036 (8.93e−15)	0.0005 (0.1757)	0.0023	12,116	0.0007 (0.3656)
4 to 3	0.0021	124,488	0.0018 (0.0001)	−0.0002 (0.5034)	0.0018	13,332	−0.0009 (0.1549)
5 to 4	0.0019	63,754	0.0006 (0.2628)	−0.0002 (0.6777)	0.0017	9,206	0.0011 (0.1695)
6 to 5	0.0011	26,564	0.0001 (0.8470)	−0.0005 (0.3203)	0.0008	4,630	0.0017 (0.0409)

contrast, the RD estimates suggest that the only significant effect is in moving from position 6 to 5, with no significant differences between pairs of positions above that. While we do not report them in the paper, we also investigated the effect of position on conversion rates, and found the effects to be insignificant in most positions, except when going from position 6 to 5, where there is a positive effect.[10] While we cannot conclusively say what mechanism drives these results, a plausible interpretation is that consumers do not see any quality signals in the specific positions, but do see a quality signal in whether the ad is in the top five positions. Such a step jump might also be facilitated by the fact that the top five positions typically lie above the page fold, and hence do not require the consumer to scroll down to see the ad.

Finally, we also examine the economic significance of the sales effects. We find that these effects are even stronger than for CTRs, with orders increasing by over 212.5% relative to the baseline, compared with the 10.7% to 20.6% impact in the case of clicks.

6.3. Advertiser-Level Effects

We next investigate how position effects vary across advertisers. Table 4 reports position effects for the keywords advertised by three of the four firms in the data. We were unable to conduct advertiser-level analysis for the fourth firm because of the significantly smaller number of observations in that case. While we are unable to name the firms, the firms labeled 1 and 2 are smaller than firm 3, which is the largest firm in the category. We find that the position effects for firms 1 and 2 are largely very similar to the effects for the pooled analysis reported earlier, with significant effects at positions 1, 3, 5, and 6. However, firm 3 has largely insignificant effects, except at position 5, which is typically the position at which consumers have to scroll down the page to see the next ad. Thus, firm 3, which is the largest of the three firms, has much weaker position effects than firms 1 and 2, which are smaller. This could reflect the substitutability between advertising and other sources of information about quality, in this case the firm size. Consumers may be using the position to infer something about the quality of the product advertised or the relevance of the ad to them from the position when the firm is smaller. Conversely, when the firm is larger, it is possible that firm size provides them this information and hence advertising has weaker or no effects. Another possibility is that consumers, on average, have greater prior experience with the larger firm than the two smaller firms. This results in weaker position effects for the bigger firm because of the substitutability between advertising and prior experience. These are very interesting results, however, with only three advertisers in our analysis, it is hard to conduct a more detailed analysis about position effects and differences in quality or experience at the advertiser level. Fortunately, there are lots of keywords in our data set, making a more detailed analysis possible at the keyword level.

We further investigate the relationship between prior consumer experience and position effects by using information on the *Quality Scores* for keyword/advertising combinations as a proxy for prior consumer experience. Because the search engine assigns a *Quality Score* to a particular advertiser for a given keyword based on past click-through history, an advertiser on whose ad more consumers have clicked in the past receives, on average, a higher *Quality Score* than another advertiser with fewer clicks. We therefore split the data into two parts. In one part the *Quality Score* of the advertiser for the keyword is higher than the overall median *Quality Score*. In the other part, it is below the median. These estimates are reported in Table 5. We find that position effects for the low *Quality Score* keywords look similar to the pooled results, with significant position effects at

[10] These results are available from the authors on request.

Table 4 Advertiser Level Position Effects

	Firm 1 (CTR %)			Firm 2 (CTR %)			Firm 3 (CTR %)		
Position	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)
2 to 1	2.2190	1,157	0.4614 (0.0108)	1.9488	781	0.4821 (0.0177)	2.8730	2,176	0.3555 (0.2314)
3 to 2	1.1711	3,703	0.1609 (0.4004)	1.0441	4,467	−0.1431 (0.1846)	1.6288	5,209	0.0833 (0.4999)
4 to 3	0.9814	5,002	0.1588 (0.0406)	1.1225	5,520	0.1182 (0.0423)	0.9428	3,433	0.0840 (0.1520)
5 to 4	0.7479	3,390	0.0756 (0.3014)	0.9126	3,910	0.0970 (0.1582)	0.7809	1,809	0.0363 (0.3116)
6 to 5	0.7627	1,649	0.1862 (0.0596)	0.8391	1,928	0.0544 (0.0400)	0.7291	715	0.1589 (0.0431)
7 to 6	0.7409	712	0.1421 (0.0685)	0.8062	792	0.1120 (0.0319)	0.4803	266	0.2575 (0.1652)
8 to 7	0.2246	326	0.0574 (0.8062)	0.5721	392	0.1578 (0.2116)	−0.2922	118	−0.4110 (0.1755)

positions 1, 3, and 5. On the other hand, the position effects are largely insignificant for the high quality score keywords, with significant effects at the 90% level only at position 6. To the extent that quality score is a proxy for prior consumer experience, this provides support to the argument that prior experience and advertising act as substitutes. Note that this is an argument about consumer experience in the aggregate sense, not for a particular consumer.

The results on low versus high *Quality Score* keywords is consistent with the findings of Blake et al. (2015), although that paper looks at the question of overall effects of participating in search advertising auctions versus not participating, and uses individual consumer-level variation in past experience. They find that search ads have a greater effect on consumers who had not purchased on eBay before and had a smaller frequency of purchase in the past. Higher *Quality Scores* are consistent with a higher level of aggregate consumer experience in the past. Thus, our finding that position effects are stronger for low *Quality Score* keywords is consistent with their finding on the relationship between past consumer experience and advertising effects.

6.4. Brand vs. Category Keywords

We next investigate moderation in position effects based on the degree of specificity of the keyword phrases themselves. For the purpose of this analysis, we first classified keyword phrases as brand phrases if there was any reference to a specific brand or product in the keyword phrase. All other phrases were termed as category phrases. For example the keyword "*Rayban sunglasses*" would be classified as a brand phrase whereas the keyword "*Sunglasses*" would be categorized as a category phrase. For this

Table 5 Position Effects on CTRs: Low vs. High Quality Score Keyword Phrases

	Low quality score (CTR %)			High quality score (CTR %)		
Position	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)
2 to 1	1.2323	2,370	0.5725 (0.0308)	4.1504	2,510	0.2819 (0.4691)
3 to 2	0.8887	6,930	−0.0118 (0.8908)	2.1678	7,310	0.0884 (0.4754)
4 to 3	0.6503	7,608	0.0809 (0.0188)	1.5447	7,334	0.1955 (0.1567)
5 to 4	0.5732	4,686	−0.0279 (0.5770)	1.3554	4,710	0.0906 (0.3659)
6 to 5	0.4070	2,244	0.1761 (0.0185)	1.1410	2,324	−0.0041 (0.9677)
7 to 6	0.2814	974	−0.0112 (0.8946)	0.7245	902	0.3351 (0.0653)
8 to 7	0.2158	220	0.2734 (0.2505)	0.6057	186	−0.1199 (0.5986)

classification, we used a textual analysis of the keyword phrases, with the algorithm searching for occurrences of brand names and specific product identifiers (product name, model number, etc.) in the keyword phrase.[11] Table 6 reports the position effects separately for brand and category keywords. Consistent with our expectation, we find weaker effects for brand keywords than for category keywords. For brand keywords, there are significant effects only at positions 5 and 6, which are the typical positions where consumers have to scroll down to view the next listing. For category keywords, in addition to these positions, there are significant effects at positions 1

[11] Specific details on the procedure used for this classification are available from the authors on request.

Table 6 Position Effects on CTRs: Brand vs. Category Keywords

	Brand keywords (CTR %)			Category keywords (CTR %)		
Position	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)
2 to 1	2.4785	2,626	0.3480 (0.1252)	2.2894	1,930	0.6177 (0.0211)
3 to 2	1.3927	9,276	0.2397 (0.7604)	1.2318	5,574	0.0491 (0.5868)
4 to 3	1.0449	8,418	0.1505 (0.1274)	1.0017	5,938	0.0740 (0.0281)
5 to 4	0.9350	5,320	0.0112 (0.8570)	0.8219	4,188	0.1547 (0.1711)
6 to 5	0.8255	2,630	0.0975 (0.0505)	0.7498	1,768	0.1325 (0.0465)
7 to 6	0.6952	1,158	0.0550 (0.0525)	0.5507	654	0.2870 (0.0398)
8 to 7	0.4970	276	−0.0843 (0.5768)	0.3589	210	0.1134 (0.5905)

Table 7 RD Estimates of Position Effects on CTRs: Broad vs. Exact Match

	Broad match (CTR %)			Exact match (CTR %)		
Position	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)
2 to 1	1.7456	3,262	0.2296 (0.1899)	3.1752	1,580	0.8675 (0.0173)
3 to 2	1.16665	12,352	0.0497 (0.2300)	2.2306	1,670	0.1508 (0.4845)
4 to 3	0.9915	13,164	0.1034 (0.0178)	1.4834	1,084	0.1037 (0.6027)
5 to 4	0.8316	8,858	0.0486 (0.2579)	1.1118	338	−0.1014 (0.6239)
6 to 5	0.7368	4,298	0.1113 (0.0347)	0.7879	218	0.1999 (0.4463)
7 to 6	0.5807	1,700	0.1515 (0.0488)	0.7241	142	−0.3060 (0.6988)
8 to 7	0.3220	340	0.0231 (0.8641)	0.4415	64	−0.1394 (0.8101)

and 3; this mirrored the pooled results. Furthermore, the magnitudes of the effects both in absolute terms and as a percentage of the baseline are higher for category keywords than brand keywords. These findings are consistent with the notion that consumers searching for more specific products and brands have, on average, more information about what they are looking for than when they search using more general keyword phrases. Given that they have other information, they may use the position in the search advertising listings to a lesser extent than when their information is less specific. Once again, these results are consistent with Blake et al. (2015), even though that study is focused on overall effects of search advertising rather than position effects specifically. Their finding, like ours, is that the advertising effects are weaker for brand keywords than for nonbrand keywords.

6.5. Broad vs. Exact Match Types

We report the RD estimates for broad and exact match types for CTRs in Table 7. The comparisons of these two types of match types are consistent with our expectations. For broad match types, there are significant effects at positions 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. In general, position effects are more pervasive for broad match ads compared to exact match ads. This is consistent with our expectation that the headline and ad copy in exact match ads is more targeted compared to a broad match ad. To our knowledge, this is the first time an empirical study has documented the differences between advertising response for broad and exact match ads.

6.6. Weekends vs. Weekdays

The results for the position effects separated by weekday and weekend are reported in Table 8. The weekday results are largely similar to the pooled results, with a significant effect at positions 1, 3, and 5. The weekend effects are less significant in general but also show differences in the position effects. The only significant results 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 explanation for weekend effects in offline retail categories in Warner and Barsky (1995). The weekend effects described here also provide indirect support for the search cost explanation for position effects per se, while simultaneously not conclusively proving its existence or ruling out the presence of other explanations. If position effects are driven even partially by a sequential search mechanism, with consumers sequentially moving down the list of search advertising results until their expected benefit from the search is lower than their cost of further search, it is logical to conclude that they would search more when search costs are lower. Because search costs are plausibly lower on weekends, due to greater availability of time, this would lead to position effects lower down the list on weekends than on weekdays, which is what we find in our analysis.

Table 8 RD Estimates of Position Effects on CTRs: Weekday vs. Weekend

	Weekday (CTR %)			Weekend (CTR %)		
Position	Baseline	Num. of obs.	Estimate (*p*-value)	Baseline	Num. of obs.	Estimate (*p*-value)
2 to 1	2.4659	3,374	0.4994 (0.0145)	1.9813	1,356	0.4267 (0.1687)
3 to 2	1.2464	10,128	0.1018 (0.1356)	1.4592	3,998	−0.0190 (0.8770)
4 to 3	1.0104	10,168	0.1493 (0.0072)	1.1172	3,978	0.0314 (0.7008)
5 to 4	0.8240	6,844	0.0236 (0.6579)	0.9557	2,540	0.1811 (0.0279)
6 to 5	0.7637	3,384	0.1508 (0.0140)	0.7843	1,166	0.0721 (0.4919)
7 to 6	0.5989	1,300	0.1049 (0.2058)	1.0221	516	0.2005 (0.0485)
8 to 7	0.3302	328	−0.0340 (0.8199)	0.5391	142	0.0153 (0.4645)

6.7. Robustness

In comparing the position effects across firms, one possibility is that the different firms advertise on different keyword phrases, and hence the across-firm differences reflect differences across the keywords in reality. To rule out this explanation, we redid the analysis for a set of keywords that all three firms had advertised on. The downside to this analysis is that the number of observations is reduced. We are thus unable to report the estimates for positions 8 to 7 and 7 to 6 due to the paucity of observations in those positions. We find that the results for the top five positions (see Table 9) are consistent with the analysis presented earlier, with significant position effects at the same positions as before (except at position 5 for firm 3, which is significant at the 90% level instead of the earlier 95% level). We conduct similar robustness checks for the analysis of broad versus exact match ads (see Table 10), and for weekdays versus weekends (see Table 11), and find that the results are robust to keeping the set of keywords fixed across match type and weekday/weekend, respectively.

6.8. Practical Implications

Our approach to measuring the causal effects of position on clicks and sales is of practical relevance for both advertisers and the search engine itself. For advertisers, there are clear implications from our findings for the bidding strategies for search advertising. The advertiser from whom we obtained the data for our study sets its bids so as to maintain advertising expense as a percentage of sales at a target level. This target level itself was arrived at in consultations between the firm's marketing and finance teams. The advertiser estimated the value it placed on each click (which in turn was based on the expected profits generated through a click), and based its bids on this value. The advertiser then took the position it obtained from the auction as a given. This is not an unusual strategy for search advertising, given that a generalized second-price auction is used to sell search advertising spots. Advertisers may, therefore, have incentives to base their bids on their valuation for a click.[12] However, given our results, advertisers could benefit by targeting a position rather than bidding their valuations. For instance, we find that for the advertisers in our data set, there is no statistically significant effect of moving from position 3 to position 2. Thus, there is no benefit in terms of clicks or sales of being in position 2 relative to position 3. However, it is more costly to be in position 2 than in position 3. In our data, for instance, the average cost per click in position 2 is about 40.07 cents and at position 3, it is about 35.88 cents. Thus, there is almost an 11.7% difference in cost between these two positions. If the advertiser were to take the position effects into consideration while making their bidding decisions, there would be potentially large savings by moving ads to a lower position when there are no position effects (e.g., position 3 to position 2), with no effect on clicks or sales. In positions where there is a statistically significant positive effect of moving up a position (e.g., position 2 to position 1), the advertiser would need to trade off the increased clicks/sales from moving up a position with the increased costs. This "bidding for a position" strategy as opposed to one that bases the bids purely on the advertiser's valuation for clicks is possible only if the advertiser has credible estimates of the causal effect of position. Our estimates on the causal effects of position can thus be valuable to advertisers who want to optimize their bidding strategies.

Our approach is of value to the search engine. Search engines such as Google award positions based on a modified second-price auction. The search engine adjusts the bids with a *Quality Score* it assigns to each advertiser for each keyword. An important input to obtaining this score is a causal position effect. Search engines currently use controlled experiments to obtain the position effects, where the positions for the ads for a keyword phrase are randomized for a subset of the searches for the keyword phrase. This is costly for the search engine as these experimental pages do not generate revenues. In addition to the revenue impact of these experiments, they might provide relatively coarse estimates of the causal effects if

[12] Unlike a second-price auction, the generalized second-price auction mechanism used for search advertising does not have a truth telling equilibrium where advertisers bid their valuations. Yet as Edelman et al. (2007) show, the auction corresponds closely to a mechanism that generates similar outcomes as the second-price auction, in that the bids are based on the bidders' own valuations.

Table 9 **Firm-Specific Position Effects for Keywords Advertised by Firms 1–3**

	All firms (CTR %)			Firm 1 (CTR %)			Firm 2 (CTR %)			Firm 3 (CTR %)		
Position	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)
2 to 1	2.5248	1,210	0.4765 (0.0103)	1.9906	201	0.5317 (0.0484)	2.4748	258	0.4362 (0.0518)	3.5878	519	0.5292 (0.4704)
3 to 2	1.4315	5,874	0.1461 (0.8892)	1.3710	885	0.1674 (0.4487)	1.1795	2,332	0.0787 (0.6251)	2.1735	1,418	0.1308 (0.6738)
4 to 3	1.0207	7,406	0.1881 (0.0081)	1.1911	1,137	0.2906 (0.0187)	1.0973	3,292	0.0716 (0.0273)	0.8955	1,289	0.1752 (0.1820)
5 to 4	0.7977	5,084	0.1077 (0.2530)	0.8624	803	0.0978 (0.8722)	0.7407	2,396	0.1642 (0.5110)	0.6187	778	0.1015 (0.2802)
6 to 5	0.7639	2,404	0.2086 (0.0074)	0.7919	402	0.4438 (0.0016)	0.6881	1,155	0.1564 (0.0043)	0.6106	278	0.2241 (0.0663)

Table 10 **Position Effects for Keywords Advertised for Both Broad and Exact Match**

	Pooled estimates (CTR %)			Broad match (CTR %)			Exact match (CTR %)		
Position	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)
2 to 1	2.7278	2,342	0.4701 (0.0047)	1.7058	1,060	0.3427 (0.2133)	3.3658	1,256	0.8598 (0.0409)
3 to 2	1.6270	6,394	0.0546 (0.5992)	1.13388	4,830	0.0577 (0.6178)	2.2120	1,444	0.2060 (0.3767)
4 to 3	1.1770	4,348	0.0439 (0.0551)	0.9514	3,468	0.0907 (0.0311)	1.4221	918	0.0115 (0.6107)
5 to 4	0.8677	2,246	0.1732 (0.6055)	0.8371	1,886	0.1019 (0.7691)	0.9412	288	0.3739 (0.8665)
6 to 5	0.7342	892	0.1562 (0.0632)	0.7041	706	0.1363 (0.0890)	0.7956	190	0.1788 (0.5306)

the experiments are of a small scale (only a small proportion of pages served are experimental) and require pooling across keywords. Our RD approach allows search engines to obtain these causal position effects without conducting randomized experiments. Search engines can save their experimental capacity to conduct analysis that supplements the RD estimates or to inform other areas where alternatives to randomized experiments might not be feasible. The exact cost savings would be a function of the percentage of a search engine's queries that are used to conduct randomized experiments for obtaining position effects. If, for instance, the experiment is conducted for a random sample of 1% of queries, the opportunity costs for Google are about $100 million at the quarterly level (given revenues of $14.9 billion for Google in the quarter ending September 30, 2013, 68% of which came from search).

7. Conclusion

In this paper, we investigate the important issue of the causal effect of position in search advertising on outcomes such as website visits and sales. We present a novel RD-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 data set of advertising in a durable goods category by a focal advertiser as well as its major competitors on the Google search engine. The application of RD requires that the researcher observe the *AdRank* of competing retailers, which is the score used by Google to decide position. Typically, only Google observes the *AdRank* of competing advertisers. Hence, we would be unable to apply RD to measuring causal position effects using data from only one advertiser who observes only his own *AdRank*. However, because our data contains information including *AdRank* of firms that competed with this firm in Google's search advertising listings, we could set up an RD design.

We find that there are significant position effects, and that these would be significantly misstated by analysis that does not account for position selection and even more by sophisticated analysis that accounts for selection on observables. We find that the position effects are of economic significance, increasing

Table 11 Position Effects for Keywords Advertised Both on Weekdays and Weekends

	Pooled estimates (CTR %)			Weekday (CTR %)			Weekend (CTR %)		
Position	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)	Baseline	Num. of obs.	Estimate (p-value)
2 to 1	2.3034	4,504	0.4913 (0.0052)	2.4294	3,168	0.5197 (0.0142)	1.9813	1,356	0.4267 (0.1687)
3 to 2	1.2905	13,202	0.0520 (0.4082)	1.2404	9,262	0.0854 (0.2466)	1.4592	3,998	−0.0190 (0.8770)
4 to 3	1.0441	13,264	0.0943 (0.0453)	1.0119	9,388	0.1225 (0.0326)	1.1172	3,978	0.0314 (0.7008)
5 to 4	0.8706	8,764	0.0634 (0.2836)	0.8475	6,382	0.0150 (0.7925)	0.9557	2,540	0.1811 (0.0279)
6 to 5	0.7715	4,158	0.1121 (0.0424)	0.7608	3,030	0.1391 (0.0310)	0.7843	1,166	0.0721 (0.4919)
7 to 6	0.6437	1,644	0.1536 (0.0673)	0.5834	1,166	0.1441 (0.1289)	1.0221	516	0.2005 (0.0485)

the CTRs by 10%–20% in positions where they are significant, but by a lot less than the naive estimates that do not account for selection. We document moderators for these effects. Position effects, in general, are weaker for smaller firms, for keywords with a lower amount of prior consumer experience, and where the specificity of the search phrase is higher. We find important differences in these effects between broad and exact match keywords, with significant effects only at the topmost position for exact match, and significant effects only in lower positions for broad match keywords. Finally, we document important weekend effects in this context. Position effects are weaker on the weekend. This result is consistent with the idea that consumers' search costs are lower on weekends.

The results of our empirical analysis would be of great interest to managers who are setting firms' online advertising strategies. Furthermore, they should be of interest to the academic audience since we point to significant selection issues in this context and to a viable way to correct for them. The finding that value varies by position and that position effects vary by firm should be of interest for theoretical work in this area. 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 especially at the scale necessary for uncovering the impact of various moderators. We hope that this study leads researchers to pay greater attention to concerns about selection and endogeneity, which have received less than the necessary focus in the field of online advertising specifically and Web analytics in general.

Finally, we discuss the limitations of this paper. In general, RD designs can be demanding on the data. Fortunately, our empirical analysis is in a data-rich environment, allowing us to conduct a relatively rich analysis of main effects and moderators. However, in spite of the large data set, there are some constraints in what can be estimated. For instance, the sales orders results cannot be estimated for all of the positions for which CTR results can be estimated due to smaller volumes of orders relative to click-throughs. The advertiser-specific effects can be estimated reliably for only three of the four firms. The robustness checks, where we compare position effects across advertisers (or other such comparisons) with a common set of keywords cannot be reliably estimated for all positions. Because there is a variation in the amount of data available for the various estimates present, comparisons of estimates based on their significance levels needs to account for the amount of data available for the estimates. We have attempted to make transparent to the reader the volume of data used for every estimate that we present, so that the reader can draw their conclusions from the estimates with full information. Finally, the results we present here are based on the analysis of one large category of consumer durables. While there is some degree of generalizability of the results, particularly in the sense that the economic underpinnings of the effects and their moderators are similar, the reader should exercise caution in extrapolating the specific results to other categories. Future research using other data sets and methodologies could be helpful in investigating the effects that we have been unable to investigate in our paper.

Acknowledgments

The authors thank an anonymous data provider for sharing the data and for useful discussions. The authors also thank Wes Hartmann, Peter Lenk, Puneet Manchanda, and Ken Wilbur, and the participants of seminars at Columbia University, Google, Ohio State University, Santa Clara University, University of Rochester, University of Southern

California, University of California at Los Angeles, University of California at San Diego, University of Texas at Austin, Washington University in St. Louis, and the Adobe Data Sciences Symposium, the Bay Area Marketing Symposium, the UT Dallas FORMS Conference, and the 2012 Marketing Science Conference for useful comments. All remaining errors are the authors'.

References

Ackerberg DA (2001) Empirically distinguishing informative and prestige effects of advertising. *RAND J. Econom.* 32(2):316–333.

Agarwal A, Hosanagar K, Smith MD (2011) Location, location, location: An analysis of profitability of position in online advertising markets. *J. Marketing Res.* 48(6):1057–1073.

Athey S, Ellison G (2011) Position auctions with consumer search. *Quart. J. Econom.* 126(3):1213–1270.

Blake T, Nosko C, Tadelis S (2015) Consumer heterogeneity and paid search effectiveness: A large scale field experiment. *Econometrica* 83(1):155–174.

Busse M, Silva-Risso J, Zettelmeyer F (2006) $1,000 cash back: The pass-through of auto manufacturer promotions. *Amer. Econom. Rev.* 96(4):1253–1270.

Busse M, Simester D, Zettelmeyer F (2010) The best price you'll ever get: The 2005 employee discount pricing promotions in the U.S. automobile industry. *Marketing Sci.* 29(2):268–290.

Chen Y, He C (2011) Paid placement: Advertising and search on the Internet. *Econom. J.* 121(556):F309–F328.

Cook TD, Campbell DT (1979) *Quasi-Experimentation* (Houghton Mifflin Harcourt, Boston).

Deighton J, Henderson C, Neslin S (1994) The effects of advertising on brand switching and repeat purchasing. *J. Marketing Res.* 31(1):28–43.

Ebbes P, Wedel M, Bockenholt U (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. *Quant. Marketing Econom.* 3(4):365–392.

Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* 97(1):242–259.

Fan J, Gijbels I (1996) *Local Polynomial Modeling and Its Applications* (Chapman & Hall, London).

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

Guan Z, Cutrell E (2007) An eye tracking study of the effect of target rank on Web search. *Proc. SIGCHI Conf. Human Factors Comput. Systems* (ACM, New York), 417–420.

Hahn J, Todd P, van der Klaauw W (2001) Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica* 69(1):201–209.

Hoch SJ, Ha Y-W (1986) Consumer learning: Advertising and the ambiguity of product experience. *J. Consumer Res.* 13(2): 221–233.

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

Imbens GW, Lemieux T (2008) Regression discontinuity designs: A guide to practice. *J. Econometrics* 142(2):615–635.

Jerath K, Ma L, Park Y-H, Srinivasan K (2011) A "position paradox" in sponsored search auctions. *Marketing Sci.* 30(4):612–627.

Kalyanam K, Borle S, Boatwright P (2010) The impact of ad position in search engine marketing: The effect of brand name and price. Working paper, Santa Clara University, Santa Clara, CA.

Katona Z, Sarvary M (2010) The race for sponsored links: Bidding patterns for search advertising. *Marketing Sci.* 29(2):199–215.

Kihlstrom RE, Riordan MH (1984) Advertising as a signal. *J. Political Econom.* 92(3):427–450.

Lamberti J (2003) Measuring the effectiveness of sponsored search marketing. Discussion paper, Interactive Advertising Bureau, New York.

Lee DS, Lemeiux T (2010) Regression discontinuity designs in economics. *J. Econom. Literature* 48(2):281–355.

Lodish L, Abraham M, Kalmensen S, Livelsberger J, Lbetkin B, Richardson B, Stevens M (1995) How TV advertising works: A meta-analysis of 389 real-world split cable TV advertising experiments. *J. Marketing Res.* 32(3):125–139.

Ludwig J, Miller D (2007) Does head start improve children's life chances? Evidence from a regression discontinuity design. *Quart. J. Econom.* 122(1):159–208.

Nair HS, Hartmann WR, Narayanan S (2011) Identifying causal marketing mix effects using a regression discontinuity design. *Marketing Sci.* 30(6):1079–1097.

Narayanan S, Kalyanam K (2014) Regression discontinuity with estimated score. Working paper, Stanford University, Stanford, CA.

Narayanan S, Manchanda P (2009) Heterogeneous learning and the targeting of marketing communication for new products. *Marketing Sci.* 28(3):424–441.

Nelson P (1974) Advertising as information. *J. Political Econom.* 82(4):729–754.

Parker PM, Gatignon H (1996) Order of entry, trial diffusion and elasticity dynamics: An empirical case. *Marketing Lett.* 7(1): 95–109.

Rutz OJ, Bucklin RE (2011) From generic to branded: A model of spillover in paid search advertising. *J. Marketing Res.* 48(1): 87–102.

Rutz OJ, Trusov M (2011) Zooming in on paid search ads—A consumer-level model calibrated on aggregated data. *Marketing Sci.* 30(5):789–800.

Sethuraman R, Tellis GJ (1991) An analysis of the trade-off between advertising and price discounting. *J. Marketing Res.* 28(3): 160–174.

Shadish WR, Cook TD, Campbell DT (2001) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Cengage Learning, Boston).

Sheth JN, Mittal B, Newman BI (1999) *Customer Behavior: Consumer Behavior and Beyond* (Dryden Press, Fort Worth, TX.).

Thistlethwaite DL, Campbell DT (1960) Regression-discontinuity analysis: An alternative to the ex post facto experiment. *J. Educational Psych.* 51(6):209–317.

van der Klaauw W (2008) Regression discontinuity analysis: A survey of recent developments in economics. *Labour: Rev. Labour Econom. Indust. Relations* 22(2):219–245.

Varian HR (2007) Position auctions. *Internat. J. Indust. Organ.* 25(6):1163–1178.

Varian HR (2009) Online ad auctions. *Amer. Econom. Rev.* 99(2): 430–434.

Warner EJ, Barsky RB (1995) The timing and magnitude of retail store markdowns: Evidence from weekends and holidays. *Quart. J. Econom.* 110(2):321–352.

Weitzman ML (1979) Optimal search for the best alternative. *Econometrica* 47(3):641–654.

West PM, Brown CL, Hoch SL (1996) Consumption vocabulary and preference formation. *J. Consumer Res.* 23(2):120–135.

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