NET Institute*

www.NETinst.org

Working Paper #07-23

Revised April 2008

Search Engine Advertising: Pricing Ads to Context

Avi Goldfarb
University of Toronto

Catherine Tucker
MIT

Search Engine Advertising:
Pricing ads to context

Avi Goldfarb and Catherine E. Tucker[*]

April 2008

Abstract

Each search term put into a search engine produces a separate set of results. Correspondingly, each of the sets of ads displayed alongside the results is priced using a separate auction. We investigate how bids for these context-based ads depends on the difficulty of making a match. This contrasts with the existing literature that focuses on the effect of match quality. We examine advertising prices paid by lawyers for 139 Google search terms in 195 locations. Other things being equal, the fewer searches there are on a term, the higher the price. To identify a causal relationship between match-difficulty and prices paid, we exploit a natural experiment in "ambulance-chaser" regulations across states. When lawyers cannot contact a client by mail and matching becomes more difficult, the relative price per ad click is $0.93 higher. We check the robustness of this result by performing a falsification test using a different ambulance-chaser regulation. Our results suggest that prices are higher for context-based ads when the difficulty of both online and off-line matching increases. This highlights that a major reason why search advertising is profitable is because its use of context can monetize the "long tail" by reducing friction in the matching process.

1. Introduction

In 1998, Goto.com[1] introduced two novel features to search advertising markets that had not been tried before in other advertising markets: (1) The use of electronic auctions and (2) the ability to provide and price advertisements based on search terms or "keywords." The practical implications of these new features were as follows. A personal injury lawyer and an immigration lawyer decide to advertise their services. Using normal media channels, like newspapers, magazines, Yellow Pages or a banner ad on a website, they would pay the same price, given the advertisement's physical size, placement and the audience size and demographics. However, if they used Google, which uses a similar system to Goto.com's, it would be a different story. The personal injury lawyer, after placing a bid online, would pay on average $26.18 every time someone clicked on her ad alongside a search for "personal injury lawyer". However, the immigration lawyer using Google would have to pay on average only $7.48 per click for the same sized ad, displayed alongside a search for "immigration lawyer."

So far, the academic literature on search-engine ad pricing has focused on the question of how search auctions reflect the "match-value" of such ads, in terms of how much a firm values the match with the consumer. By contrast, in this paper we investigate how search auction prices reflect the "match-difficulty" of such ads; that is, how difficult it is for firms to reach customers through online and off-line methods. This allows us to explore the extent to which search engine advertising reaps the benefits of monetizing the "long tail" of advertising, where hard-to-match obscure clients and obscure products find each other.

We first establish that there is a negative relationship between search volume and prices. Although our finding this relationship does suggest that when a match is difficult (i.e. few

---

[1]Goto.com was renamed Overture in 2001 and purchased by Yahoo in 2003. Prior to this time, search engine ads were priced by impressions and demographics.

customers are searching for that service), search ad prices will be higher, it is not conclusive because search volume may be endogenous.

To address this identification challenge, we look for exogenous variation in how difficult it is to make a match. We find such variation in state bar regulations that prohibit "ambulance-chasing behavior" for the $40 billion sector of trial lawyers. Many states have laws that prevent lawyers contacting potential clients using written media for a few months after the accident. This makes it harder for vendors to match with clients. We use data from market research conducted by a lawyer website portal. We have data on estimated auction prices for with 139 different searches for various legal service "keywords" in 195 regional city markets. We regress a keyword's estimated price per click on fixed effects for each location and keyword, and an indicator variable for whether the keyword is affected by state regulations.

We find that in locations with solicitation regulations, injury keywords cost advertisers an extra $1.01 (roughly 7%) relative to the price of other keywords (such as "tax lawyer") in that state, compared with the price premium of personal injury keywords in non-regulated states. We perform a battery of tests to check the robustness of this result. These include a falsification test for the endogeneity of the state law, using an alternate law that is similarly motivated. Overall, our results suggests that when advertisers cannot reach customers through alternative advertising platforms and matching is harder, ad prices rise. The search engine enables these difficult matches cheaply, so it can profit from the "long tail" of advertising. The relatively frictionless nature of context-based ad pricing means that the search engine, the advertisers, and the customers gain from advertiser-client matches that are otherwise problematic or costly using alternative on-line or off-line methods.

2. Related Literature

We examine how much context-based ad pricing in media platforms depends on the difficulty of making a match. We build on four distinct literatures: (1) two-sided advertising markets, (2) online advertising, (3) online/offline substitution, and (4) legal services marketing.

Previous research on two-sided media platforms like search engines has modeled the use of content to attract customers and consequently advertisers (Anderson and Gabszewicz 2005; Xie and Chen 2007). However, this literature promotes models where advertisers pay more the more eyeballs they reach, rather than where payments vary with match-difficulty.[2] The empirical literature has echoed this focus. Wilbur (2007) shows that TV ad prices increase with audience size. Busse and Rysman (2005) show that yellow pages ad prices increase with ad size and, by implication, exposure. By contrast, this paper emphasizes that prices can rise when advertisers expect less exposure, because there are few potential clients on the other side of the market.

Our empirical focus on "match-difficulty" illuminates a nascent theoretical literature on the potential benefits of "targeting" advertising. Work by Iyer, Soberman, and Villas-Boas (2005) demonstrates the theoretical advantages that such targeting has for firms, while Gal-Or and Gal-Or (2005) use the example of customized television advertising to show that better targeting of advertisements increases customer welfare. Chen and He (2006) have extended this targeting literature to paid search.

Our emphasis on horizontal differentiation in match-*difficulty* across keywords contrasts with most of the literature, which focuses on how prices in position auctions reflect "match-quality" or expected match profitability. A major contribution of this literature has been to use

---

[2] The consequences of such inflexible pricing policies are set out by Baye and Morgan (2001). They show that when a media platform sets a fixed advertising fee, this fee will exceed the socially optimal level and many potential advertisers will opt out.

heterogeneity in match-quality to explain why firms pay more *per click* to be displayed first (Edelman, Ostrovsky, and Schwarz (2007)). Other theoretical and empirical articles generalize the second-price auction to take account of match-quality (e.g. Ganchev, et al. 2007), and therefore take "match-difficulty" as exogenous. Associated research, such as (Wilbur and Zhu 2007)'s work on click fraud, examines how match-quality affects search advertising decisions.

The empirical literature on search advertising in marketing has also focused on the effects of match-quality for search advertising. For example, Rutz and Bucklin (2007) and Ghose and Yang (2007) have shown the effects of different keywords on performance and cross-selling opportunities. The rest of the empirical literature on online advertising has focused on banner ads and email marketing (such as Manchanda, et al. 2006; Chatterjee, Hoffman and Novak 2003; and Ansari and Mela 2003), perhaps because these predate keyword advertising.[3]

We also add to a growing literature on the relationship between offline and online options and marketing outcomes (Jank and Kannan (2006); Forman, Ghose, and Goldfarb (2007)). We add to this literature, that studies decisions by consumers to go online, by showing that firms' decisions to go online to make a match are dependent on how hard it is offline to make a match.

Finally, our work is related to an older but substantial marketing literature on legal services advertising (Smith and Meyer 1980, Kotler and Connor 1977, Darden, Darden and Kiser 1981). This literature was inspired by the deregulation of legal services advertising in 1977 and examined the consequences from both firm and consumer perspectives of the introduction of advertising and marketing by lawyers. We follow in this tradition by studying the relationship between regulation and lawyer advertising prices.

---

[3] The first banner advertisement (for Zima alcoholic beverage) appeared on Wired Magazine's Hotwired website in 1994. While OpenText briefly experimented with something like search advertising in 1996, it was not successfully implemented until Goto.com applied it in 1998. Prior to the establishment of auctions as the way to price keywords, Yahoo charged a fixed rate for banner advertisements placed near popular keywords.

3. Data on Advertising Prices for Lawyer Services

We use marketing research data collected by a platform that brings together clients and lawyers. These data were collected from Google's "Traffic Estimator Tool", which provides potential advertisers with a guide to the auction prices they would expect to pay for different keywords in different locations. The traffic estimator provides (given enough data points) a range of prices that other advertisers have paid recently for an ad being in positions 1-3 in a certain city and the search volume associated with that price range. Our data contain projections for 139 keywords for 195 geographic areas defined by Google to closely resemble (consolidated) metropolitan statistical areas. In order to use our natural experiment of state-level restrictions, we exclude metropolitan statistical areas that cross state lines, like Burlington, VT – Plattsburg, NY and New Bedford, MA – Providence, RI. Appendix Table 1 provides descriptive statistics for the data used in the study. Appendix Table 2 provides a complete list of the keywords used.

There are two major challenges to using this data: Interpreting price data from an auction mechanism, and missing data. We discuss each in turn. In using data from the Traffic Estimator Tool, we use the exact information advertisers use in setting their bid prices. Since 2002, Google and Yahoo have sold keywords using second-price sealed bid auctions instead of less stable first-price auctions (Edelman, Ostrovsky and Schwarz 2007). However, the form of second-price auction used obscures how bids translate to prices. An advertiser places a bid based on its maximum willingness to pay for an ad to appear next to a specific search term for a specific geographical location. Google then bills a sum lower than this maximum price whenever the ad is clicked. However, an advertiser is not necessarily paying the second price that was bid in that particular auction. Instead, keyword prices post-bidding are adjusted for the quality of the website buying the keyword, click-fraud, and the clicks–to-impression ratio, with no information

given to advertisers (or researchers) about the precise formulas used. In this paper, we use "estimated prices" data for Google that abstract from this ex-post quality adjustment.[4]

The average price per click for the different keywords and different types of keywords varies greatly. Table 1 presents some initial results showing the relationship between search volume and keyword price. In these regressions, we used fixed effects to control for inherent differences in propensity to pay across keywords and across locations. For example, the keyword fixed effects control for the fact that ad prices for searches for food poisoning attorneys are less than 3 percent of ad prices for searches for aviation attorneys, as the expected payout of an aviation accident lawsuit is higher. The dummy for location captures city-specific factors such as comfort with the internet, which may explain why areas such as Greenwood, MI have click prices that are on average one-third of those in Newark, NJ. Our initial results in column 1 show a strong negative relationship between the average price per click and the reported search volume. To check that this was not an artifact of the linear specification, we also used log values of the price per click. These showed that our results hold for both percentage changes and levels.

Table 1: Variation of CPC with Search Volume

Dependent Variable	Price Per Click	Logged Price Per Click	Price per Click Missing CPC data coded as zero (Tobit)
Search Volume	-1.229***	-0.0975***	-8.909***
	(0.155)	(0.0168)	(0.188)
Observations	12271	12264	26964
R-squared	0.809	0.890	0.274[5]

Fixed effects for each region and each keyword included

---

[4] Google accounts for two-thirds of the search market (October, 2007).

[5] Pseudo R-Squared

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We face a missing data challenge, in that Google reports the cost per click range when they have enough historical data.[6] We ran further regressions to evaluate whether the missing data were systematically connected to the type of keyword or to the presence of the solicitation regulations we use later in the paper for identification. We found no statistically significant evidence that they were. This lack of systematic correlation and the work of Little (1992) suggests that missing data are not driving our results. In addition, we ran Tobit specifications including the missing data that allowed for censoring of keyword price at the bottom of the observed range. The results, reported in column 3 of Table 1, have the same sign but are of larger magnitude than those reported in column 1. This suggests that if anything the missing data bias our estimates downwards.

Another challenge of using these price data is that Google gives a price range, but not an indication of the distribution of prices paid between these lower and upper cutoffs. We report results for the midpoint of this range. We have repeated all our specifications using both the upper and lower limits, and obtained qualitatively similar results.

The results in Table 1 show that there is a negative correlation between how many clients are searching (our proxy for the ease of an off-line match) and the price of the keyword. Of course there is a difficulty in putting a causal interpretation on this negative correlation, despite the inclusion of fixed effects for search phrase and location. There may be unobserved factors for that location that both reduce search volume and increase the profitability of a client lead, for select keywords. For example, suppose this could be a location where mountainous roads meant

---

[6] A similar data sparseness issue was addressed by Rutz (2007), who uses Bayesian methods to help estimate search word performance for hotel search advertising data. Unfortunately for our purposes, Google has not embraced this methodology.

that trucks and consequently truck accidents were rare, but that when they did happen they were very serious. To tackle this inherent identification challenge, we use a natural experiment.

4. Variation in Restrictions on Lawyer Behavior

Trial lawyers earned $40 billion in 2004, an amount that is over 50 percent higher than Microsoft or Intel and twice that of Coca-Cola (National Review 2004). The size of this market makes studying advertising strategies in this industry independently important. However, for our purposes of identifying how the difficulty of off-line matches affects search advertising prices there are two other attractive features of this industry: Differences in state-level bar exams and the small-scale nature of personal-injury lawyer practices keep markets local[7] and there is variation in rules regarding off-line solicitation across states. We use this variation in ambulance-chasing solicitation regulations to establish whether context-based ad pricing is more profitable when off-line matching is more difficult. Each regulation gives us a natural experiment with a treatment group of locations affected by the regulation and a control group of locations that are not affected. To control for systematic differences between regulated and unregulated states, we contrast keyword prices affected by regulation with keyword prices that are unaffected by the state regulations in regulated states. Therefore, we estimate how much affected keywords diverge in price from unaffected keywords in regulated locations relative to unregulated locations.

Law firms have only been allowed to advertise nationwide since 1977, when the Supreme Court ruled to allow legal advertising in <u>Bates v. the State Bar of Arizona</u>. This case brought to

---

[7] Though several states have reciprocity agreements with lawyers in other states, the small scale nature of most personal injury claims means that cases are tried locally by local lawyers.

an end a state bar association tradition that it was not seemly for lawyers to advertise their services in newspapers, on television, or through other channels. This deregulation prompted a spate of empirical evaluation by marketing scholars (Smith and Meyer 1980, Kotler and Connor 1977, Darden, Darden and Kiser 1981) on legal service advertising. However, the notion that there are some types of marketing communications that demean the status of the law persists in local state bar regulations. In particular, some state bars prohibit lawyers from writing to potential clients who have recently sustained an accident or injury.

A typical text in a state bar manual is found in a section entitled "solicitation", and reads:

"A lawyer shall not send, or knowingly permit to be sent, on a lawyer's behalf or on behalf of the lawyer's firm or on behalf of a partner, an associate, or any other lawyer affiliated with the lawyer or the lawyer's firm, a written communication to a prospective client for the purpose of obtaining professional employment if the written communication concerns an action for personal injury or wrongful death arising out of, or otherwise related to, an accident or disaster involving the person to whom the communication is addressed or a relative of that person, unless the accident or disaster giving rise to the cause of action occurred more than *X* days before the mailing of the communication"

Table 2 records all regulations as of April 2007 where a state bar forbade written communication to potential clients. There is a little variation over how long the states prohibited contact (the mode is 30 days), but the regulations are similar. There was no statistically significant relationship between the enacting of a law and the number of lawyers per dollar of gross state product, the number of civil suits per head, state population, or state GDP (Appendix Table 3).

Table 2: Bar regulations prohibiting contact with clients

State	Personal injury laws/rules
Alabama	No written communication allowed 30 days for personal injury or wrongful death
Arizona	No written communication allowed 30 days for personal injury or wrongful death
Arkansas	No written communication allowed 30 days for wrongful death
Colorado	No written communication allowed 30 days for personal injury or death
Connecticut	No written communication allowed 40 days for personal injury or death
Florida	No written communication allowed 30 days for personal injury or wrongful death
Georgia	No written communication allowed 30 days for personal injury or wrongful death
Hawaii	No written communication allowed 30 days for personal injury or wrongful death

Louisiana	No written communication allowed 30 days for personal injury or wrongful death
Missouri	No written communication allowed 30 days for personal injury or wrongful death (accident or disaster)
Nevada	Must wait 45 days after any known event before written communication
New York	No written communication for 30 days for personal injury or wrongful death unless law says need to file in 30 days in which case cannot solicit for 15 days
South Carolina	No written communication allowed 30 days for personal injury or wrongful death
Tennessee	No written communication allowed 30 days for workers' comp, personal injury, or wrongful death
Wyoming	For written communications, need to wait 30 days after "occurrence" before soliciting a specific client

Personal injury keywords can be objectively identified because bar associations uses a precise legal definition to define what is a personal injury case and what is not. *Personal injury* is damage to an individual rather than property, and is taken to cover accidents, medical negligence, and industrial diseases contracted by workers at their workplace. The personal injury keywords we identified cover regular accidents as well as industrial diseases such as mesothelioma where regulations apply after diagnosis or death.[8] There are, however, a few cases where there may be both personal injury and injury to property in a civil suit. For example "toxic mold attorneys" may litigate for both personal injury damages and property damages. We tried including and excluding these "combined" civil cases, and achieved qualitatively similar results.

5. Estimation Strategy and Results

Using data on the prices of keywords across cities, we examine the responsiveness of keyword prices to this variation in how easy it is to make a match "off-line". Descriptive statistics of keyword prices across regulatory regimes suggest that the regulations have an effect: keyword prices are 28 cents higher in states with solicitation regulation. These differences may, however, be a result of unobservable differences in willingness to pay across keywords and locations. To control for these unobservable differences, we include a series of fixed effects (i.e.

---

[8] The keywords, and whether they were categorized as personal injury keywords, are listed in Appendix Table 2.

dummy variables) for each location l and each keyword k and focus on the interaction between whether a keyword relates to personal injury and whether there is personal injury regulation in that state. The location fixed effects allow us to control for all city-level differences in numbers of lawyers, wealth, and litigiousness. The keyword fixed effects allow us to control for all keyword-level differences. Therefore, this empirical strategy allows us to control for *differences* in prices that occur because personal injury keywords are different from other keywords, and also *differences* in prices that occur because states that enact personal injury regulation are different from states that do not; this is known as a "differences in differences" approach.[9] Usually in differences-in-differences researchers take the approach of using a prior time periods not affected by the policy to control for geographical cross-sectional variation in customer behavior. By contrast, in this paper in place of a time series we use *other* keywords to control for this cross-sectional variation in consumer behavior. As long as there is no other systematic reason why personal injury keywords should be differently priced to non-personal injury keywords in states with regulation, we can interpret the interactions β as measuring the causal effect of the regulations on prices.

$Cost\ per\ Click_{kl}$

$$= \beta (PersonalInjuryWord_k) X (Solicitation\ Restricted_l)$$

$$+ \text{Keyword}_k + City_l + \varepsilon_{kl}$$

[1]

We estimate equation [1] using a variety of distributional and specification assumptions. Table 3 displays results for our main specification. The estimates for the interactions suggest that

---

[9] This use of differences in differences is a similar idea to the specifications used by Chevalier and Mayzlin (2006) in their study of online book reviews and Busse, Risso and Zettelmeyer (2006) in their study of pass-through of auto manufacturer promotions.

both solicitation regulations and contingency fee limits affect the prices that lawyers pay for personal injury search terms.

Table 3: Main Results

	(1)	(2)	(3)
Personal injury keyword *and* Law	1.013**	1.419**	0.866*
restricting solicitation	(0.496)	(0.573)	(0.447)
Observations	12271	5067	12271
R-squared	0.808	0.762	0.808
Sample	All	Accident Words and Non-Specific Words	Broader Definition of Accident Word

All regressions include a full set of fixed effects for each city and each keyword.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The presence of a solicitation regulation is associated with a $1.01 increase in the price for a personal injury keyword. These values are economically important relative to average keyword prices of $9.28. The significance of these estimates is robust to various specifications of the error structure. These results suggest that when state bar regulation makes it harder to contact personal injury victims by other marketing communications channels, lawyers are willing to pay relatively more for personal injury search advertising keywords.

We conducted a number of robustness checks on our results. For the independent variables, we wanted to verify that it was not an idiosyncratic definition of "personal injury keyword" that led to our results. To allow for a broader definition of personal injury, we also tried a definition including "any violation of an individual's right, other than his or her rights in

property." This added the keywords associated with "dog bites", "mold", "toxic mold", "premises liability", "food poisoning" and "nursing home abuse" to the treatment group. The results reported in column 3 of Table 3 are similar to before, if slightly less precise. We also ran a more limited regression that used only the non-specialty keywords in Appendix Table 2 as controls. The results reported in column 2 of Table 3 are very similar to the main results.

6. A Falsification Check for the Endogeneity of the State Law

The interpretation of our results relies on the assumption that the enacting of bar regulation over personal injury lawsuits is exogenous. One concern is, however, that these laws might reflect particular market conditions for lawyers. For example, solicitation regulation might be more likely in areas where lawyers are more aggressive at seeking clients. However, these aggressive lawyers could also be more likely to win cases and consequently value a match more highly, driving up bid prices and confounding our results.

As an initial check, we studied the correlation between the enactment of a law and state characteristics. We found no statistically significant relationship between solicitation restrictions and the number of lawyers per dollar gross state product, the number of civil suits per capita, the state's population, or gross state product per capita (Appendix Table 3).

To further verify that this is not driving our results, we study the effects of an alternate, similarly motivated law as a placebo/falsification test. We use the example of "contingency fee limits." These are also enacted by states in response to "aggressive behavior" on the part of lawyers. If the endogeneity of such "ambulance-chaser" regulations explains our results then we would expect such laws to also be associated with a negative effect on keyword price. If, on the other hand we are measuring the effect of the law on ad prices, we would expect such regulation

to be associated with a positive effect on keyword price. This gives us a way of testing whether the endogeneity of the law is driving our results.

A contingency fee is a fee payable only in the case of a favorable result. Table 4 displays the contingency fee limits across states based on data from the Pacific Research Institute's "U.S. Tort Liability Index". While there is substantial variation in the laws' texts, all the laws ultimately limit how profitable it is to represent a personal injury client.

Table 4: Contingency Fee Limits

State	Law
Alaska	Requires that contingent fees be calculated exclusive of punitive damages. [Alaska Stat. § 9.60.080.]
Illinois	Limits contingent fees to 33.3% of the first $150,000 recovered, 25% of the next $850,000 recovered, and 20% of any amount recovered over $1 million. [735 Ill. Comp. Stat Ann. § 5/2 –1114.]
Maine	Limits contingent fees in professional liability cases to 33.3% of the first $100,000 recovered, 25% of the next $100,000 recovered, and 20% of any amount recovered over $2 million. Permits a judge to allow fees in excess of these amounts in special circumstances. [Me. Rev. Stat. Ann. tit. 24 § 2961.]
Nebraska	Allows a court to review contingent fees in medical and professional liability cases. [Neb Stat. § 44-2834.]
Oklahoma	Limits contingent fees to 50% of a plaintiff's recovery. [Okla. Stat. Ann. tit.5, § 7.]
Wisconsin	Limits contingent fees to 1/3 of the first $1 million recovered, 25% of the first $1 million recovered if liability is stipulated within 180 days of filing of the original complaint and not within 60 days of first day of trial, and 20% for amounts exceeding $1 million recovered. Allows a judge to exceed these amounts in exceptional circumstances. [Wisc. Stat. Ann. § 655.013.]

Source: Pacific Research Institute

Table 4 reports the results for an identical specification to Table 3

Table 5: Falsification check using contingency fee limits

	(1)	(2)	(3)

Personal injury keyword *and* Law limiting Contingency Fees	-2.169** (0.837)	-2.450** (0.976)	-1.962** (0.806)
Observations	12271	5067	12271
R-squared	0.808	0.762	0.808
Sample	All	Accident Words and Non-Specific Words	Broader Definition of Accident Word

All regressions include a full set of fixed effects for each city and each keyword.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

State contingency fee limits are associated with a $2.27 decrease in the upper range price of personal injury keywords and a $1.86 decrease for the lower range of prices. The robustness checks using just limited numbers of keywords and a broader definition of an accident word again support the results. In particular, the limited number of keyword results rules out the possibility that rules such as Alaska's that encompass all punitive damages are leading to us mis-specify our control group of words. This result is again robust to various specifications of the error structure. This suggests that when there are no state contingency fee limits to reduce the profitability of lawsuits, context-based ad pricing allows search engines to charge higher prices. Thus, these context-based ad prices are extremely sensitive to the profitability of the end customer.

7. Further Investigation

We use the effect of state solicitation regulation as a proxy for match difficulty. We further stratified our results by other indicators of the difficulty of the matching process. Table 6 shows

the results. We only find a positive and significant effect for solicitation regulation when the market is "thin." When search volume is estimated as low (less than one search a day), there is a negative effect from the solicitation regulation. When the market is thicker (more than one search a day), the difficulty of matching off-line produces a negligible effect on prices. Similarly, when we break up our results by "lawyer spending," which is an alternate measure of the litigiousness of that state or the thickness of the legal market, we find similar results. The difficulty of making a match off-line affects prices only when spending on lawyers is low, or in our interpretation the population is less litigious making the market thinner.

Table 6: Further exploration

	(1)	(2)	(3)	(4)
Personal injury keyword *and* Law	0.313	0.971**	0.239	1.693***
restricting solicitation	(0.459)	(0.485)	(0.795)	(0.631)
Observations	1753	10518	6148	6123
R-squared	0.971	0.802	0.801	0.816
Sample	Search Volume >=1	Search volume<1	Above median lawyer spending/GSP	Below median lawyer spending/GSP[+]

All regressions include a full set of fixed effects for each city and each keyword.

+ Resident and active attorneys in 2005 per dollar of gross state product. Source: Pacific Research Institute

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

8. Conclusion

A new and growing literature asks how search engine pricing reflects match quality. We take a different approach, asking how search engine auction pricing reflects match-*difficulty*. We present initial evidence that suggests that search volume is negatively correlated with prices. This

suggests a correlation between difficulties in finding clients and higher search ad prices. However, correlation is not necessarily causative. To identify a causal relationship between the difficulty of finding clients and search ad-prices, we sought an exogenous shifter of match difficulty. We found such a shifter in the form of ambulance-chasing regulations that in several states prohibit lawyers from contacting potential clients using written media. When lawyers are not allowed to contact a personal injury or wrongful death client by mail, the price of a personal injury keyword is $1.01 higher in that state relative to other personal injury keywords controlling for location fixed effects. We perform multiple robustness checks for this result, including a placebo test to check for the endogeneity of state laws using contingency fee limits. Our findings suggest that search engines can monetize the difficulties that vendors have finding clients in thin markets. A search engine's ability to use context-based ads to automate without friction the match of obscure clients and vendors allows them to profit from the "long tail" of advertising.

There are both managerial and policy implications to this research. Managerially, our results suggest that context-based pricing is an effective marketing strategy for extracting rents from advertisers because it enables frictionless matching and lower search costs for vendors seeking clients. It is therefore not clear that extending electronic auctions to other advertising networks without context-based advertising in place will necessarily reach the "long tail" of advertising. For example, it is not clear that Google's plans to bring online auctions to TV ads and conduct these auctions on the basis of "daypart, geography and […] demographic" will be as successful as at promoting frictionless matching as its prior online search auctions that are conducted using context-based pricing.

Our findings also have anti-trust implications when it comes to defining markets for search engines. They suggest that the existence of off-line markets (in our case direct response

advertising) that facilitate matching can reduce a search engine's pricing power. This means that anti-trust authorities should look broadly when thinking about the market definition for search advertising.

Bibliography

Anderson, Simon P, and Jean Jaskold Gabszewicz. "The Media and Advertising: A Tale of Two-Sided Markets"." *CEPR Discussion Papers 5223*, 2005.

Anderson, Eric, Nathan Fong, Duncan Simester and Catherine Tucker "How does an Obligation to Collect Sales Tax Affect Consumer and Firm Behavior?" Mimeo, MIT (2007)

Ansari, Asim, and Carl Mela. "E-Customization." *Journal of Marketing Research* 40, no. 2 (2003): 131-145.

Bagwell, Kyle. "The Economic Analysis of Advertising." In *Handbook of Industrial Organization*. 2003.

Baye, Michael R., and John Morgan. "Information Gatekeepers on the Internet and the Competitiveness of Homogenous Product Markets." *American Economic Review* 91, no. 3 (2001): 454-474.

Busse, Meghan, and Marc Rysman. "Competition and Price Discrimination in Yellow Pages Advertising." *RAND Journal of Economics* 36, no. 2 (2005): 378-390.

Busse, Meghan, Jorge Silva-Risso Risso, and Florian Zettelmeyer. "$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions." *American Economic Review* 96, no. 4 (2006): 1253-1270.

Chatterjee, Patrali, Donna L. Hoffman, and Thomas P. Novak. "Modeling the Clickstream: Implications for Web-Based Advertising Efforts." *Marketing Science* 22, no. 4 (2003): 520-541.

Chen, Yongmin, and Chuan He. *Paid Placement: Advertising and Search on the Internet.* Net Institute Working paper #06 03, 2006.

Chen, Yuxin, and Jinhong Xie. "Cross-Market Network Effect with Asymmetric Customer Loyalty: Implications for Competitive Advantage." *Marketing Science* 26 (2007: 52-66

Chevalier, Judith, and Dina Mayzlin. "The Effect of Word of Mouth on Sales: Online Book Reviews." *Journal of Marketing Research* 18, no. 3 (2006).

Darden, Donna K., William R. Darden, and G. E. Kiser. "The Marketing of Legal Services." *Journal of Marketing* 45, no. 2 (1981): 123-134.

Economist. "Who's Afraid of Google?" September 1 (2007): 9.

Edelman, Ben, Michael Ostrovsky, and Michael Schwarz. "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review* 97, no. 1 (2007): 242-259.

Forman, Chris, Anindya Ghose, and Avi Goldfarb. "Competition between local and electronic markets: How the benefit of buying online depends on where you live." Working paper, University of Toronto (2007).

Gal-Or, Esther, and Mordechai Gal-Or. "Customized Advertising via a Common Media Distributor." *Marketing Science* 24, no. 2 (2005): 241-253.

Ganchev, Kuzman, Alex Kulesza, Jinsong Tan, Ryan Gabbard, Qian Liu, and Michael Kearns. *Empirical Price Modeling for Sponsored Search.* Presented at the Workshop on Sponsored Search Auctions, 2007.

Ghose, Anindya, and Sha Yang. *An Empirical Analysis of Search Engine Advertising: Sponsored Search and Cross-Selling in Electronic Markets.* Working Paper, New York University, 2007.

Grossman, Gene. M, and Carl Shapiro. "Informative Advertising with Differentiated Products." *Review of Economic Studies* 51, no. 1 (1984): 63-81.

Iyer, Ganesh, David Soberman, and Miguel Villas-Boas. "The Targeting of Advertising." *Marketing Science* 24, no. 3 (2005): 461-476.

Jank, Wolfgang, and P.K. Kannan. "Understanding Geographical Markets of Online Firms using Spatial Models of Customer Choice." *Marketing Science* 24 (2005): 623-634.

Kotler, Philip, and Richard A. Connor. "Marketing Professional Services." *Journal of Marketing* 41, no. 1 (1977): 71-76.

Little, Roderick J. A. "Regression With Missing X's: A Review." *Journal of the American Statistical Association* 87, no. 420 (1992): 1227-1237.

Manchanda, Puneet, Jean-Pierre Dubé, Khim Yong Goh, and Pradeep K. Chintagunta. "The Effect of Banner Advertising on Internet Purchasing." *Journal of Marketing Research* 43, no. 1 (2006): 98-108.

National Review. July 08, 2004.

Reuters. "EBay to auction advertising on U.S. radio stations." 2007.

Rutz, Oliver, and Randolph Bucklin. *A Model of Individual Keyword Performance in Paid Search Advertising.* Working Paper, Yale University, 2007.

Smith, Robert E., and Tiffany S. Meyer. "Attorney Advertising: A Consumer Perspective." *Journal of Marketing* 44, no. 2 (1980): 56-64.

Wilbur, Kenneth C. "An Empirical Model of Television Advertising and Viewing Markets." *Marketing Science*, 2007: Forthcoming.

Wilbur, Kenneth, and Yi Zhu. *Click Fraud.* Working Paper, University of Southern California, 2007.

Appendix

Appendix Table 1

Variable	# of Observations	Mean	Standard Deviation	Minimum	Maximum
CPC (mid-point)	12271	9.28	7.650	0	52.87
Personal Injury Keyword	26964	0.187	0.389	0	1
Law restricting solicitation	26964	0.304	0.460	0	1
Law restricting contingency fees	26964	0.103	0.303	0	1

Appendix Table 2: List of Keywords

Personal Injury Keywords	Other Keywords		

Personal Injury Keywords

Asbestos Attorney
Asbestos Lawyer
Aviation Accident Attorney
Aviation Accident Lawyer
Birth Injury Attorney
Birth Injury Lawyer
Brain Injury Attorney
Brain Injury Lawyer
Car Accident Attorney
Car Accident Lawyer
Construction Accident Attorney
Construction Accident Lawyer
Dog Bite Attorney
Dog Bite Lawyer
Food Poisoning Attorney
Food Poisoning Lawyer
Medical Malpractice Attorney
Medical Malpractice Lawyer
Mesothelioma Attorney
Mesothelioma Lawyer
Personal Injury Attorney
Personal Injury Lawyer
Truck Accident Attorney
Truck Accident Lawyer
Wrongful Death Attorney
Wrongful Death Lawyer

Neutral Keywords

Attorney
Attorneys
Law Firm
Lawsuit
Lawyer
Lawyers
Legal Aid
Legal Help
Litigation
Mediation
Mediator

Other Keywords

Adoption Attorney
Adoption Lawyer
Alimony Attorney
Alimony Lawyer
Arson Attorney
Arson Lawyer
Assault Attorney
Assault Lawyer
Bankruptcy Attorney
Bankruptcy Lawyer
Child Abuse Attorney
Child Abuse Lawyer
Child Support Attorney
Child Support Lawyer
Computer Crime Attorney
Computer Crime Lawyer
Contract Attorney
Contract Lawyer
Credit Card Fraud Attorney
Credit Card Fraud Lawyer
Custody Attorney
Custody Lawyer
Divorce Attorney
Divorce Lawyer
Domestic Violence Attorney
Domestic Violence Lawyer
Drug Possession Attorney
Drug Possession Lawyer
Dui Attorney
Dui Lawyer
Dwi Attorney
Dwi Lawyer
Embezzlement Attorney
Embezzlement Lawyer
Employment Attorney
Employment Lawyer
Estate Planning Attorney
Estate Planning Lawyer
Extortion Attorney
Extortion Lawyer
Family Law Attorney
Family Law Lawyer
Forgery Attorney
Forgery Lawyer
Identity Theft Attorney
Identity Theft Lawyer
Immigration Attorney
Immigration Lawyer

Insurance Fraud Attorney
Insurance Fraud Lawyer
Intellectual Property Attorney
Intellectual Property Lawyer
Landlord Attorney
Landlord Lawyer
Living Will Attorney
Living Will Lawyer
Mold Attorney
Mold Lawyer
Money Laundering Attorney
Money Laundering Lawyer
Nursing Home Abuse Attorney
Nursing Home Abuse Lawyer
Oui Attorney
Oui Lawyer
Patent Attorney
Patent Lawyer
Perjury Attorney
Perjury Lawyer
Premises Liability Attorney
Premises Liability Lawyer
Prenuptial Attorney
Prenuptial Lawyer
Probate Attorney
Probate Lawyer
Prostitution Attorney
Prostitution Lawyer
Real Estate Attorney
Real Estate Lawyer
Robbery Attorney
Robbery Lawyer
Securities Fraud Attorney
Securities Fraud Lawyer
Sexual Assault Attorney
Sexual Assault Lawyer

Shoplifting Attorney
Shoplifting Lawyer
Tax Attorney
Tax Lawyer
Tenant Attorney
Tenant Lawyer
Theft Attorney
Theft Lawyer
Toxic Mold Attorney
Toxic Mold Lawyer
Traffic Violation Attorney
Traffic Violation Lawyer
Visa Attorney
Visa Lawyer
Workers Compensation Attorney
Workers Compensation Lawyer
Wrongful Termination Attorney
Wrongful Termination Lawyer

Appendix Table 3: Correlation Coefficients for State Laws and State Characteristics

	Gross state product per capita	State population	Average CPC upper bound	Average CPC lower bound	Presence of solicitation regulation	Presence of contingency fee limit	Resident and active attorneys per dollar of state GSP	Total state trial-courts' incoming civil cases per 1000 population
GSP per capita	1.00							
State population	-0.05	1.00						
	(0.70)							
Average CPC upper bound	-0.12	-0.11	1.00					
	(0.42)	(0.46)						
Average CPC lower bound	0.34***	0.30**	0.16	1.00				
	(0.01)	(0.03)	(0.27)					
Solicitation regulation	-0.06	0.03	0.07	0.15	1.00			
	(0.65)	(0.81)	(0.63)	(0.29)				
Contingency fee limit	0.34***	-0.13	-0.16	-0.05	-0.26*	1.00		
	(0.02)	(0.37)	(0.25)	(0.74)	(0.07)			
Resident and active attorneys	0.03	0.27*	-0.17	-0.09	0.03	0.07	1.00	
	(0.84)	(0.06)	(0.23)	(0.53)	(0.84)	(0.63)		
Civil cases	0.25*	0.03	-0.09	0.10	0.00	-0.16	0.01	1.00
	(0.09)	(0.84)	(0.56)	(0.49)	(0.99)	(0.27)	(0.95)	