Click Fraud

Kenneth C. Wilbur
Assistant Professor of Marketing
Marshall School of Business
University of Southern California
3660 Trousdale Parkway, ACC 306E
Los Angeles, California 90089
Phone: 213-740-7775
Fax: 213-740-7828
Email: kwilbur@usc.edu

Yi Zhu
Doctoral Student
Department of Economics
University of Southern California
3620 South Vermont Avenue, KAP 300
Los Angeles, California 90089
Phone: 213-740-2106
Fax: 213-740-8543
Email: zhuy@usc.edu

Acknowledgements: We are grateful for extensive comments and discussions with Simon Anderson, Martin Byford, Joshua Gans, Allen Weiss, Linli Xu, and seminar participants at Stanford GSB, the UC/USC Marketing Colloquium, Santa Clara University, Melbourne Business School, the Second Workshop on Game Theory in Marketing, and the 2007 INFORMS Marketing Science Conference.

Click Fraud

Abstract

"Click fraud" is the practice of deceptively clicking on search ads with the intention of either increasing third-party website revenues or exhausting an advertiser's budget. Search advertisers are forced to trust that search engines do everything possible to detect and prevent click fraud even though the engines get paid for every undetected fraudulent click. We seek to answer whether it is in a search engine's interest to prevent click fraud.

We find that, under full information in a second price auction, if x% of clicks are fraudulent, advertisers will lower their bids by x%, leaving the auction outcome and search engine revenues unchanged. However, if we allow for uncertainty in the amount of click fraud or change the auction type to include a click-through component, search engine revenues may rise or fall with click fraud. A decrease occurs when the keyword auction is relatively competitive, as advertisers lower their budgets to hedge against downside risk. If the keyword auction is less competitive, click fraud may transfer surplus from the winning advertiser to the search engine. This last result suggests that the search advertising industry may benefit from using a neutral third party to audit search engines' click fraud detection algorithms.

Keywords: *Advertising, Auctions, Click Fraud, Game Theory, Internet Marketing, Search Advertising*

Search advertising revenues grew from virtually nothing in 1996 to more than $7 billion in 2006, constituting 43% of online advertising revenues (Advertising Age 2006). The primary benefits of search advertising for advertisers are its relevance and accountability. It tends to reach consumers as they enter the market for the advertised product, and advertisers' ability to track consumers' actions online allows for accurate measurements of advertising profitability.

The downside of this accountability is a practice known as "click fraud." Website publishers or rival advertisers may impersonate consumers and click search ads, driving up advertising costs without increasing sales, effectively stealing a firm's paid advertising inventory. The Click Fraud Network, which defines itself as "a community of online advertisers, agencies and search providers," estimated that 16.2% of all search engine paid clicks, and 28.1% of all content network paid clicks, in the third quarter of 2007 may have been fraudulent. Discussions with executives in the search advertising industry indicate that the amount of click fraud varies widely across industries and keywords. The perceived threats of click fraud may outweigh the benefits of using search advertising for some firms in high-risk categories.

73% of search advertisers say that click fraud is a concern (Advertising Age 2006). The question of click fraud is vexing because search engines cannot give advertisers full information about how they detect and prevent click fraud. Doing so would be tantamount to providing unscrupulous advertisers with directions on how to commit click fraud. Advertisers are therefore forced to trust that search engines do their utmost to prevent click fraud, even though the search engines get paid every time they fail to detect a fraudulent click. This trust was called into question in 2006 when Google CEO Eric Schmidt was quoted saying "Eventually the price that the advertiser is willing to pay for the conversion will decline because the advertiser will realize that these are bad clicks. In other words, the value of the ad declines. So, over some amount of time, the system is, in fact, self-correcting. In fact, there is a perfect economic solution, which is to let it happen." (Ghosemajumder 2006) His remarks were interpreted as suggesting that market forces would eliminate any negative effects of click fraud in the long run, possibly undermining the need for click fraud detection.[1]

The primary objective of this paper is to understand how click fraud affects search engines' advertising revenues. We also hope to gain insights into what actions search engines may be able to take to mitigate click fraud. We present an analytical model of the auction market for search advertising keywords and then introduce the possibility that third-party websites or rival bidders may engage in click fraud. The strengths of our model are its parsimony and generality as firms' search advertising objectives and the degree of competition in keyword

[1] In 2005, Google CFO George Reyes said "I think something has to be done about [click fraud] really, really quickly, because I think, potentially, it threatens our business model." (Stone 2005)

auctions vary widely across keywords.

We find that, in a second-price auction, when firms know that $x\%$ of all clicks will be fraudulent, they lower their bids by $x\%$. In equilibrium, this adjustment leaves advertising expenditures and the auction result unchanged. However, when the amount of click fraud is uncertain or when the auction contains a click-through rate component, search engine revenues may increase or decrease with click fraud. A decrease may occur in relatively competitive keyword auctions as high bidders hedge their advertising budgets to protect against the threat of a high realization of click fraud. On the other hand, advertising revenues may increase in relatively uncompetitive auctions if the foregone profits of exiting the ad auction outweigh the effects of click fraud, resulting in a transfer from very profitable advertisers to the search engine.

The surge in internet usage and advertising revenues has attracted substantial academic interest (see, e.g., He and Chen 2006, Iyer and Pazgal 2003, Manchanda et al. 2006, Prasad 2007). Research on search advertising has focused mainly on competition in advertising auctions and consumer search. Baye and Morgan (2001) analyzed a homogeneous products market organized by a search engine ("gatekeeper") and showed that the gatekeeper's incentive is to maximize consumer adoption but limit the number of advertisers using the platform since it can extract more revenues when competition among advertisers is lessened. Chen and He (2006) analyzed optimal consumer search and advertiser bid strategies and showed that advertisers' bid order mirrors their products' relevance order. Consumers then optimally engage in sequential search. Borgers et al. (2007), Edelman, Ostrovsky, and Schwarz (2007) and Varian (2007) analyze equilibria in sophisticated auction mechanisms similar to those used by search engines.

Empirical work on search advertising has focused mainly on the link between keyword prices and advertiser profitability. Goldfarb and Tucker (2007) showed that keyword prices increase in advertisers' profitability of advertising, and decrease with the availability of substitute advertising media. Rutz and Bucklin (2007a) developed a model to enable advertisers to decide which keywords to keep in a campaign, and showed that keyword characteristics and ad position influence conversion rates. Rutz and Bucklin (2007b) showed that there are spillovers between search advertising on branded and generic keywords, as some customers may start with a generic search to gather information, but later use a branded search to complete their transaction. Ghose and Yang (2007) empirically analyzed a model of consumer search and advertiser behavior, linking keyword characteristics to purchase rates and evaluating the optimality of advertiser bids.

We are not aware of any previous analyses of the economic effects of click fraud. We begin by discussing the institutional details of the industry that guide our analysis.

1 Industry Background

In this section, we describe the market for search advertising, types of click fraud, advertiser perceptions of click fraud, and issues in click fraud detection and measurement.

1.1 The Search Advertising Marketplace

Search advertising, also known as "cost-per-click" (CPC) or "pay-per-click" advertising, is sold on a per-click basis. Advertisers bid on a word or phrase related to their business and enter a maximum advertising budget per time period. When consumers enter that "keyword" into a search engine or read a third-party webpage relevant to the keyword, the advertiser's ad then may be displayed along with the consumer's search results or webpage content. If the consumer clicks on the advertiser's ad, she is redirected to a web address chosen by the advertiser, and the advertiser is charged a fee. Advertising costs and quantity of searches available vary widely across keywords.

Search advertising was pioneered by a firm named GoTo.com, which was later renamed Overture and acquired by Yahoo. Overture sold keywords in a public-information, first-price auction. It later changed its auction mechanism to a private-value variation on Vickrey's (1961) second-price auction, the Generalized Second Price auction described by Edelman, Ostrovsky, and Schwarz (2007). The market leaders are Google, Yahoo, and Microsoft with 64%, 22%, and 6% of clicks, respectively.[2]

Keyword prices vary according to advertiser profitability, media competition, and keyword characteristics. Though not representative, the keyword "mesothelioma attorney" cost an average of $35 per click, but region-specific keyword costs reached as high as $80 per click (Goldfarb and Tucker 2007). Rutz and Bucklin (2007b) illustrate the dramatic differences between keywords containing branded and generic terms. In a search advertising campaign for a hotel chain, branded keywords on Google created 3.5 million impressions, with a click-through rate of 13.3% and a cost per reservation of $2.76. Generic keywords generated 19.9 million impressions, with a click-through rate of 0.3% and a cost per reservation of $61.71.

Search ads are typically ranked according to some function of advertisers' willingness to pay and the ads' value to searching consumers. Google's early ranking algorithm was to multiply the advertiser's bid per click by its "click through rate," the number of consumers who clicked on the ad divided by all consumers who saw the ad. This tended to increase the utility of search ads, increasing customer traffic and acceptance of advertising. There is some evidence that higher ad positions are more desirable since not all consumers read through all of the ads. For example, Wilk (2007) reported that 62% of all searchers do not

[2] Source: http://hitwise.com/datacenter/searchengineanalysis.php, accessed November 2007.

read past the first page of ads, and 23% do not read past the first few ads. He also noted that consumers often refine their search if they do not find a good ad among the first few slots. Chen and He (2006) find that a higher ad listing sends a quality signal to uninformed consumers. Rutz and Bucklin (2007a) and Ghose and Yang (2007) demonstrate empirically that higher ad positions result in higher conversion rates.

In 2006, Google added a "quality score" to its ranking function. The quality score is a function of click-through rate, search term relevance, ad text, and ad landing page, but the specific function is not publicly available. Yahoo added a click-through component to its ranking algorithm in 2007 (Shields 2007).

Other forms of online advertising include cost-per-thousand (CPM), in which websites are compensated on an impression basis, and cost-per-action (CPA), in which advertisers pay per sale or lead. Prasad (2007) discussed "impression fraud," a problem in CPM advertising that is conceptually similar to click fraud but operationally different. CPA advertising has the potential to resolve click fraud concerns, but has a principal/agent problem in which advertisers are incented to conceal customer leads and conversions from the search engine. Google piloted a CPA beta test in 2007 but participating advertisers were required to use Google software to track their conversions. It may be that if advertisers reveal enough revenue information to the search engine to resolve the principal/agent problem, the search engine would be able to design its auction mechanism to extract maximal advertising revenues. We speculate that CPA will cannibalize some CPC revenues, but we do not expect it to completely replace the CPC business model.

1.2 Types of Click Fraud

Search advertisers are charged when their ads are clicked, regardless of who does the clicking. Clicks may come from potential customers, employees of rival firms, or computer programs. We refer to all clicks that do not come from potential customers as "click fraud."

Click fraud is sometimes called "invalid clicks" or "unwanted clicks." This is partly because the word "fraud" has legal implications that may be difficult to prove or contrary to the interests of some of the parties involved. Google calls click fraud "invalid clicks" and says it is "clicks generated through prohibited methods. These prohibited methods include but are not limited to: repeated manual clicks, or the use of robots, automated clicking tools, or other deceptive software."[3] Google acknowledges and describes the risks posed by click fraud in its annual reports.

There are two main types of click fraud:

[3] Source: https://www.google.com/adsense/support/bin/answer.py?answer=16737. Accessed January 2007.

- **Inflationary click fraud:** Search advertisements often appear on third-party websites and compensate those website owners on a per-click basis or with a share of advertising revenues. These third parties may click the ads to inflate their revenues.

- **Competitive click fraud:** Advertisers may click rivals' ads with the purpose of driving up their costs or exhausting their ad budgets. When an advertiser's budget is exhausted, it exits the ad auction. A common explanation for competitive click fraud is that firms have the goal of driving up rivals' advertising costs, but such an explanation may not be subgame perfect. If committing competitive click fraud is costly, then driving up competitors' costs comes at the expense of driving down one's own profits. A more convincing explanation may be found in the structure of the ad auction. When a higher-bidding advertiser exits the ad auction, its rival may claim a better ad position without paying a higher price per click.

There are myriad other types of click fraud, such as fraud designed to boost click-through rates, to invite retaliation by search engines against rival websites, or to do malicious harm based on philosophical or economic grounds. These other types are thought to be relatively infrequent, so we do not consider them in this paper.

1.3 Advertiser Perceptions of Click Fraud

Search advertisers say click fraud is troubling. Advertising Age (2006) reported the following results of a survey of search advertising agencies:

> "In your experience, how much of a problem is click fraud with regard to paid placement?"
> 16% "a significant problem we have tracked"
> 23% "a moderate problem we have tracked"
> 35% "we have not tracked, but are worried"
> 25% "not a significant concern"
> 2% "never heard of it"
>
> "Have you been a victim of click fraud?"
> 42% Yes
> 21% No
> 38% Don't know

"What type of click fraud did you experience?"

78% Inflationary click fraud

53% Competitive click fraud

Google implicitly acknowledged the problem when it paid $90 million to settle a click-fraud lawsuit, *Lane's Gifts v. Google*, in July 2006.

1.4 Click Fraud Detection and Prevention

Search engines implicitly acknowledge they cannot fully detect click fraud. Google states: "[our] proprietary technology analyzes clicks and impressions to determine whether they fit a pattern of use intended to artificially drive up an advertiser's clicks or impressions, or a publisher's earnings. Our system uses sophisticated filters to distinguish between clicks generated through normal use by users and clicks generated by unethical users and automated robots, enabling us to filter out most invalid clicks and impressions."[4] Thus, they imply that they do not detect fraudulent clicks that do not fit a pattern. We surmise it is especially difficult to detect invalid clicks if they come from IP addresses that are used by many people or if the invalid clicks are designed to resemble clicks generated by normal human use.

Most search engines claim to offer advertisers some basic protections against click fraud, though they do not explain specifically how they identify fraudulent clicks. Tuzhilin (2006) defined the "fundamental problem of click fraud prevention:" a search engine can not explain specifically how it detects click fraud to its advertisers without providing explicit instructions to unscrupulous advertisers on how to avoid detection. Advertisers are forced to either blindly trust that search engines seek to prevent click fraud or they may hire third party firms to detect click fraud and pursue refunds for any such fraud detected.

Empirical research on click fraud's effects will have two challenges. The first challenge is that probabilistic judgments are required to detect click fraud, as a smart click fraudster would design its fraudulent clicks to complicate detection. For example, fraudulent clicks may be generated by a widely-distributed "botnet" (Daswani and Stoppelman 2007) and designed to mimic human use. The second challenge is that if click fraud can be detected by the researcher, it also could have been detected by the advertiser. Search engines' standard business practice is to refund advertising expenditures when advertisers present evidence of undetected click fraud, so empirical evidence of click fraud's effects could potentially be impacted by advertiser detection of click fraud. We speculate these concerns could be resolved by analyzing data from a company that previously did not try to detect click fraud,

[4] Source: https://www.google.com/adsense/support/bin/answer.py?answer=9718&ctx=sibling. Accessed January 2007.

or perhaps using an experimental approach.

2 A Baseline Model of Search Advertising

We begin with a simple setting to establish how the market operates in the absence of click fraud. This aids interpretation of equilibrium results when we introduce inflationary and competitive click fraud in later sections.

Clicks We assume there is a fixed period of length one. n customers click and clicks arrive at a constant rate $\frac{1}{n}$. Firm $j \in \{1, 2\}$ receives π_{jW} per customer click when its ad is in the top spot, and π_{jL} otherwise. We define $\Delta_j = n(\pi_{jW} - \pi_{jL})$ as the total value to advertiser j of remaining in the top spot for the entire period of time. It must be that $\min(\Delta_1, \Delta_2) > 0$ else firms will never enter positive bids.

Search Advertising Technology Each firm enters a bid per click, b_j, for a single advertising slot sold by a monopoly gatekeeper. The high bidder claims the slot and pays the low bidder's bid per click.[5] The high bidder then enters a capacity, $K_j \geq 1$, the maximum number of clicks for which it is willing to pay. If the total number of clicks exceeds K_j, the high bidder exits the advertising market when its capacity has been exceeded. The low bidder then claims the top spot at the next-highest advertiser's per-click bid, which we normalize to 0.

Given that the advertiser's maximum expenditure is the product of K and b, the advertiser's capacity choice is equivalent to choosing an advertising budget, as is required by all major search engines (Google, Yahoo, MSN). In the absence of an ad budget, we presume an advertiser could choose to stop remitting payments to the search engine at some point, which would give the same effect. The capacity-setting assumption simplifies the analysis, but the results would be unchanged under a budget-setting assumption.

Edelman, Ostrovsky, and Schwarz (2007) showed that, with two bidders and one slot, the Generalized Second Price auction used by Google and Yahoo reduces to a Vickrey auction. We will appeal several times to the standard result that, in a second price auction, it is optimal for firms to bid their reservation price.

We assume the two firms' ads have identical click-through rates, so the high bidder wins the top ad position. We relax this assumption in section 5.

[5] Results in sections 2-4 continue to hold if the auction winner is determined by the advertising budget, or by the product of bid and total clicks, rather than by a comparison of per-click bids alone.

Structure of the Game The game is played in two stages. First, each firm enters a bid per click and observes its position, with the high bidder in the top spot. In the second stage, the high bidder chooses its capacity. The reason for this structure is that, in reality, each firm may discover its rival's bid immediately by varying its own bid and observing whether its ad position changes, but a rival's capacity may only be discovered if it has been exhausted. We seek a subgame perfect equilibrium in pure strategies under full information.

Equilibrium We analyze second-stage profits, then first stage actions. When firm j wins the auction (W), its profit is

$$\Pi_{jW} = \begin{cases} n\left(\pi_{jW} - b_k\right) & \text{when } K_j \geq n \\ n\pi_{jW}\frac{K_j}{n} + \left(1 - \frac{K_j}{n}\right) n\pi_{jL} - b_k K_j & \text{when } K_j < n \end{cases} \tag{1}$$

where $\frac{K_j}{n}$ is the fraction of customer clicks firm j receives while on top, in the event it does not remain on top for the entire time period. If firm k wins the auction and sets a capacity $K_k \geq n$, firm j receives $\Pi_{jL} = n\pi_{jL}$

In the first stage, each firm will anticipate the second-stage outcome and choose the bid b that makes it indifferent between winning and losing the auction. Thus b_j is chosen to equate $\Pi_{jW} = \Pi_{jL}$.

We summarize equilibrium behavior in proposition 1.

Proposition 1 *In the absence of click fraud, firm j bids $\frac{\Delta_j}{n}$ per click and wins the advertising auction if $\Delta_j > \Delta_k$. The auction winner remains on top for the entire time period and earns a profit of $n\pi_{jW} - \Delta_k$. The gatekeeper earns* $\min\{\Delta_1, \Delta_2\}$.

Proof. Π_{jW} does not change with K_j when $K_j \geq n$. When $K_j < n$, Π_{jW} changes linearly with K_j at rate $(\pi_{jW} - \pi_{jL} - b_k)$. If firm j wins, it cannot be that $b_k > \pi_{jW} - \pi_{jL}$, since firm j would be better off losing in this scenario. Thus, when firm j wins the auction, $\frac{\partial \Pi_{jW}}{\partial K_j} > 0$ for $K_j \leq n$, so firm j sets $K_j \geq n$ and earns $\Pi_{jW} = n(\pi_{jW} - b_k)$. $\Pi_{jW} = \Pi_{jL}$ then gives $b_j = \frac{\Delta_j}{n}$. ■

Proposition 1 serves as a useful benchmark to which we compare equilibria under inflationary and competitive click fraud.

3 Search Advertising with Inflationary Click Fraud

We now introduce inflationary click fraud into the baseline model. Major search engines pay third-party websites to display search ads relevant to their site content. Inflationary click fraud results when those website owners click the ads to inflate their advertising revenues.

We first analyze the individual website's problem of choosing a click fraud level. Next, we solve for equilibrium bids and capacities, given a known amount of inflationary click fraud. Finally, we consider the case of stochastic inflationary click fraud.

3.1 Websites' Choice of Inflationary Click Fraud

Ads associated with a particular keyword are placed on I third-party websites, indexed by w. We consider two different compensation schemes. If website w is paid γ_w per click generated, its revenues are $\gamma_w(n_w+r_w)$, where n_w is the number of customer clicks generated through the website, and r_w is its inflationary click fraud level. If website w is paid a fraction $\delta_w \in (0,1)$ of the ad revenues it generates, its revenues are $\delta_w b(n_w + r_w)$, where b is the advertiser's payment per click. b does not vary across websites and is a function of inflationary click fraud, with $\frac{\partial b}{\partial r_w} < 0$ (as shown in section 3.2).

We assume the cost of r_w fraudulent clicks is an increasing and convex function $c_w(r_w)$ since a greater number of fraudulent clicks increases the risk that the search engine will detect the fraudulent activity. The search engine could then retaliate by excluding the website from its content network or initiating a costly legal action against the website if click fraud constitutes a breach of contract.

Under a per-click compensation scheme, website w's profits are

$$\pi_w = \max_{r_w} \gamma_w(n_w + r_w) - c_w(r_w) \tag{2}$$

yielding a first-order condition $\gamma_w = c'_w(r^*_w)$ and a choice of $r^*_w = c'^{-1}_w(\gamma_w)$ in equilibrium.

Under a revenue-sharing compensation scheme, website w's profits are

$$\pi_w = \max_{r_w} \delta_w b(n_w + r_w) - c_w(r_w) \tag{3}$$

and site w's first-order condition, $\delta_w b + \delta_w(n_w + r_w)\frac{\partial b}{\partial r_w} = c'_w(r_w)$, yields a unique r^*_w.

Proposition 2 *Holding click fraud constant, for $\gamma_w = \delta_w b$, per-click and revenue-sharing compensation schemes yield identical payouts to websites. Allowing for endogenous click fraud, the revenue-sharing compensation scheme will reduce inflationary click fraud, as it incents content network partners to partially internalize the effect of inflationary click fraud on advertisers' bids.*

Proof. The right-hand sides of site w's first-order conditions are identical under the two compensation schemes, but the left-hand side is strictly lower under the revenue-sharing compensation scheme, since $\frac{\partial b}{\partial r_w} < 0$. ■

The larger the derivative $\frac{\partial b}{\partial r_w}$ is in absolute value, the less inflationary click fraud the third-party websites will produce. It is typically thought that search ads are displayed on a large number of websites in search engines' content networks, suggesting that $\frac{\partial b}{\partial r_w}$ is small. The implications of this result for search engines are clear: click fraud is reduced when search ads are not rotated across a large number of websites, and each website is compensated with a percentage of the advertising revenues it generates.

If $\frac{\partial^2 b}{\partial r_w^2} > 0$ (which we later show to be the case under certain assumptions), then $\frac{\partial b}{\partial r_w}$ is greatest in absolute value when $I = 1$. Thus, websites' incentives to internalize the effects of their inflationary click fraud on search engine revenues are maximized when other sites' click fraud levels do not affect their share of advertising revenues. This suggests that search engines should allow advertisers to enter site-specific keyword bids b_w to maximally reduce sites' incentives to engage in click fraud. While it may be difficult for a human to manage site-specific bids when I is large, software could be designed to accomplish this task.

While a revenue-sharing compensation scheme suggests an equilibrium relationship between bids and inflationary click fraud levels, we are not going to model this relationship explicitly. Current search engine policies make information transmission between websites and advertisers prohibitively difficult. Search engines do not reveal the distribution of γ_w or δ_w to either side. In addition, websites do not know b, and advertisers do not know $c_w(r_w)$, nor how they vary across websites or advertisers. (But note that websites need not know $\frac{\partial b}{\partial r_w}$ exactly for Proposition 2 to hold; they need only know $\frac{\partial b}{\partial r_w} < 0$.) Finally, advertisers have only limited information about where their ads will appear, and website owners do not know in advance what ads will appear on their sites.

We proceed under each of two assumptions: either both advertisers can anticipate the inflationary click fraud level $r = \sum_w r_w$, or they share a common belief $f(r)$ about its distribution. Advertisers may share a common belief about $f(r)$ because they both observe the low bidder's bid, which is the only bid that is paid, so it is the only bid that influences r.

Next, we analyze the equilibrium effects of inflationary click fraud on advertisers' bidding strategies.

3.2 Deterministic Inflationary Click Fraud

We assume that customers generate $n = n_0 + \sum_w n_w$ clicks, where n_0 is the number of customer clicks that come directly from the search engine. Website owners generate r fraudulent clicks. Both advertisers can anticipate r. We relax this assumption in section 3.3.

Equilibrium Equation (4) gives firm j's profit when it wins the auction.

$$\Pi_{jW} = \begin{cases} n\pi_{jW} - b_k (n+r) & \text{when } K_j \geq n+r \\ n\pi_{jW}\frac{K_j}{n+r} + \left(1 - \frac{K_j}{n+r}\right) n\pi_{jL} - b_k K_j & \text{when } K_j < n+r \end{cases} \tag{4}$$

The first segment of the profit function represents the outcome in which firm j stays on top for the entire time period. The second segment occurs if firm j's capacity will be exhausted at some point, in which case it is on top for $n\frac{K_j}{n+r}$ customer clicks, and has exited the auction for the remaining $\left(1 - \frac{K_j}{n+r}\right) n$ customer clicks.

Proposition 3 summarizes equilibrium behavior.

Proposition 3 *When inflationary click fraud is deterministic and known to both bidders, and there is no competitive click fraud, firm j bids $\frac{\Delta_j}{n+r}$ and wins the advertising auction if $\Delta_j > \Delta_k$. Advertisers reduce their bids by a proportion of $\frac{r}{n+r}$, pricing out the effect of click fraud. Firm j remains on top for the entire time period and earns a profit of $n\pi_{jW} - \Delta_k$. The gatekeeper's revenues are* $\min\{\Delta_1, \Delta_2\}$, *as in the baseline model.*

Proof. Π_{jW} does not change with K_j when $K_j \geq n+r$. When $K_j < n+r$, Π_{jW} changes linearly with K_j at rate $\left(\frac{\pi_{jW}-\pi_{jL}}{n+r} - b_k\right)$. If firm j wins, it cannot be that $b_k > \frac{\pi_{jW}-\pi_{jL}}{n+r}$, since firm j would be better off losing in this scenario. Thus, when firm j wins the auction, $\frac{\partial \Pi_{jW}}{\partial K_j} > 0$ for $K_j \leq n+r$, so firm j sets $K_j \geq n+r$ and earns $\Pi_{jW} = n\pi_{jW} - b_k(n+r)$. $\Pi_{jW} = \Pi_{jL}$ then gives $b_j = \frac{\Delta_j}{n+r}$. ■

The auction mechanism completely internalizes the effect of inflationary click fraud when the number of fraudulent clicks is known to both bidders. Advertiser profits are unaffected; bids adjust endogenously to counter the detrimental effects of the fraudulent clicks. The gatekeeper's revenues are unchanged, though its profits may fall if it makes larger transfers to third-party websites.

3.3 Stochastic Inflationary Click Fraud

It is perhaps more intuitive to assume that advertisers do not know how many fraudulent clicks will occur since they may not know the distribution of δ_w, γ_w, n_w, or $c_w(r_w)$ across websites or where their ads will appear. We assume here that advertisers maximize expected profits under a common belief about the probability density $f(r)$ of the inflationary click fraud level r.

Capacity Choice We now add uncertainty about r into firm 1's profit function. If $K_1 < n$, firm 1's capacity will be exhausted for any realization of r. When $K_1 \geq n$, firm 1's capacity is only exhausted for some realizations of r. Equation 5 displays firm 1's profit function when it wins the auction.

$$\Pi_{1W} = \begin{cases} \int_0^\infty \left[n\pi_{1W}\frac{K_1}{n+r} - b_2 K_1 + \left(1-\frac{K_1}{n+r}\right)n\pi_{1L}\right] f(r)\,dr & \text{when } K_1 < n \\ \int_0^{K_1-n}\left[n\pi_{1W} - b_2(n+r)\right] f(r)\,dr + \\ \quad + \int_{K_1-n}^\infty \left[n\pi_{1W}\frac{K_1}{n+r} + \left(1-\frac{K_1}{n+r}\right)n\pi_{1L} - b_2 K_1\right] f(r)\,dr & \text{when } K_1 \geq n \end{cases} \tag{5}$$

The uncertainty in the first segment of the profit function concerns the number of clicks for which the firm will remain on top. On the second segment of the profit function, the first term is the firm's expected profits when it remains on top, weighted by the probability that r is small enough that firm 1 is never knocked off. The second term is the firm's expected profits in the event its capacity is exhausted, weighted by the probability that r is large enough to exhaust the firm's capacity.

Figure 1 depicts Π_{1W}. For $K_1 < n$, Π_{1W} changes linearly with K_1 at a constant rate $\int_0^\infty \frac{\Delta_1 f(r)dr}{n+r} - b_2$.

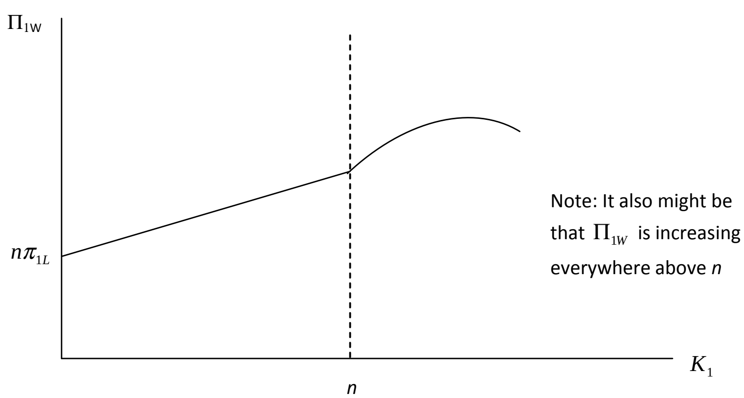


Figure 1: Firm 1's Profit Function

For $K_1 \geq n$,

$$\begin{aligned} \frac{\partial \Pi_{1W}}{\partial K_1} &\equiv MR(K_1) - MC(K_1) \\ &= \int_{K_1-n}^\infty \frac{\Delta_1}{n+r} f(r)\,dr - b_2[1-F(K_1-n)]. \end{aligned} \tag{6}$$

Both $MR(K_1)$ and $MC(K_1)$ are decreasing in K_1. $\frac{\partial \Pi_{1W}}{\partial K_1}$ is continuous at $K_1 = n$ though its slope changes at this point.

Firm 1 will choose a K_1 larger than n if $\int_0^\infty \frac{\Delta_1 f(r)dr}{n+r} > b_2$. This holds in equilibrium when $b_1 > b_2$. K_1 is therefore determined by the first-order condition on the second segment of the profit function.

$$\int_{K_1-n}^{\infty} \frac{\Delta_1}{n+r} f(r)\, dr - b_2[1 - F(K_1 - n)] \geq 0 \tag{7}$$

K_1 will be finite if $MR(K_1)$ crosses $MC(K_1)$. At $K_1 = n$, $MR(K_1)$ is above $MC(K_1)$ and steeper than $MC(K_1)$. As K_1 increases, $\frac{dMR(K_1)}{dK_1} = -\frac{\Delta}{K_1} f(K_1 - n)$ and $\frac{dMC(K_1)}{dK_1} = -b_2 f(K_1 - n)$, so $MR(K_1)$ later becomes flatter than $MC(K_1)$. If $MR(K_1)$ does not cross $MC(K_1)$, $K_1 = \infty$. Figure 2 shows the case when K_1 is finite. Appendix 1 proves that the choice of K_1 is unique when equation (7) holds with equality. If $K_1 < \infty$, K_1 is increasing in Δ_1.

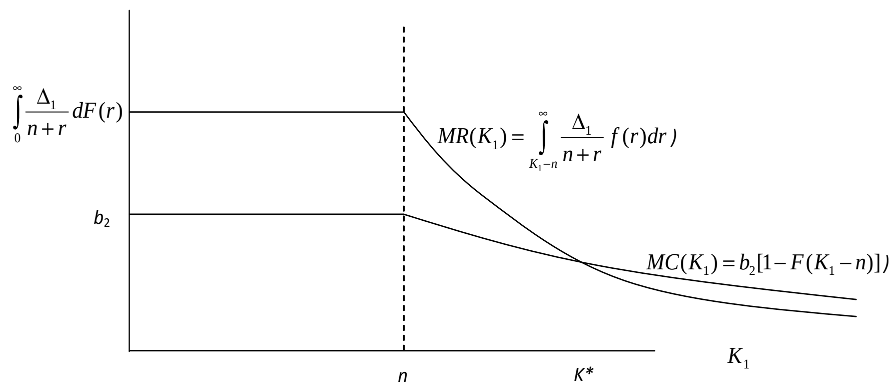


Figure 2: Firm 1's Choice of K_1

Bids As before, we calculate b_1 as the per-click payment that makes firm 1 indifferent between acquiring the advertising right and not acquiring it. Thus b_1 is found by setting $\Pi_{1W} = \Pi_{1L}$. We consider two cases: $n < K_1 < \infty$ and $K_1 = \infty$.

In the first case, K_1 will be finite when the firms are sufficiently similar that $MC(K)$ does not lie everywhere below $MR(K)$. To aid interpretation of the results, we assume symmetry between the two firms, $\Delta_1 = \Delta_2 = \Delta$, implying $K_1 = K_2 = K$ and $b_1 = b_2 = b$. We find b by equating firm 1's expected winning profits to its expected losing profits, but we now must consider that when firm 1 loses, it will claim the top spot when $n + r > K_2$. Thus

$$\Pi_{1W} = \int_0^{K-n} [n\pi_{1W} - b\,(n+r)]\, f(r)\, dr + \int_{K-n}^{\infty} (\Delta \frac{K}{n+r} + n\pi_{1L} - bK) f(r)\, dr\ , \tag{8}$$

$$\Pi_{1L} = \int_0^{K-n} (n\pi_{1L}) f(r)\, dr + \int_{K-n}^{\infty} (n\pi_{1W} - \Delta \frac{K}{n+r}) f(r)\, dr \tag{9}$$

and b is determined by the equality of Π_{1W} and Π_{1L}.

Proposition 4 *When inflationary click fraud is stochastic, there is no competitive click fraud, and firms are identical and set a finite K, expected gatekeeper revenues are strictly lower than the baseline model.*

Proof. Gatekeeper revenues are equal to firm 1's expenditure $\int_0^{K-n}[b(n+r)]f(r)\,dr + \int_{K-n}^{\infty} bKf(r)\,dr$

$$\begin{aligned}
&= \int_0^{K-n}(n\pi_{1W})f(r)\,dr + \int_{K-n}^{\infty}\left[\Delta\frac{K}{n+r} + n\pi_{1L}\right]f(r)\,dr - \\
&\quad - \int_0^{K-n} n\pi_{1L}f(r)\,dr - \int_{K-n}^{\infty}(n\pi_{1W} - \Delta\frac{K}{n+r})f(r)\,dr \\
&= \Delta\left\{2\left[\int_0^{K-n} f(r)\,dr + \int_{K-n}^{\infty}(\frac{K}{n+r})f(r)\,dr\right] - 1\right\}
\end{aligned} \tag{10}$$

Note that $\int_{K-n}^{\infty}(\frac{K}{n+r})f(r)\,dr < \int_{K-n}^{\infty} f(r)\,dr$, since $\frac{K}{n+r} < 1$ for every $r \in (K-n, \infty)$, so

$$\Delta\left\{2\left[\int_0^{K-n} f(r)\,dr + \int_{K-n}^{\infty}(\frac{K}{n+r})f(r)\,dr\right] - 1\right\} < \Delta. \tag{11}$$

The right-hand side is gatekeeper revenues when inflationary click fraud is deterministic. ■

In the second case, firm 1 wins and sets a capacity $K_1 = \infty$. This occurs when $\Delta_1 - \Delta_2$ is sufficiently large that $MR(K_1)$ lies everywhere above $MC(K_1)$.

Proposition 5 *When inflationary click fraud is stochastic, there is no competitive click fraud, and firms are sufficiently dissimilar that the high bidder sets $K_j = \infty$, expected gatekeeper revenues are strictly higher than the case when inflationary click fraud is deterministic.*

Proof: See Appendix 2.

Uncertainty about the amount of inflationary click fraud may either increase or decrease gatekeeper revenues. It is likely to lower gatekeeper revenues when firms' incremental profits of winning the auction are similar. In such situations, for example in auctions for generic keywords, bidding is more intense and the auction winner realizes a smaller profit from the auction. Low profits induce the auction winner to strategically limit its capacity to avoid paying for a large number of fraudulent clicks.

Gatekeeper revenues may rise with inflationary click fraud when one firm's profits of winning are much larger than its rival's (for example in auctions for branded keywords). In this case, the high bidder gains very large rents in the baseline model, and its rents are so large that it never chooses to strategically limit its capacity. Click fraud may then have the effect of transferring some of the winner's profits to the gatekeeper.

What we learn in this section is that inflationary click fraud does not harm advertisers when they know exactly how much to expect; this seemingly verifies the executive's comment that perhaps no solution to click fraud is necessary. However, under the more realistic assumption that firms face uncertainty in the level of inflationary click fraud, we see two things. First, search engines certainly have a strong incentive to detect and limit click fraud in very competitive keyword auctions. Second, when keyword auctions are less competitive, it may be in the gatekeeper's interest to allow some click fraud.

4 Search Advertising with Inflationary and Competitive Click Fraud

We have previously considered the effects of third-party invalid clicks on market equilibria. Now we extend the analysis to consider what happens when the low bidder may click the high bidder's ad to hasten the high bidder's exit from the advertising auction.

We start by proving our earlier assertion that competitive click fraud may not be subgame perfect. In a model where the number of inflationary fraudulent clicks is known and the number of competitive fraudulent clicks is rationally anticipated, firm 1 will shade its capacity upward in equilibrium. Assuming click fraud is costly, firm 2 then will not commit any competitive click fraud.

In section 4.2, we show that uncertainty in the total number of clicks may lead to competitive click fraud in equilibrium. Competitive click fraud unambiguously decreases advertisers' bids, but it also may increase the high bidder's capacity. As we show for two special cases of the model, the net effect on gatekeeper revenues may be positive or negative.

Assumptions About Competitive click fraud We assume the low bidder chooses a level of competitive click fraud, z, at cost $c(z)$. We assume $c(z)$ is increasing and convex since a larger number of clicks will increase the probability that the high bidder or the gatekeeper can verify the identity of the firm committing click fraud and retaliate (e.g., through civil lawsuits or business channels).[6]

We assume the low bidder chooses z simultaneously with the high bidder's choice of K. The total number of clicks is now $z + n + r$. We seek a rational expectations equilibrium in pure strategies under full information: each firm anticipates its rival's action.

[6] One might also posit a competitive click fraud cost function $c(z, r)$, where $\frac{dc}{dr} < 0$, to allow for the probability of competitive click fraud detection to fall with inflationary click fraud. We expect the two types of click fraud can be independently detected, given that website owners' fraudulent clicks will come exclusively from their own sites, while competitive click fraud is more likely to occur on search engines' main pages. The results presented below are virtually unchanged under the assumption that $c(z) = c(z, r)$.

4.1 Deterministic Inflationary and Competitive Click Fraud

In the case that r is deterministic and known to both firms, firm j's profit when it wins the initial auction is

$$\Pi_{jW} = \begin{cases} n\pi_{jW} - b_k(n + z_k + r) & \text{when } K_j \geq n + z_k + r \\ n\pi_{jW}\frac{K_j}{n+z_k+r} + \left(1 - \frac{K_j}{n+z_k+r}\right) n\pi_{jL} - b_k K_j & \text{when } K_j < n + z_k + r \end{cases} \tag{12}$$

Firm k's profit when it loses the initial auction is

$$\Pi_{kL} = \begin{cases} n\pi_{kL} - c(z_k) & \text{when } K_j \geq n + z_k + r \\ n\pi_{kL}\frac{K_j}{n+z_k+r} + \left(1 - \frac{K_j}{n+z_k+r}\right) n\pi_{kW} - c(z_k) & \text{when } K_j < n + z_k + r \end{cases} \tag{13}$$

These profit functions are similar to those analyzed in section 3.2. Proposition 6 describes equilibrium behavior.

Proposition 6 *When both firms know the inflationary click fraud level r, if firm j wins the auction, then in equilibrium $K_j \geq n + z_k + r$ and $z_k = 0$. Firm j never loses the top spot, and firm k therefore does not engage in competitive click fraud.*

Proof. Suppose not. If $z_k > 0$ and $K_j < n + z_k + r$, firm j's profit is $n\pi_{jW}\frac{K_j}{n+z_j+r} + \left(1 - \frac{K_1}{n+z_k+r}\right) n\pi_{jL} - b_k K_j$. This is strictly less than the case in which $K_j \geq n + z_k + r$. Therefore, firm j will always increase K_j until $K_j \geq n + z_k + r$. Firm k's best response to this strategy is $z_k = 0$. ■

4.2 Stochastic Inflationary and Competitive Click Fraud

Here we set up the problem under the general distribution $f(r)$ and discuss results and intuition from the general model. We describe the set of equilibria in pure strategies in Appendix 3.

Profits As before, there are two parts to firm 1's profit function. When $K_1 < n + z_2$, firm 1 is always knocked off the top spot. When $K_1 \geq n + z_2$, firm 1 is only knocked off for some realizations of r.

$$\Pi_{1W}=\begin{cases}\int_0^\infty(\Delta_1\frac{K_1}{n+z_2+r}+n\pi_{1L}-b_2K_1)f(r)\,dr & \text{when } K_1<n+z_2\\ \\ \int_0^{K_1-n-z_2}\left[n\pi_{1W}-b_2(n+z_2+r)\right]f(r)\,dr & \\ \quad+\int_{K_1-n-z_2}^{\infty}\left[\frac{\Delta_1K_1}{n+z_2+r}+n\pi_{1L}-b_2K_1\right]f(r)\,dr & \text{when } K_1\geq n+z_2\end{cases}\tag{14}$$

Π_{1W} is continuous at $K_1=n+z_2$ though its slope falls at this point.

The problem facing Firm 2 in the case that it loses is choosing z_2 to maximize

$$\Pi_{2L}=\begin{cases}\int_0^\infty\left[n\pi_{2L}\frac{K_1}{n+z_2+r}-c(z_2)+\left(1-\frac{K_1}{n+z_2+r}\right)n\pi_{2W}\right]f(r)\,dr & \\ & \text{when } K_1<n+z_2\\ \\ \int_0^{K_1-n-z_2}\left[n\pi_{2L}-c(z_2)\right]f(r)\,dr+ & \\ \quad+\int_{K_1-n-z_2}^{\infty}\left[n\pi_{2W}-\frac{\Delta_2K_1}{n+z+r}-c(z_2)\right]f(r)\,dr & \text{when } K_1\geq n+z_2\end{cases}\tag{15}$$

Π_{2L} is continuous at $K_1=n+z_2$ though its slope falls at this point.

Proposition 7 *Under stochastic inflationary click fraud, gatekeeper revenues may be increasing or decreasing in the level of competitive click fraud z.*

Proof. We prove this proposition with two special cases of the model. In the extreme case that $c(z)=0$, the low bidder's best strategy is to set $z=\infty$, erasing the high bidder's profit and driving bids to zero. In section 4.3, we solve for equilibrium for a special case of $f(r)$ and $c(z)$ and show that gatekeeper revenues may increase in the level of competitive click fraud. ■

The presence of both uncertain inflationary and competitive click fraud limit the high bidder's ability to react to either one. The high bidder mitigates inflationary click fraud by limiting its capacity to protect against paying for a large realization of r. The high bidder mitigates competitive click fraud by increasing its capacity, to prevent the low bidder from knocking it off with a large z. Thus when we add both types of click fraud into the model, the auction winner cannot respond optimally to either one without being hurt by the other. We show how this mechanism may operate for a special case of $f(r)$ and $c(z)$ in section 4.3.

4.3 Special case: $f(r)$ discrete and $c(z)$ linear

In this section, we solve a special case of the model for equilibrium K_1 and z_2. We assume that $r=\underline{r}$ with probability $1-\theta$ and $r=\overline{r}$ with probability θ, where $\underline{r}<\overline{r}$. This discrete distribution $f(r)$ is the only distribution that yields analytical solutions in this model. We also assume $\Delta_1 = \Delta_2 = \Delta$, to show that auction competition is not driving the results, and $c(z) = cz$ for simplicity.

Profit and reaction functions If firm 1 wins the auction, its profit is

$$\Pi_{1W} = \begin{cases} n\pi_{1W} - b_2(1-\theta)(n+\underline{r}+z_2) - b_2\theta(n+z_2+\overline{r}); \\ \qquad\qquad\qquad\qquad \text{when } K_1 \geq n+\overline{r}+z_2 \\ \\ (1-\theta)\left[n\pi_{1W} - b_2(n+\underline{r}+z_2)\right] + \theta\left(\frac{K_1\Delta}{n+\overline{r}+z_2} + n\pi_{1L} - b_2K_1\right); \\ \qquad\qquad\qquad\qquad \text{when } n+\underline{r}+z_2 \leq K_1 < n+\overline{r}+z_2 \\ \\ (1-\theta)\left[\frac{K_1\Delta}{n+\underline{r}+z_2}\right] + \theta\left(\frac{K_1\Delta}{n+\overline{r}+z_2}\right) + n\pi_{1L} - b_2K_1; \\ \qquad\qquad\qquad\qquad \text{when } K_1 < n+\underline{r}+z_2 \end{cases} \tag{16}$$

Π_{1W} is continuous and piecewise linear. It is flat for $K_1 \geq n+\overline{r}+z_2$.

$$\frac{\partial \Pi_1}{\partial K_1} = \begin{cases} \theta\left(\frac{\Delta}{n+\overline{r}+z_2} - b_2\right), & \text{For } n+\underline{r}+z_2 \leq K_1 < n+\overline{r}+z_2 \\ \Delta\left(\frac{1-\theta}{n+\underline{r}+z_2} + \frac{\theta}{n+\overline{r}+z_2}\right) - b_2, & \text{For } K_1 < n+\underline{r}+z_2 \end{cases} \tag{17}$$

We first show it can't decrease on the first segment and then increase on the second. From the slope expressions, if it did, then

$$\Delta\left(\frac{1-\theta}{n+\underline{r}+z_2} + \frac{\theta}{n+\overline{r}+z_2}\right) < b_2 < \frac{\Delta}{n+\overline{r}+z_2} \tag{18}$$

which cannot happen since $\underline{r} < \overline{r}$.

Figure 3 shows the three possible shapes Π_1 can take in K_1.

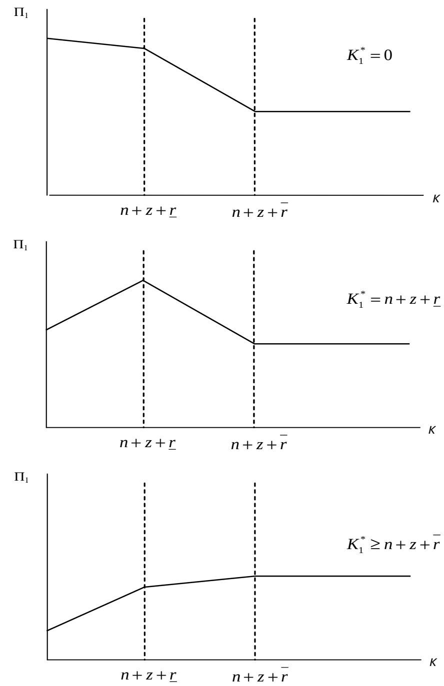


Figure 3: Possible shapes of Π_{1W}

Equilibrium capacity may be given by $K_1^* = 0$, $K_1^* = n + z_2 + \underline{r}$, or $K_1^* \geq n + z_2 + \overline{r}$, depending on the shape of Π_1. The middle case occurs when (rewriting the slope conditions and evaluating at $K_1 = n + \underline{r} + z_2$)

$$\Delta \left(\frac{1-\theta}{K_1} + \frac{\theta}{K_1 + \overline{r} - \underline{r}} \right) > b_2 > \frac{\Delta}{K_1 + \overline{r} - \underline{r}}. \tag{19}$$

If the first inequality is violated, $K_1^* = 0$. If the second inequality is violated, $K_1^* \geq n + z_2 + \overline{r}$.

Now we consider the auction loser. Firm 2's profit function is

$$\Pi_{2L} = \begin{cases} n\pi_{2L} - cz_2, & \text{when } K_1 \geq n + \overline{r} + z_2 \\ (1-\theta)n\pi_{2L} + \theta \left(\frac{-K_1\Delta}{n+\overline{r}+z_2} + n\pi_{2W} \right) - cz_2, & \text{when } n + \underline{r} + z_2 \leq K_1 < n + \overline{r} + z_2 \\ (1-\theta) \left(\frac{-K_1\Delta}{n+\underline{r}+z_2} \right) + \theta \left(\frac{-K_1\Delta}{n+\overline{r}+z_2} \right) + n\pi_{2W} - cz_2 & \text{when } K_1 < n + \underline{r} + z_2 \end{cases} \tag{20}$$

For $K_1 \geq n+\overline{r}+z_2$, $\frac{\partial \Pi_2}{\partial z_2} = -c$ and $z_2 = 0$. For $n+\underline{r}+z_2 \leq K_1 < n+\overline{r}+z_2$, $\frac{\partial \Pi_2}{\partial z_2} = \frac{\theta K_1 \Delta}{(n+\overline{r}+z_2)^2} - c = 0$ and $z_2 = \sqrt{\frac{\theta K_1 \Delta}{c}} - n - \overline{r}$. For $K_1 < n+\underline{r}+z_2$, $\frac{\partial \Pi_2}{\partial z_2} = K_1 \Delta \left[\frac{(1-\theta)}{(n+\underline{r}+z_2)^2} + \frac{\theta}{(n+\overline{r}+z_2)^2}\right] - c = 0$ defines z_2. Second-order conditions are satisfied for $K_1 < n + \overline{r} + z_2$.

Equilibrium in K_1 and z_2 We find two equilibria in pure strategies. In the first, Π_{1W} is rising in its third segment, $K_1 \geq n + \overline{r} + z_2$, and $z_2 = 0$.

The other possibility is that Π_{1W} peaks at $K_1 = n + \underline{r} + z_2$, in which case firm 2 responds according to its first-order condition. We then have

$$z_2^* = \sqrt{\frac{\theta^2 \Delta^2}{4c^2} + \theta \Delta (\underline{r} - \overline{r})} + \frac{\theta \Delta}{2c} - n - \overline{r} \tag{21}$$

$$K_1^* = \sqrt{\frac{\theta^2 \Delta^2}{4c^2} + \theta \Delta (\underline{r} - \overline{r})} + \frac{\theta \Delta}{2c} - \overline{r} + \underline{r} \tag{22}$$

There are three necessary conditions for this equilibrium. First, it must be that K_1^* is in the prescribed range, which implies $z_2^* < \overline{r} - \underline{r}$. Second, it must be that $z_2^* > 0$. Third, it must be that firm 2 prefers $\Pi_{2L}(K_1^*, z_2^*)$ to $\Pi_{2L}(K_1^*, 0)$; this implies $\theta \Delta \left(1 - \frac{K_1^*}{K_1^* + \overline{r} - \underline{r}}\right) > c z_2^*$. The equilibrium level of competitive click fraud is increasing with $\overline{r}$ and decreasing with c.

Gatekeeper revenues We now evaluate gatekeeper revenues. Total payments made to the gatekeeper in equilibrium are $b^* K^*$ as determined by firm 1's indifference to winning and losing: $\Pi_{1W}^*(K^*, z^*, b^*) = \Pi_{1L}^*(K^*, z^*)$ implies

$$n\pi_{1W} - \theta \Delta \left(1 - \frac{K^*}{K^* + \overline{r} - \underline{r}}\right) - b^* K^* = n\pi_{1L} + \theta \Delta \left(1 - \frac{K^*}{K^* + \overline{r} - \underline{r}}\right) - c z^* \tag{23}$$

or

$$b^* K^* = \Delta \left[1 - 2\theta \Delta \left(1 - \frac{K^*}{K^* + \overline{r} - \underline{r}}\right)\right] + c z^* \tag{24}$$

It can be seen that gatekeeper revenues, $b^* K^*$, rise with z^*.

5 Asymmetric Click-Through Rates

Search engines commonly use advertisers' click-through rates (CTRs) in conjunction with per-click bids to determine ad position listings (Rutz and Bucklin 2007a, 2007b). It is therefore interesting to consider whether our main results would change under a more realistic keyword auction.

We are not aware of any previous papers to solve for equilibrium bidding strategies in an auction with a CTR component, or to investigate the effects of the gatekeeper's choice of auction mechanism on advertising revenues.[7] We show that moving from a second price auction to what we term a "click-through auction" (bid-per-click*total-clicks) has three effects. First, advertisers no longer bid their reservation price, and the less profitable firm may win the auction, even when it has a lower click-through rate. Second, search engine revenues may be lower in a click-through auction than in a second-price auction. Third, inflationary click fraud may increase search engine revenues in a click-through auction, even when the amount of fraud is deterministic and known to both bidders.

5.1 Baseline Model with a Click-Through Auction

We now assume that firm j will get n_j clicks during the unit time period. What follows is identical if n_j is a fraction $\eta_j \in (0,1)$ of total potential clicks N, where $n_j = \eta_j N$ (as click-through rates are commonly defined).

It is useful to denote firm j's per-click value of winning as $v_j \equiv \pi_{jW} - \pi_{jL}$. We assume firms are numbered such that $v_1 \geq v_2$.

Firm asymmetry in clicks is different from firm asymmetry in advertising profits. It may be that firms' clicks are identical ($n_1 = n_2$) but their variable profits are different due to price or cost factors. It may also be that firms' reservation prices are identical ($v_1 = v_2$) but their clicks are different due to factors like brand recognition or ad quality. It seems likely that n_j and v_j would be positively related, but we do not require it.

We distinguish between two types of auction:

the **Second Price (SP) auction**, which we used in previous sections. The highest bid per click determines the auction winner, and the winner pays the loser's bid on each click it receives.

the **Click-Through (CT) auction**, in which the gatekeeper allocates the advertising slot to firm j if and only if $n_j b_j > n_k b_k$. The winning bidder then pays the loser's bid per click. We assume that n_1 and n_2 are known to both advertisers and the gatekeeper.

As before, we solve the second stage first. When firm 1 wins the CT auction, its profits are

$$\Pi_{1W} = \begin{cases} n_1 (\pi_{1W} - b_2) & \text{for } K_1 \geq n_1 \\ \pi_{1W} K_1 + \left(1 - \frac{K_1}{n_1}\right) n_1 \pi_{1L} - b_2 K_1 & \text{for } 1 \leq K_1 < n_1. \end{cases} \tag{25}$$

[7] Ghose and Yang (2007) include click-through rate in their empirical model, but they do not solve for equilibrium bidding strategies, they assume that rank is continuous, and that advertisers pay their own bids rather than the next-lowest bid.

First-order conditions indicate the firm will set $K_1 \geq n_1$ if $v_1 > b_2$. If firm 1 loses the auction and firm 2 sets a $K_2 \geq n_2$, its profits will be $\Pi_{1L} = n_1 \pi_{1L}$.

In the SP auction firm j chose its bid b by setting $\Pi_{jW} = \Pi_{jL}$. However, it is not optimal for both firms to bid their reservation price in the CT auction. To see this, assume that $n_1 < n_2$ and $b_2 = v_2$. If $b_1 = v_1$, firm 1 will lose the auction. However, for $b_1 = \frac{n_2}{n_1} v_1 > v_1$, firm 1 can win the auction and earn a positive profit on each click.

Lemma 1 *A weakly dominant strategy is to bid $b_j^* = \frac{n_k}{n_j} v_j$. Firm j will win the CT auction if and only if $n_k v_j > n_j v_k$.*

Proof. Consider two mutually exclusive cases. For $b_2 < v_1$, winning the auction produces a positive profit for firm 1. No $b_1 > b_1^*$ will increase firm 1's profits, while a $b_1 < b_1^*$ can only decrease profits by reversing the profitable auction outcome. For $b_2 > v_1$, firm 1 cannot profitably win the auction. $b_1 = b_1^*$ ensures it loses. No $b_1 < b_1^*$ can improve profits, while a $b_1 > b_1^*$ can only change profits by unprofitably reversing the auction outcome. The proof for firm 2 is symmetric. ■

Lemma 1 shows that the low-click firm bids more aggressively in the CT auction than in the SP auction, while the high-click firm bids more passively. This happens because the auction mechanism handicaps the high-click firm. It does so by making it susceptible to the threat of negative variable profits produced by an aggressive bid by the low-click firm.

Substituting firm j's equilibrium bid into the gatekeeper's auction mechanism $n_j b_j > n_k b_k$ implies that $v_j > b_k$. In equilibrium neither firm wins the auction at an unprofitable per-click payment. Thus when firm j wins the auction, it optimally sets a capacity $K_j \geq n_j$.

When both firms play optimal strategies, firm 1 will win if and only if $\frac{v_1}{v_2} > \frac{n_1}{n_2}$. Thus, even if firm 1 has a higher per-click profit and a higher click-through rate, it may lose the keyword auction due to its rival's ability to bid aggressively. The high-value firm is only assured of winning the auction if its relative profit advantage is larger than its relative click-through advantage.

We now compare gatekeeper revenues in the CT and SP auctions.

Proposition 8 *A switch from a Second Price auction to a Click Through auction increases gatekeeper revenues if and only if $\frac{v_1}{v_2} > \frac{n_1}{n_2} > 1$. Otherwise, the Click Through auction produces lower revenues than a Second Price auction.*

Proof. From section 2, firm 1 will always win the SP auction and pay the gatekeeper $n_1 v_2$. Consider two mutually exclusive cases. If $\frac{n_1}{n_2} < \frac{v_1}{v_2}$, firm 1 also wins the CT auction and pays the gatekeeper $n_1 b_2^* = \frac{n_1^2}{n_2} v_2$. This is more than in the SP auction if and only if $n_1 > n_2$. In the second case, if $\frac{n_1}{n_2} > \frac{v_1}{v_2}$, firm 2 wins the CT auction and pays the gatekeeper $n_2 b_1^* = \frac{(n_2)^2}{n_1} v_1$. This is less than gatekeeper revenues in the SP auction: $\frac{n_1}{n_2} > \frac{v_1}{v_2} \iff \frac{n_2^2}{n_1} v_1 < n_1 v_2$. ■

The Click Through auction only produces higher revenues than the Second Price auction if the high-value firm's relative click advantage is not too large compared to its relative profit advantage. The possibility that the CT auction can lower search engine revenues is counterintuitive and contrary to the conventional wisdom. It may suggest that Google's early adoption of the CT auction lowered advertisers' costs, encouraging them to buy more keywords on Google's platform than on rival platforms. It also may help explain Yahoo's late adoption of the CT auction, or Google's switch to its unspecified use of "quality scores" in 2006.

5.2 Deterministic Inflationary Click Fraud in a Click Through Auction

Our motivation to consider the CT auction is to determine whether it reverses our result that gatekeeper revenues may increase with click fraud. In the SP auction, both bidders can price out the effects of inflationary click fraud when its quantity is known, yielding no effect on search engine revenues. In the CT auction, however, inflationary click fraud alters the ratio of firms' click-through rates. Click fraud may increase search engine revenues when it reduces the high-value advertiser's relative advantage in clicks.

We again assume that content network websites generate r fraudulent clicks, and both bidders know r. The gatekeeper now awards the advertising slot to firm j if and only if $(n_j + r)b_j > (n_k + r)b_k$. The previous analysis indicates that firm 1 will win if and only if $\frac{v_1}{v_2} > \frac{n_1+r}{n_2+r}$. (To see this, relabel each firm's click level with $n_j' = n_j + r$.) The gatekeeper's revenues are $\frac{(n_1+r)^2}{n_2+r} v_2$ if $\frac{v_1}{v_2} > \frac{n_1+r}{n_2+r}$, or $\frac{(n_2+r)^2}{n_1+r} v_1$ otherwise.

Proposition 9 *When firm j wins the CT auction, if deterministic inflationary click fraud does not reverse the auction result, gatekeeper revenues will increase if $r > n_j - 2n_k$.*

Proof. When firm j wins the CT auction it pays the gatekeeper $(n_j + r)b_k^* = \frac{(n_j+r)^2}{n_k+r} v_k$. Taking the derivative shows that gatekeeper revenues are increasing in r if and only if $r > n_j - 2n_k$. ■

Note, we have ignored the possibility that the click fraud level r responds to advertisers' bids. What would happen if we allowed r to depend on the auction winner? Let us assume that $n_1 > n_2$, $r_1 > r_2$, and $\frac{v_1}{v_2} > \frac{n_1+r_1}{n_2+r_2}$, so the high-profit firm is also the high-traffic firm and the auction winner. Gatekeeper revenues are then $\frac{(n_1+r_1)^2}{n_2+r_2} v_2$, greater than gatekeeper revenues for any common click fraud level $r \in (r_2, r_1)$.

Inflationary click fraud has two effects in the Click Through auction. First, it lowers the threshold at which a high-value, high-click firm wins the auction, expanding the parameter space in which the high-value firm wins. Second, it alters the low-value firm's ability to threaten the high-value firm. The smaller is the high-value firm's relative click advantage, the more likely a given level of click fraud will be beneficial to the search engine.

We have shown that our main result, that search engine revenues may increase with click fraud, is not an artifact of our assumed auction mechanism. We proceed with the normative implications of our analysis.

6 Managerial Implications

Our analysis has produced several results that could influence search engines' and advertisers' business practices. We note that our implications are subject to the limitations discussed in the final section.

Content network management We showed that third-party websites' incentives to engage in click fraud are greater when a per-click compensation scheme is used in place of a revenue-sharing compensation scheme, and when search ads are rotated across a large number of websites. Content networks should not only adopt these strategies, they should make them public to increase transparency and build advertiser confidence.

We found that content network partners' incentives to engage in click fraud are minimized when advertisers may enter site-specific bids. Any site that generates a large amount of inflationary click fraud would then be penalized through a lower site-specific bid. We are not aware of any content networks that currently allow advertisers to enter site-specific bids in CPC auctions, but it seems within the realm of technical possibility.

Note, we have not modeled websites' choice to enter or remain in a content network. It may be that decreasing websites' incentives to commit click fraud could also reduce search engines' inventory of customer clicks by encouraging websites to enroll in competing content networks or reducing their incentives to invest in content.

Advertiser Information We showed that click fraud does advertisers no harm when advertisers have full information in a second-price auction. This suggests that search engines should take actions to increase the amount of information at advertisers' disposal. Specifically, they can issue keyword-specific reports on how and when they punish advertisers and websites suspected of engaging in click fraud, issue keyword-specific reports on when and how much click fraud they detect, and give advertisers information about the identity and frequency of the content network sites on which their ads will appear.

Tuzhilin's "Fundamental Problem of Click Fraud Prevention" Tuzhilin (2006) defined the "fundamental problem of click fraud prevention." Search engines may try vigorously to detect and prevent click fraud, but they cannot tell advertisers specifically how they do so, as this would constitute explicit instructions on how to avoid click fraud detection. It also may be that in identifying fraudulent clicks, search engines must make probabilistic judgments balancing "false positives" against "false negatives." Presumably, search engines would prefer to minimize false negatives, while advertisers would prefer to maximize false positives.

To resolve this problem, we suggest that the search advertising industry form a neutral third party to authenticate search engines' click fraud detection efforts. Such a party could maintain the confidentiality needed by search engines while allaying advertisers' concerns.

Similar third parties are used in other media industries. For example, Nielsen Media Research's audience measurements underpin transactions between television networks and advertisers, the Audit Bureau of Circulations authenticates newspapers' and magazines' subscription figures, and comScore and other companies measure website audiences for display (CPM) advertising transactions. In the absence of such a neutral third party, it may be possible to design some creative incentive-compatible contracts to provide verifiable evidence of click quality. For example, if human searchers are each assigned individual-specific accounts, advertisers could enter different bids for clicks made from individuals' accounts, and "anonymous" clicks.

Our result that search engines are sometimes helped, and sometimes hurt, by click fraud reinforces the need for such a neutral third party. Advertisers may perceive the risk that search engines do not apply the same click fraud detection algorithms to all keyword auctions. Our results suggest that a profit-maximizing search engine might exert maximal efforts to prevent click fraud in competitive keyword auctions but do less to prevent click fraud in relatively uncompetitive auctions such as those for branded keywords. Or it may be that search engines try vigorously to prevent click fraud but are unable to credibly convey the depths of their efforts to concerned advertisers.

We have only modeled one gatekeeper; would competition between gatekeepers resolve the click fraud problem? We think not, for two reasons. First, the "fundamental problem of click fraud detection" would still prevent search engines from sending credible signals to advertisers about their click fraud detection efforts. Second, so long as advertisers realize profits per click, and consumers are distributed across search engines, the profit-maximizing advertiser is likely to buy keywords from all search engines (though its bid may vary across search engines).

Will click fraud destroy the market? Our results suggest it seems unlikely that click fraud will ever completely destroy the search advertising industry. First, we found that when advertisers have full information in a SP auction, they can strategically adjust their bids and advertising budgets to mitigate the effects of click fraud. Second, so long as search engines are able to maintain a positive probability of detecting some click fraud and punishing those responsible, we will see limited click fraud in equilibrium. It seems the CPC business model will likely remain viable in the long run.

7 Discussion

We have presented the first analysis of the effects of inflationary and competitive click fraud on search advertising markets. We found that, when advertisers know the level of inflationary click fraud in a second price auction, they lower their bids to the point that click fraud has no impact on total advertising expenditures. However, when the level of inflationary click fraud is uncertain, total advertising expenditures may rise or fall. They rise when the keyword auction is relatively less competitive since advertising is so profitable for the high bidder that it is willing to pay to remain on top for any realization of click fraud. Advertising expenditures may fall when the keyword auction is more competitive since the high bidder faces higher advertising costs and therefore shades its capacity downward to protect against paying for large levels of inflationary click fraud. Even when inflationary click fraud is known to both bidders in a click-through auction, it can enhance the low-value firm's ability to bid aggressively, thereby increasing gatekeeper revenues.

We also analyzed the effects of competitive click fraud in the second price auction. We found that when inflationary click fraud is deterministic, a high-bidding firm may effectively deter its rival from committing click fraud by choosing a large capacity. However, when the number of clicks is stochastic, the high bidder may shade its bid downward and the low bidder may then profitably engage in competitive click fraud. We showed that gatekeeper revenues may be increasing or decreasing in the level of competitive click fraud.

As in all models, we have made several simplifying assumptions. Two assumptions in particular suggest directions for future research. The first is the assumption that the gatekeeper offers only one advertising slot. Edelman, Ostrovsky, and Schwarz (2007) show that when only one search ad is available, the auction mechanism used in practice is identical to a standard Second Price auction. However, when more than one ad is offered, the auction mechanism is what they term a "Generalized Second Price" ("GSP") auction. Unlike the SP auction, the GSP does not have an equilibrium in dominant strategies, and firms do not engage in truth-telling. This technical concern limits our ability to make predictions about firms' equilibrium click fraud strategies in a multiple-slot auction. We suspect that adding more slots and advertisers to the model would increase click fraud, since more advertisers would stand to gain from knocking off the highest bidder.

The second assumption that could be relaxed in future work is the assumption of a single gatekeeper. Expanding the analysis to multiple gatekeepers could introduce elements of two-sided market competition. Search engines may choose their policies based on the possibility of advertiser, searcher, or content-network website defection to competitors. In this paper we have not considered that advertiser adoption of a search engine's platforms could be a function of its business model.

References

[1] Advertising Age. 2006. *Search Marketing Fact Pack 2006.*

[2] Anderson, S. P., R. Renault. 1999. Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model. *RAND Journal of Economics* **30**(4) 719-735.

[3] Anderson, S. P., R. Regis. 2001. Consumer Information and Firm Pricing: Negative Externalities from Improved Information. *International Economic Review* **41**(3) 721-742.

[4] Balasubramanian, S. 1998. Mail versus Mall: A Strategic Analysis of Competition between Direct Marketers and Conventional Retailers. *Marketing Science* **17**(3) 181-195.

[5] Baye, M. R., J. Morgan. 2001. Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets. *American Economic Review* **91**(3) 454-474.

[6] Borgers, T., I. Cox, M. Pesendorfer, V. Petricek. 2007. Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence. Mimeo. University of Michigan.

[7] Bradlow, E. T., D. C. Schmittlein. 2000. The Little Engines that Could: Modeling the Performance of World Wide Web Search Engines. *Marketing Science* **19**(1) 43-62.

[8] Chatterjee, P., D. L. Hoffman, T. P. Novak. 2003. Modeling the Clickstream: Implications for Web-Based Advertising Efforts. *Marketing Science* **22**(4) 520-541.

[9] Chen, Y., C. He. 2006. Paid Placement: Advertising and Search on the Internet. Mimeo, University of Colorado.

[10] Danaher, P., G. W. Mullarkey. 2003. Factors Affecting Online Advertising Recall: a Study of Students. *Journal of Advertising* **43**(3) 252-267.

[11] Daswani, N., M. Stoppelman. 2007. The Anatomy of Clickbot.A. Mimeo, Google Inc.

[12] Dreze, X., A. Bonfrer. 2006. Real-Time Evaluation of Email Campaign Performance. Mimeo, University of Pennsylvania.

[13] Dreze, X., F. Zufryden. 1998. Is Internet Advertising Ready for Prime Time? *Journal of Advertising Research* **38**(3) 7-18.

[14] Dukes, A., E. Gal-Or. 2003. Negotiations and exclusivity contracts for advertising. *Marketing Science* **22**(2) 222-245.

[15] Edelman, B., M. Ostrovsky, M. Schwarz. 2007. Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords. *American Economic Review* **97**(1).

[16] Ghosemajumder, Shuman. 2006. Let Click Fraud Happen? Um, No. http://googleblog.blogspot.com/2006/07/let-click-fraud-happen-uh-no.html. Accessed Nov. 4 2007.

[17] Ghose, A., S. Yang. 2007. An Empirical Analysis of Paid Placement in Online Advertising. Mimeo. New York University.

[18] Goldfarb, A., C. E. Tucker. 2007. Search Engine Advertising: Pricing Ads to Context. Mimeo. University of Toronto.

[19] He, C., Y. Chen. 2006. Managing e-Marketplace: a Strategic Analysis of Nonprice Advertising. *Marketing Science* **25**(2) 175-187.

[20] Iyer, G., A. Pazgal. 2003. Internet Shopping Agents: Virtual Co-Location and Competition. *Marketing Science* **22**(1) 85-106.

[21] Lal, R., M. Sarvary. 1999. When and How is the Internet Likely to Decrease Price Competition? *Marketing Science* **18**(4) 485-503.

[22] Manchanda, P., J. Dube, K. Y. Goh, and P. K. Chintagunta. 2006. The Effect of Banner Advertising on Internet Purchasing. *Journal of Marketing Research* **43**(1) 98-108.

[23] Mayzlin, D. 2006. Promotional Chat on the Internet. *Marketing Science* **25**(2) 155-163.

[24] Prasad, A. 2007. Website Traffic Claims and Internet Advertising. Mimeo, University of Texas at Dallas.

[25] Rutz, O. J., R. E. Bucklin. 2007a. A Model of Individual Keyword Performance in Paid Search Advertising. Mimeo, Yale University.

[26] Rutz, O. J., R. E. Bucklin. 2007b. From Generic to Branded: a Model of Spillover Dynamics in Paid Search Advertising. Mimeo, Yale University.

[27] Shields, M. 2007. Yahoo's 'Panama' Shows Click-Rate Lift. *Mediaweek.com* February 26.

[28] Stone, B. 2005. When Mice Attack. Newsweek 145(4) 52.

[29] Telang, T., P. Boatwright, T. Mukhopadhyay. 2004. A Mixture Model for Internet Search Engine Visits. *Journal of Marketing Research* **41**(May) 206-214.

[30] Tuzhilin, A. 2006. The Lane's Gift v. Google Report. Mimeo, New York University.

[31] Varian, H. 2006. Position Auctions. *International Journal of Industrial Organization* **25** (6) 1163-1178.

[32] Vickrey, W. 1961. Counterspeculation, Auctions and Competitive Sealed Tenders. *Journal of Finance* **16**(1) 8-37.

[33] Wilk, K. 2007. Rank Does Matter. *iMedia Connection* February 21.

Appendix 1

Here we prove that, if $K_1 < \infty$, it is unique. For $K_1 > n$, Firm 1's second-order condition is

$$\frac{\partial^2 \Pi_{1W}}{\partial K^2} = \left(\frac{-\Delta_1}{K_1} + b_2 \right) f(K_1 - n) \tag{26}$$

We can show that if the first-order condition is satisfied, the second-order condition is strictly negative, implying Π_{1W} is strictly concave. $\frac{\partial \Pi_{1W}}{\partial K} = 0$ implies

$$\frac{\Delta_1}{b_2} = \frac{[1 - F(K_1 - n)]}{\int_{K_1 - n}^{\infty} \frac{f(r)dr}{n+r}} \tag{27}$$

Substituting this into the second-order condition gives

$$\frac{\partial^2 \Pi_{1W}}{\partial K^2} = \left[\frac{-[1 - F(K_1 - n)]}{\int_{K_1-n}^{\infty} \frac{K_1 f(r) dr}{n+r}} + 1 \right] b_2 f(K_1 - n) \tag{28}$$

when the first-order condition holds with equality. Under the bounds of the integral, we have $\frac{K_1}{n+r} \leq 1$ for every term $r \geq K_1 - n$. Thus

$$\left[\int_{K_1-n}^{\infty} \frac{K_1}{n+r} f(r) dr - [1 - F(K_1 - n)] \right] < \left[\int_{K_1-n}^{\infty} f(r) dr - [1 - F(K_1 - n)] \right] = 0 \tag{29}$$

Therefore the second-order condition is strictly satisfied whenever the first-order condition holds with equality.

Appendix 2

Here we prove that, under stochastic inflationary click fraud and no competitive click fraud, when $K_1 = \infty$, expected gatekeeper revenues may be larger than in the baseline model. We have

$$\Pi_{1W} = \int_0^{\infty} [n\pi_{1W} - b_2 (n + r)] f(r) dr \tag{30}$$

and

$$\Pi_{1L} = \int_0^{K_2-n} (n\pi_{1L}) f(r) dr + \int_{K_1-n}^{\infty} (n\pi_{1W} - \Delta_1 \frac{K_1}{n+r}) f(r) dr \tag{31}$$

when $n \leq K_2 < \infty$. Gatekeeper revenue when firm 1 wins is $\int_0^{\infty} b_2 (n + r) f(r) dr$, so we need to find b_2.

Firm 2 chooses b_2 to set $\Pi_{2W} = \Pi_{2L}$. From above, we have

$$\Pi_{2W} = \int_0^{K_2-n} [n\pi_{2W} - b_2 (n + r)] f(r) dr + \int_{K_2-n}^{\infty} (\Delta_2 \frac{K_2}{n+r} + n\pi_{2L} - b_2 K_2) f(r) dr \tag{32}$$

and $\Pi_{2L} = n\pi_{2L}$. b_2 is chosen to satisfy $\Pi_{2W} = \Pi_{2L}$, or

$$\begin{aligned} &\int_0^{K_2-n} b_2 (n + r) f(r) dr + \int_{K_2-n}^{\infty} b_2 K_2 f(r) dr \\ = \; &\Delta_2 (\int_0^{K_2-n} f(r) dr + \int_{K_2-n}^{\infty} \frac{K_2}{n+r} f(r) dr). \end{aligned} \tag{33}$$

From Firm 2's FOC in K_2, we have

$$\int_{K_2-n}^{\infty}\left[\frac{\Delta_2}{n+r}-b_1\right]f(r)\,dr=0,\text{ and }b_1>b_2 \tag{34}$$

so

$$\int_{K_2-n}^{\infty}\left[\frac{\Delta_2}{n+r}-b_2\right]f(r)\,dr>0. \tag{35}$$

Therefore, we know that

$$\begin{aligned}
&\int_0^{K_2-n}\left[b_2(n+r)\right]f(r)\,dr\\
=\;&\Delta_2\int_0^{K_2-n}f(r)\,dr+K_2\int_{K_2-n}^{\infty}\left(\frac{\Delta_2}{n+r}-b_2\right)f(r)\,dr\\
>\;&\Delta_2\int_0^{K_2-n}f(r)\,dr.
\end{aligned} \tag{36}$$

We can now look at expected gatekeeper revenues,

$$\begin{aligned}
&\int_0^{\infty}b_2(n+r)f(r)\,dr\\
>\;&\int_0^{\infty}\frac{\Delta_2\int_0^{K_2-n}f(r)\,dr}{\int_0^{K_2-n}(n+r)f(r)\,dr}(n+r)f(r)\,dr\\
=\;&\Delta_2\frac{E(n+r)}{E(n+r|n+r<K_2)}\\
>\;&\Delta_2.
\end{aligned} \tag{37}$$

Δ_2 is gatekeeper revenues in the baseline model, so we have shown that gatekeeper revenues are strictly larger when the two firms are sufficiently different that the high bidder sets an infinite capacity in spite of uncertain inflationary click fraud.

Appendix 3

Here we solve for the set of post-auction equilibria in pure strategies in the general model under symmetric click-through rates. Advertisers' beliefs about the distribution of inflationary click fraud are $f(r)$, and the low bidder may engage in competitive click fraud at cost $c(z)$. We start by drawing the high bidder's reaction function in K, the low bidder's reaction function in z, and finally analyze where they may cross. We assume, without loss of generality, that the firms are numbered such that firm 1 wins the advertising auction.

For $K_1 < n + z_2$, Π_{1W} changes linearly with K_1 at rate $\int_0^\infty \frac{\Delta_1 f(r)dr}{n+z_2+r} - b_2$. For $K_1 > n + z_2$, the rate of change is strictly greater:

$$\frac{\partial \Pi_{1W}}{\partial K_1} = \int\limits_{K_1-n-z_2}^{\infty} \frac{\Delta_1 f\,(r)\,dr}{n+z_2+r} - b_2[1 - F(K_1 - n - z_2)] \tag{38}$$

Note that

$$\int_0^\infty \frac{\Delta_1 f\,(r)\,dr}{n+z_2+r} \begin{array}{c} < \\ = \\ > \end{array} b_2 \text{ then } \begin{cases} K_1 = 0 \\ K_1 \in [0, n+z_2] \\ K_1 > n + z_2 \end{cases} \tag{39}$$

In the first case, firm 1 has to pay more per click than it earns while it is on top, so it never sets a positive capacity. In the second case, firm 1 earns zero net profit per click while on top so it may set any capacity up to $n + z_2$. In the third case, firm 1 profits from remaining on top and sets $K_1 > n + z_2$. We focus on this final case in what follows, as the first two cases are not subgame perfect. If firm 2 is the low bidder, it must not be the case that b_2 exceeds the high bidder's equilibrium profit per click.

As in section 3.3, firm 1's choice of $K_1 > n + z_2$ may or may not yield a finite K_1. If the firm's first-order condition is satisfied, its second-order condition implies K_1 is a unique maximum; if not, $K_1 = \infty$. The proof is parallel to that presented in Appendix 1, so it is omitted here.

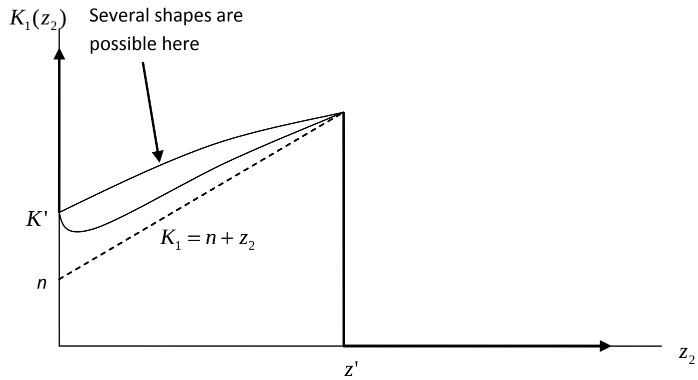
We have characterized firm 1's response to z_2 and shown that K_1 is unique when $K_1 > n + z_2$. The next question is whether K_1 is increasing or decreasing in z_2. We can apply the implicit function theorem to $\frac{\partial \Pi_{1W}}{\partial K_1} = 0$ to find that

$$\frac{\partial K^*}{\partial z} = 1 - \frac{K_1 \Delta_1 \int_{K_1-n-z_2}^{\infty} (n+z_2+r)^{-2} f(r)dr}{(\Delta_1 - K_1 b)\, f(K_1 - n - z_2)}, \text{ when } K_1 > n + z_2 \tag{40}$$

The numerator is positive, and the denominator is also positive (this is implied by $\frac{\partial^2 \Pi_{1W}}{\partial K^2} < 0$). For z_2 such that

$$K_1 \Delta_1 \int\limits_{K_1-n-z_2}^{\infty} (n+z_2+r)^{-2} f(r)dr > (\Delta_1 - K_1 b)\, f(K_1 - n - z_2) \tag{41}$$

then K_1 slopes downward in z_2; otherwise it is increasing. Figure 4 shows firm 1's reaction function in this case.



$$z' = z \text{ s.t.} \int_0^\infty \frac{\Delta_1 f(r)dr}{n+r+z} = b_2 \qquad K' = K \text{ s.t.} \int_{K-n}^\infty \frac{\Delta_1 f(r)dr}{n+r} = b_2[1-F(K-n)]$$

Figure 4: Firm 1's Reaction Function under Stochastic Inflationary and Competitive Click Fraud

The most striking thing about figure 4 is the possibility that for z_2 large enough, firm 1's optimal capacity is zero. Next we analyze firm 2's choice of z_2.

We can show that the shape of firm 2's profit function implies a unique maximum z^*. For $K_1 < n + z_2$,

$$\frac{\partial^2 \Pi_{2L}}{\partial z^2} = -2n(\Delta_2) \int_0^\infty \frac{f(r)dr}{(n+z_2+r)^3} - c''(z_2) \tag{42}$$

which is strictly negative for any z. For $K_1 \geq n + z_2$,

$$\frac{\partial^2 \Pi_{2L}}{\partial z^2} = n(\Delta_2) \left[\frac{f(K_1 - n - z_2)}{K^2} - 2 \int_{K_1-n-z_2}^\infty \frac{f(r)dr}{(n+z_2+r)^3} \right] - c''(z_2) \tag{43}$$

Which may be positive or negative, depending on $f()$. For $K_1 < n + z_2$, Π_{2L} is strictly concave, and for $K_1 \geq n + z_2$, Π_2 may be concave or convex (depending on K_1 and f). Figure 5 depicts the three possible shapes of firm 2's profit function.

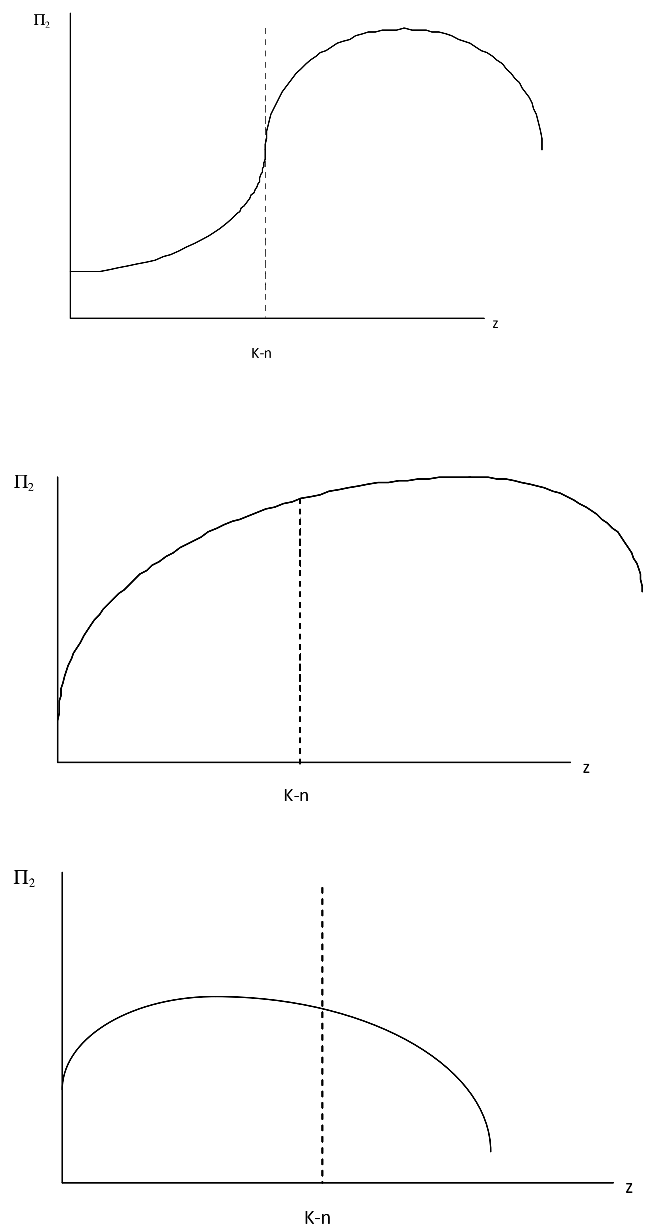
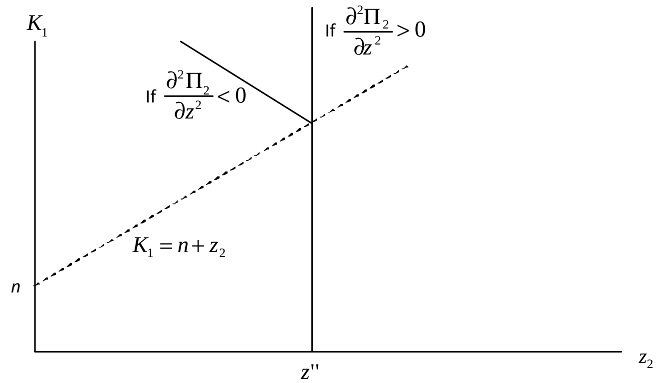


Figure 5: Firm 2's Profit Function under Stochastic Inflationary and Competitive Click Fraud (three possible shapes)

There are three relevant cases for firm 2's choice of z_2. First, it might be that the costs of committing click fraud are sufficiently high that $z_2^* = 0$. Second, Π_{2L} could be convex for $K_1 > z_2 + n$ or globally concave with a maximum in the range $K_1 < z_2 + n$; then the optimal choice is the z_2 that satisfies $\Delta_2 \int_0^\infty (n + z_2 + r)^{-2} f(r) dr = c'(z_2)$. Third, Π_{2L} may be globally concave with a maximum in the range $K_1 \geq z_2 + n$. In this final case, we can apply the implicit function theorem to $\frac{\partial \Pi_{2L}}{\partial z} = 0$ to find that $\frac{\partial K}{\partial z^*} < 0$. Figure 6 shows firm 2's reaction function $z_2(K_1)$.



$$z'' = z \text{ s.t. } \int_0^\infty \frac{\Delta_2 f(r)dr}{(n+r+z)^2} = c'(z)$$

Figure 6: Firm 2's Reaction Function under Stochastic Inflationary and Competitive Click Fraud

There are five qualitatively different equilibria in pure strategies in z and K, given b, $f(r)$, and $c(z)$. First, the click fraud cost may be sufficiently high that $z_2 = 0$. In this case $K_1 = K'$ in figure 4. Second, we might have Π_{2L} such that its global max is in the range $K_1 \geq z_2 + n$ and a crossing between $K_1(z_2)$ and $z_2(K_1)$ above $K_1 = z_2 + n$. We will then find a unique (z_2, K_1) combination where $z_2 \in (0, z'')$ and $K_1 > n + z''$. Third, we could have no crossing above $K_1 > n + z_2$ and $z'' > z'$. This would result in $z_2 = z''$ and $K_1 = 0$. Fourth, we could have no crossing above $K_1 > n + z_2$ and $z'' = z'$. Then we would get $z_2 = z''$ and $K_1 = K_1(z'')$. Finally, we could have no crossing above $K_1 > n + z_2$ and $z'' < z'$. This would yield $K_1 = 0$.