Essays on Mobile Advertising and Commerce

A dissertation presented

by

David Yu-Chung Chen

to

The Committee for the Ph.D. in Science, Technology and Management

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Science, Technology and Management

Harvard University
Cambridge, Massachusetts

May 2009

Professor David C. Parkes
Professor Peter A. Coles
Dissertation advisors

Professor Benjamin G. Edelman
Committee member

David Yu-Chung Chen
Ph.D. candidate

Essays on Mobile Advertising and Commerce

Abstract

The mobile industry holds the promise of increasing connectivity, productivity, and entertainment as mobile devices become ever more ubiquitous and powerful. Many of the market structures that enable mobile commerce are still under rapid development and afford us opportunities to ask new questions about market design. This thesis examines three such markets and mechanisms that drive advertising platforms and application stores for mobile commerce.

In the first essay, I analyze a mobile web advertising auction that employs a proportional allocation rule in which advertisements are shown with frequencies proportional to the bids. Proportional allocation is used to address the space constraints of the mobile environment and the accompanying ad fatigue. I show that the second-price

rule currently used in real-world auctions admits no pure-strategy Nash equilibrium. I propose the use of a first-price rule and prove the existence of a unique pure-strategy Nash equilibrium. This reverses the sponsored search result in which the second-price auction has pure-strategy Nash equilibria while the first-price auction does not. I also show that by tuning a single parameter in the allocation rule, the auctioneer can make trade-offs between revenue and efficiency.

In the second essay, I examine an optimize-and-dispatch scheme for delivering pay-per-impression advertisements in online and mobile advertising. Using traffic predictions based on historical traffic patterns, the platform provider seeks to allocate future inventory to advertisers such that commitments are fulfilled in expectation, and no single advertiser bears too much of the burden if actual traffic diverges from predicted traffic. I propose a maximum entropy approach and provide theoretical analysis and simulation to show how it accomplishes these goals.

In the final essay, I analyze the market used in mobile device application stores that enable the widespread distribution of third-party software. I characterize the conditions under which an application store seeking to maximize revenue should rank applications based on revenue versus download volume. I also present empirical data from the Apple iPhone App Store to illustrate some of the general features of this type of market.

Contents

Acknowledgments

Graduate school has been an incredible learning experience that would not have been possible without the help of many. First and foremost, I thank the members of my dissertation committee: David Parkes, Peter Coles, and Ben Edelman. David has generously met with me weekly for years, and has watched me mature as a researcher since the beginning of my program. He has been the most significant influence on how I view research, and his insights never cease to amaze me. Peter has spent many an afternoon helping me brainstorm and work through challenging problems. Ben has an incredible knack for seeing the big picture and helping me connect my work with real-world problems. Without their collective guidance, I could not have made it this far.

I am grateful for the countless faculty members at the Harvard Business School and the School of Engineering and Applied Sciences who have helped me along the way. A partial list could not exclude Carliss Baldwin, Yiling Chen, Tom Eisenmann, Lee Fleming, Andrei Hagiu, Marco Iansiti, HT Kung, and Jan Rivkin. I have also learned a tremendous amount from my graduate school friends and colleagues including Florin Constantin, Jacomo Corbo, Adam Juda, Richard Lai, Robin Lee, Merlina Manocaran, Sven Seuken, Marcin Strojwas, Allan Sumiyama, Mark Szigety, Jason

Woodard, Haoqi Zhang, and Feng Zhu. Members of the EconCS research group also deserve mention as they have provided crucial and insightful feedback on all of my work, often from the earliest stages.

These final two years have been incredibly productive and intense and have, I suspect, caused a great deal more stress for my partner Jane than even myself. For all her loving encouragement and patience, I will always be grateful. Finally, I am deeply thankful to my parents Sam and Theresa and my sister Anna for a lifetime of love and faith in me.

Chapter 1

Proportional Allocation Share Auctions

I analyze a mobile web advertising auction that employs a proportional allocation rule in which advertisements are shown with frequencies proportional to the bids. Proportional allocation is used to address the space constraints of the mobile environment and the accompanying ad fatigue.

I show that the second-price rule currently used in real-world auctions admits no pure-strategy Nash equilibrium. I propose the use of the first-price proportional allocation auction (FPP), and show the existence of a unique pure-strategy Nash equilibrium in a one-shot complete information game. In a dynamic game of incomplete information, I demonstrate in simulation that bids converge to this Nash equilibrium quickly. This reverses the sponsored search result in which the second-price auction has pure-strategy Nash equilibria while the first-price auction does not.

I also show that by tuning a single parameter in the FPP allocation rule, the auctioneer can make trade-offs between revenue and efficiency.

1.1 Introduction

Like their counterparts in the online world, many mobile web companies depend on advertising as a major source of revenue. A number of established players like Google and Yahoo, along with upstarts such as AdMob or Quattro Wireless, have begun acting as advertising intermediaries, connecting advertisers with mobile websites. Due to the highly-constrained environment, most mobile websites choose to only display a single advertisement on each page. The Generalized Second-Price slot auction (GSP), which has emerged as the standard auction for selling advertising in the online world, does not necessarily translate naturally into the mobile setting where there is only one slot for sale.

GSP maximizes expected revenue to the advertising platform and prices with a second-price rule. Research on sponsored search auctions[1] has uncovered a number of factors that influence expected revenue, including budget constraints (Abrams et al. (2007), Feldman et al. (2007)), the need for exploration and exploitation (Gonen and Pavlov (2007), Pandey and Olston (2006)), and ad fatigue (Abrams and Vee (2007)). While all these factors remain important in mobile web advertising, the issue of ad fatigue is especially relevant for a number of reasons. First, users can quickly tire of seeing the same ads if there is only one slot per page. Fatigue can be a concern of similar severity with online contextual ads (e.g. Google AdSense), but another characteristic of the mobile ad environment is that contextual matching is

[1] For a good overview of results from the sponsored search literature, see Lahaie et al. (2007).

difficult due to the fact that mobile webpages are on average shorter than their online counterparts. Addressing ad fatigue as a first-order concern helps protect against imprecise contextual matching.

In practice, mobile advertising platforms have approached this new setting by using a proportional share scheme, where bidders receive a share of the impressions proportional to their bid. The share is a fractional allocation representing the probability that a particular ad will be shown when mobile web surfers arrive at the webpage. Of course, only one ad is actually shown for any given pageview, but over longer periods, the fraction of time that a particular ad is shown converges to the bidder's allocated share.

Real-world auctions have generally used a second-price rule, but I show in this paper that there are no pure-strategy Nash equilibria with the second-price rule. I propose the use of the first-price proportional allocation auction (FPP) and prove that there is a unique pure-strategy Nash equilibrium in a one-shot complete information game. In a dynamic game of incomplete information, I show in simulation that bids converge to this Nash equilibrium quickly. Note that this result is the opposite of the sponsored search result where GSP has Nash equilibria and the first-price auction has bidding cycles (Edelman and Ostrovsky (2007)). I will consider this further in the Discussion (Section 1.6).

I also show that by tuning a single parameter in the FPP allocation rule, the auctioneer can make trade-offs between revenue and efficiency. I compare FPP with the maximally-efficient Vickrey-Clarke-Groves (VCG) auction, and find in simulation that FPP efficiency is quite efficient relative to VCG. In addition, the FPP allocation rule has the benefit of being transparent to the bidder, whereas the VCG allocation rule can be critiqued for being difficult for bidders to understand.

1.1.1 Related Literature: Proportional Share Mechanisms

Kelly (1997) considers the use of markets to address the problem of network congestion. In particular, network users bid on resources in a mechanism where each bid represents a users's willingness to pay for total network usage. The link manager computes a clearing price so that every user pays the same unit price, which has the effect of allocating resources in proportion to bids. For example, suppose that 2 users submit bids of 2 and 3, and the link has capacity 20. The link manager computes a per unit price of $(2+3)/20 = 0.25$, implying that bidder 1 receives $2/0.25 = 8$ in capacity and bidder 2 receives $3/0.25 = 12$ in capacity. Note that the allocations of 8 and 12 are proportional to the original bids 2 and 3.

Kelly proves that when bidders are price-takers with strictly concave, increasing utility, there exists a competitive equilibrium (Mas-Colell et al. (1995)) where bidders maximize their payoffs and the network sets a market-clearing per-unit price. He also shows that the resulting allocation maximizes aggregate user utility. Note that the auctioneer does not price differentiate as per-unit prices are uniform across bidders. Non-price-discrimination results in "aggregate feedback," which is considered a feature in the context of communication networks, as it allows for scalability (Shenker (1990)). Kelly's non-game-theoretic model differs fundamentally from my first-price proportional share auction in that my auction price differentiates and my bidders are strategic. The concept of competitive equilibrium does not apply in my setting since competitive equilibrium requires a market-clearing price and my bidders pay different prices.

Kelly's work is the basis of much further research in proportional allocation mechanisms, most notably Johari and Tsitsiklis (2004), Johari and Tsitsiklis (2007), and Sanghavi and Hajek (2004). In Johari and Tsitsiklis (2004), the authors relax Kelly's assumption that bidders are price-takers, and allow bidders to anticipate their effect on the price set by the network. They find that aggregate utility at the Nash equilibrium is no worse than 75% of the maximal aggregate utility. Note that their work still requires one network-wide unit price for all bidders, resulting in a different Nash equilibrium from the one I find in FPP.

In Johari and Tsitsiklis (2007), the authors further relax the single price assumption and consider "scalar-parametrized" mechanisms, which restrict users to one-dimensional bids. This class of mechanisms includes the proportional share allocation auctions I study. They present the VCG mechanism applied to this setting, verifying that full efficiency is achievable at Nash equilibrium. I will show in Section 1.5.1 that FPP is quite efficient relative to VCG.

1.2 First-Price Rule

In this section, I present a complete information model of FPP, following the sponsored search literature (Edelman et al. (2007)). I assume that all bidder values are common knowledge since over time, advertisers can learn each others' values by experimenting with their bids.

I find an upper bound on bids in equilibrium as a function of bidder value by analyzing what a utility-maximizing bidder would do in equilibrium. This upper bound represents the minimum amount by which any bidder will shade his bid; such shading is expected in any first-price auction.

Next I prove the existence of exactly one unique pure-strategy Nash equilibrium and provide a closed-form representation of a symmetric Nash equilibrium for the case when all bidders have the same value. Bids in this symmetric Nash equilibrium become arbitrarily close to the previously calculated upper bound as the number of bidders increases. Analytically computing the Nash equilibrium with arbitrary bidder values appears to be challenging, but simulations show that Nash equilibria in the general case are quite close to the the symmetric Nash equilibrium.

1.2.1 First-Price Model

There are N bidders in a one-shot, private value, complete information auction. Bidder i has value v_i per click and submits a positive bid per click b_i. Bidder i's allocation is proportional to his bid:

$$x_i(b_i) = \frac{b_i^\alpha}{\sum_j b_j^\alpha} = \frac{b_i^\alpha}{b_i^\alpha + b_{-i}}$$

where $\alpha > 0$, and $b_{-i} = \sum_{j \neq i} b_j^\alpha$ is the sum of all other bidders' bids (after taking the bids to the power α). $x_i(b_i)$ represents the proportion of the total impressions given to to bidder i. It is easy to verify that $\sum_i x_i(b_i) = 1$, meaning that all impressions are allocated.

I am interested in bidder clickthrough rate (CTR) since bidder utility is a function of CTR. I model bidder i's CTR by multiplying $x_i(b_i)$ by $q_i \in [0, 1]$, which is bidder i's quality score. The CTR can be interpreted as the probability of a click given a quality of 1. In Section 1.4, I also allow for CTRs to vary based on ad fatigue, meaning that bidders with higher share allocations will have lower CTRs, all else equal.

In practice, for $\alpha > 1$, I add the restriction $b_i > 1$ to maintain the interpretation that using a higher exponent α is meant to weight greater bids more heavily. I will refer to the case of $\alpha = 1$ as **linear proportional allocation** and $\alpha = 2$ as **squared proportional allocation**. Note that α need not be an integer.

Consider a first-price rule where bidder i's per-click price is his bid: $p_i(b_i) = b_i$.

I assume that utility is quasi-linear:

$$\begin{aligned} \text{utility} &= \text{CTR} * (\text{value} - \text{price}) \\ u_i(b_i) &= [x_i(b_i)q_i](v_i - p_i(b_i)) \\ &= \frac{b_i^\alpha}{b_i^\alpha + b_{-i}} q_i(v_i - b_i) \end{aligned}$$

1.2.2 Best Response Strategies, First-Price Rule

In this section, I establish some general observations about how bidders will bid in equilibrium in FPP.

Bidder i picks bid b_i to maximize utility u_i, so the first-order necessary condition for a maximizing interior solution gives the best response correspondence:

$$\begin{aligned} \frac{\partial u_i}{\partial b_i} = \frac{q_i b_i^{\alpha-1}\left(-b_i^{\alpha+1} - b_{-i}(\alpha+1)b_i + b_{-i}\alpha v_i\right)}{(b_i^\alpha + b_{-i})^2} &= 0 \\ \Leftrightarrow -b_i^{\alpha+1} - b_{-i}(\alpha+1)b_i + b_{-i}\alpha v_i &= 0 \end{aligned} \tag{1.1}$$

Note that in Equation 1.1, q_i drops out completely, so the bidder's best response is not dependent on his ad's quality.

In the following lemma, I show that the best response *correspondence* is in fact a best response *function*; that is, Equation 1.1 has exactly one root in the relevant range $[0, v_i]$. I am only interested in the range $[0, v_i]$ since this is a first-price auction in which bids above v_i yield negative utility and bids below 0 yield no allocation.

Lemma 1. *For bidder i, given any vector of others' bids $\mathbf{b_{-i}}$, where $\mathbf{b_{-i}} = (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_N)$, there exists exactly one best-responding bid b_i^* in $[0, v_i]$.*

Proof. I am interested in showing that $\partial u_i / \partial b_i$ has exactly one root in $[0, v_i]$. Let $g(b_i) = -b_i^{\alpha+1} - b_{-i}(\alpha+1)b_i + b_{-i}\alpha v_i$. Then, $\partial u_i/\partial b_i$ will have a root wherever g has a root unless the denominator $(b_i^\alpha + b_{-i})^2$ is 0.

I now show that g has exactly one root in $[0, v_i]$. By the Descartes Sign Rule, the maximum possible number of positive real roots of a polynomial is equal to the number of sign changes starting from the sign of the coefficient of the highest power and proceeding to the lowest power.

In this case, g has one sign change, so it has a maximum of 1 possible positive root. Thus, $\partial u_i/\partial b_i$ also has a maximum of 1 possible positive root. Next, I show that $\partial u_i/\partial b_i$ always has at least 1 positive root.

For any small positive ϵ, $\left.\frac{\partial u_i}{\partial b_i}\right|_{b_i=\epsilon} > 0$. This is because the expression is dominated by the positive $b_{-i}\alpha v_i$ term while the negative terms all include the small b_i. Note that I cannot simply evaluate $\left.\frac{\partial u_i}{\partial b_i}\right|_{b_i=0}$ since $\frac{\partial u_i}{\partial b_i}$ has a hole at $b_i = 0$.

Next, I find that

$$\left.\frac{\partial u_i}{\partial b_i}\right|_{b_i=v_i} = -\frac{q_i v_i^{\alpha-1}\left(v_i^{\alpha+1} + b_{-i}(\alpha+1)v_i\right)}{(v_i^\alpha + b_{-i})^2} < 0$$

since every term is positive.

$\partial u_i/\partial b_i$ is a rational function (the ratio of two polynomial functions) and is thus continuous at each point in its domain. Since the sign of $\partial u_i/\partial b_i$ switches as b_i goes from small ϵ to v_i, the intermediate value theorem implies that there must be at least one positive root. Combining this fact with the observation from Descartes' Sign Rule that $\partial u_i/\partial b_i$ has at most 1 positive root, I conclude that $\partial u_i/\partial b_i$ must have exactly 1 positive root, and thus there is exactly one root in $[0, v_i]$. □

By the following theorem, I establish that any best-responding bid is in the interval $\left(0, \frac{\alpha}{\alpha+1}v_i\right)$.

Theorem 1.2.1. *In the FPP auction, a utility-maximizing bidder i will always pick bid $b_i > 0$ such that*

$$b_i < \frac{\alpha}{\alpha+1}v_i$$

for any $\alpha > 0$ and for any $b_{-i} > 0$. Bidder i will shade down his bid from true value v_i by a factor of $\frac{\alpha}{\alpha+1}$ or more regardless of the other bids.

Proof. Equation 1.1 is a polynomial in b_i of degree $(\alpha + 1)$, so it can only be solved explicitly for small integral values of α. However, with some rearranging I arrive at the result. I rewrite Equation 1.1 as:

$$-b_i^{\alpha+1} - b_{-i}[(\alpha+1)b_i - \alpha v_i] = 0 \tag{1.2}$$

Since $-b_i^{\alpha+1} < 0$ and $-b_{-i} < 0$, the only way (1.2) can be feasible is if the expression in brackets is negative, namely:

$$(\alpha+1)b_i - \alpha v_i < 0$$

$$\Leftrightarrow b_i < \frac{\alpha}{\alpha+1}v_i$$

$\square$

An implication of this theorem is that as α increases, the upper bound to each bidder's best response increases. Intuitively, larger values of α increase the marginal value of submitting a higher bid since a higher bid receives an even larger allocation with a higher α.

To graphically illustrate this theorem, I explicitly solve for b_i in (1.1) for small integral values of α. For example, if $\alpha = 1$, the quadratic formula yields one positive solution:

$$b_i = -b_{-i} + \sqrt{b_{-i}^2 + v_i b_{-i}}$$

If $\alpha = 2$, the cubic formula yields one positive solution:

$$b_i = (b_{-i}v_i + \sqrt{b_{-i}^3 + v_i^2 b_{-i}^2})^{\frac{1}{3}} - (-b_{-i}v_i + \sqrt{b_{-i}^3 + v_i^2 b_{-i}^2})^{\frac{1}{3}}$$

In Figure 1.1, I show the bidder's optimal bid as a function of the sum of the other exponentiated bids for the case when bidder value v_i is 6. I examine exponents α of 1 (linear proportional allocation) and 2 (squared proportional allocation). From the graph, for $\alpha = 1$, bidder i's optimal bid approaches the asymptote $\frac{1}{2}v_i = 3$ from below and for $\alpha = 2$, the optimal bid approaches the asymptote $\frac{2}{3}v_i = 4$. There are no values of b_{-i} for which the bidder will ever bid above $\frac{\alpha}{\alpha+1}v_i$. Re-stated, this means that with $\alpha = 1$, no bidder will ever bid greater than half his value and with $\alpha = 2$, no bidder will ever bid greater than two-thirds his value.

In Figure 1.2, I examine the optimal bid for the case of $\alpha = 1$, varying v_i and b_{-i}. The level curve that is found by intersecting the graph in Figure 1.2 with the plane $v_i = 6$ corresponds exactly with the dotted line for $\alpha = 1$ in Figure 1.1. The

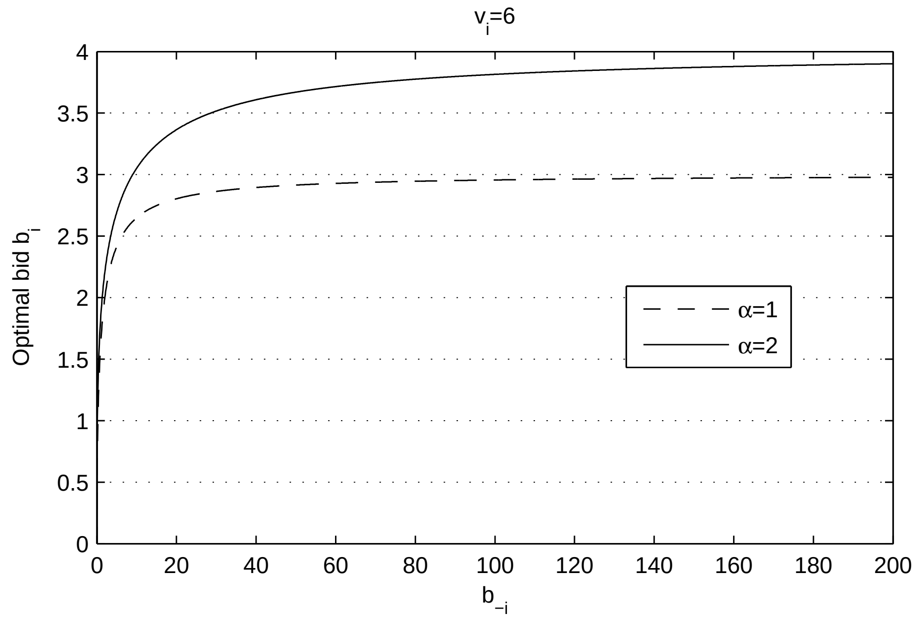


Figure 1.1: As b_{-i} increases, the optimal bid for bidder i approaches $\frac{1}{2}v_i$ for $\alpha = 1$, and $\frac{2}{3}v_i$ for $\alpha = 2$.

analogous graph for $\alpha = 2$ is similar in shape to Figure 1.2, and approaches the higher asymptote of $\dfrac{2}{3}v_i$. For large B, observe that the level curve that is found by intersecting the graph with the plane $b_{-i} = B$ approaches the line $b_i = \dfrac{1}{2}v_i$.

In simulation, I have found that the $\dfrac{\alpha}{\alpha+1}v_i$ asymptote is reached very quickly with as few as 5 competing bidders with comparable values. As the exponent α increases, the shading factor $\dfrac{\alpha}{\alpha+1}$ becomes arbitrarily close to 1, meaning that with a high exponent, bidders are effectively incented to bid their true values. Of course, this is unlikely to be desirable because large exponents tend to heavily favor higher bidders, effectively cutting out lower bidders.

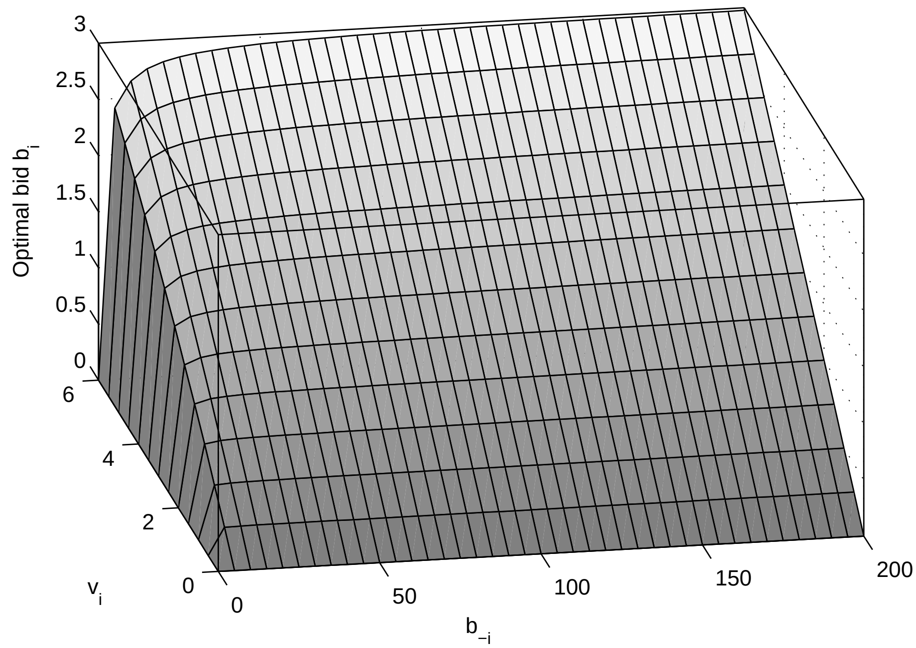


Figure 1.2: For $\alpha = 1$, as b_{-i} increases, the optimal bid for bidder i approaches $\frac{1}{2}v_i$.

1.2.3 Nash Equilibrium, First-Price

I begin this section by proving the existence of a unique Nash equilibrium and then giving a closed-form description of the Nash equilibrium in which all bidders have the same value.

After that, I will show two results about strategies in the Nash equilibrium. First, bids are ordered by value, i.e., bidders with higher value bid higher. Second, if I denote the ratio of bid to value as the **shading factor**, I find that shading factors are in reverse order from value, i.e. bidders with higher value shade more.

Theorem 1.2.2. *For all v_i, there exists one unique pure-strategy Nash equilibrium* *in the FPP auction if a minimum bid is enforced of no more than M for all i, where*

$$M = \frac{(N-1)\alpha v_i}{(N-1)\alpha + N}$$

Proof. I begin by applying Brouwer's fixed point theorem to show the existence of a pure-strategy Nash equilibrium. Brouwer's fixed point theorem is valid because I showed in Lemma 1 that a bidder's best response is always a single point rather than a set of points.

Until now, I have referred loosely to the "best response function" as the solution in b_i to

$$-b_i^{\alpha+1} - b_{-i}(\alpha+1)b_i + b_{-i}\alpha v_i = 0$$

I will now refer to this as the **individual best response function** and to the N-tuple consisting of N individual best response functions as the **best response function**.

Brouwer's fixed point theorem states that every continuous function from the closed unit ball D^N to itself has at least one fixed point.

For this proof, I assume that in practice there is a fixed minimum bid m (for example, one cent). I will show that as long as m is reasonably small (less than $v_i/3$), there will be one unique Nash equilibrium.

I need to show that the best response function maps from the space $R = [m, v_1] \times \ldots \times [m, v_N]$ to itself and is continuous. I prove these 2 properties in the following 2 lemmas.

Lemma 2. *The best response function maps from the space $R = [m, v_1] \times \ldots \times [m, v_N]$ to itself.*

Proof. The minimum bid is m by assumption, and the maximum bid for bidder i is v_i

in order to yield non-negative utility; thus, the domain the of best response function is R. I need to show that the image of the best response function is a subset of R. The v_i upper bound of the interval follows from Theorem 1.2.1 since every bidder i will always bid below $\frac{\alpha}{\alpha+1}v_i$, which is less than v_i for any $\alpha > 0$.

Now I suppose that every bidder bids the minimum m and I characterize the conditions under which best responses will be at least m. If all bidders bid the minimum m, then for any bidder i, $b_{-i} = \sum_{j \neq i} m^\alpha = (N-1)m^\alpha$. Bidder i's best response is the solution in b_i to $\partial u_i / \partial b_i = 0$, which simplifies in this case to the following:

$$-b_i^{\alpha+1} + (N-1)m^\alpha[-(\alpha+1)b_i + \alpha v_i] = 0 \tag{1.3}$$

By Lemma 1, $\partial u_i / \partial b_i$ has exactly one root in $[0, v_i]$. To the left of this root, $\partial u_i / \partial b_i > 0$ since the root represents the point of maximum utility. This means that if I substitute m for b_i in $\partial u_i / \partial b_i$ and find that the expression is at least 0, it must be the case that the best response is at least m:

$$\begin{aligned}
-m^{\alpha+1} + (N-1)m^\alpha[-(\alpha+1)m + \alpha v_i] &\geq 0 \\
\Leftrightarrow -m + (N-1)[-(\alpha+1)m + \alpha v_i] &\geq 0 \\
\Leftrightarrow m \leq \tfrac{(N-1)\alpha v_i}{(N-1)\alpha+N} &= M \qquad (1.4)
\end{aligned}$$

Expression 1.4 establishes that as long as the minimum bid is less than or equal to M, the best response function will map each bidder's initial minimum bid to a higher bid, falling in the set R.

In practice, N is always at least 2 (since there are at least 2 bidders in the auction) and α is at least 1, so (1.4) becomes

$$m \leq \frac{v_i}{3}$$

Larger values of N and α allow the use of even higher minimum bids.

Next, I show that the best response function is increasing in b_{-i}; this means that since I have established that the best response to minimum bidding is at least the minimum bid, best responses will always be at least the minimum bid.

Consider how bidder i would best-respond given two different vectors of others' bids, $\mathbf{b_{-i}}$ and $\mathbf{c_{-i}}$ such that $b_{-i} = c_{-i} + \epsilon$ where $b_{-i} = \sum_{j \in \mathbf{b_{-i}}} b_j^\alpha$ and c_{-i} is defined similarly. I want to show that bidder i's best response b_i^* to vector $\mathbf{b_{-i}}$ is greater than his best response c_i^* to $\mathbf{c_{-i}}$.

If we assume that it is not, then $b_i^* < c_i^*$, so there exists a $\delta > 0$ such that $b_i^* + \delta = c_i^*$. Bidder i's best response function in response to b_{-i} is:

$$-(b_i^*)^{\alpha+1} + (c_{-i} + \epsilon)[-(\alpha+1)b_i + \alpha v_i] = 0 \tag{1.5}$$

Bidder i's best response function in response to c_{-i} is:

$$-(b_i^* + \delta)^{\alpha+1} + c_{-i}[-(\alpha+1)(b_i + \delta) + \alpha v_i] = 0 \tag{1.6}$$

If Equation 1.6 is subtracted from Equation 1.5, the left-hand side is greater than 0, yielding a contradiction. Thus the best response function must be increasing in b_{-i}. □

Lemma 3. *The best response function is continuous.*

Proof. The graph of Figure 1.1 shows that the **individual best response function** f_i is continuous. In this proof, I show that the **best response function** f is continuous.

Given strategy profile $\mathbf{c}$ and any $\epsilon > 0$, I need to show that $d(f(\mathbf{b}), f(\mathbf{c})) < \epsilon$ whenever $d(\mathbf{b}, \mathbf{c}) < \delta$ for some $\delta > 0$, where I use the ∞-norm distance metric $d(\mathbf{x}, \mathbf{y}) = \max_i(|x_i - y_i|)$.

I write $\mathbf{b}$ in terms of $\mathbf{c}$ by defining γ_i: $\forall i, \quad b_i = c_i + \gamma_i$. This means that $b_{-i} = \sum_{j \neq i} b_j = c_{-i} + \sum_{j \neq i} \gamma_j$.

$$
\begin{aligned}
d(f(\mathbf{b}), f(\mathbf{c})) &= \max_i \left(\left| f_i(b_{-i}) - f_i(c_{-i}) \right| \right) \\
&= \max_i \left(\left| f_i(c_{-i} + \sum_{j \neq i} \gamma_j) - f_i(c_{-i}) \right| \right) \\
&\leq \max_i \left(\left| \sum_{j \neq i} \gamma_j \right| \right) \qquad (1.7) \\
&\leq (N-1) * \max_i(|\gamma_i|) \\
&= (N-1) * d(\mathbf{b}, \mathbf{c})
\end{aligned}
$$

Equation 1.7 follows from the fact that the slope of the individual best response function is between 0 and 1 when a minimum bid is imposed; in other words, $\mathbf{b}$ and $\mathbf{c}$ are to the right of the curve in Figure 1.1. Given any $\epsilon > 0$, set $\delta = \epsilon/(N-1)$ and continuity follows. $\square$

Since continuity and being a fixed point are invariant under homeomorphisms, Brouwer's fixed point theorem also applies if the domain is a set homeomorphic to the closed unit ball. In particular, the closed unit ball is homeomorphic to the hyperrectangle $R = [m, v_1] \times \ldots \times [m, v_N]$. Lemmas 2 and 3 demonstrate that the preconditions of Brouwer's fixed point theorem are met, so Brouwer's fixed point theorem guarantees the existence of a fixed point in this auction. Since a Nash equilibrium is by definition a fixed point of the best response function, I have shown the existence of a Nash equilibrium.

Next, I demonstrate that there can only be 1 unique Nash equilibrium. Assume to the contrary that there are 2 Nash equilibria $\mathbf{x} = (x_1, \ldots, x_N)$ and $\mathbf{y} = (y_1, \ldots, y_N)$. I write $\mathbf{x}$ in terms of $\mathbf{y}$ by defining ϵ_i as follows:

$$\forall i, \;\; x_i^\alpha = y_i^\alpha + \epsilon_i$$

This implies the following:

$$\begin{aligned} x_i &= (y_i^\alpha + \epsilon_i)^{\frac{1}{\alpha}} \\ x_i^{\alpha+1} &= (y_i^\alpha + \epsilon_i)^{\frac{\alpha+1}{\alpha}} \end{aligned}$$

I also write x_{-i} and y_{-i} as the following:

$$\begin{aligned} x_{-i} &= \sum_{j \neq i} x_j^\alpha = \sum_{j \neq i} (y_j^\alpha + \epsilon_j) \\ y_{-i} &= \sum_{j \neq i} y_j^\alpha \end{aligned}$$

By the definition of Nash equilibrium, for all i, x_i and y_i solve the best response function:

$$-x_i^{\alpha+1} - x_{-i}\left[(\alpha+1)x_i - \alpha v_i\right] = 0 \tag{1.8}$$

$$-y_i^{\alpha+1} - y_{-i}\left[(\alpha+1)y_i - \alpha v_i\right] = 0 \tag{1.9}$$

With some substituting, I now rewrite (1.8) as (1.10) and (1.9) as (1.11):

$$-(y_i^\alpha + \epsilon_i)^{\frac{\alpha+1}{\alpha}} - \sum_{j \neq i}(y_j^\alpha + \epsilon_j)\left[(\alpha+1)(y_i^\alpha + \epsilon_i)^{\frac{1}{\alpha}} - \alpha v_i\right] = 0 \tag{1.10}$$

$$-y_i^{\alpha+1} - \sum_{j \neq i} y_j^\alpha \left[(\alpha+1)y_i - \alpha v_i\right] = 0 \tag{1.11}$$

If I subtract Equation 1.11 from Equation 1.10, the left-hand side is less than 0 if any of the ϵ_i's are not equal to 0, yielding a contradiction. The only way to avoid this contradiction is if every $\epsilon_i = 0$, meaning that there is one unique Nash equilibrium. $\square$

A possible concern with Theorem 1.2.2 is that it is a nonconstructive proof, so it is not known whether bidders will actually find this equilibrium in practice. In Section 1.2.2, I present simulations showing that in a repeated FPP game, bidders do indeed reach this equilibrium with as few as 2 to 3 updates in their bids.

Next, I present a closed-form representation of the Nash equilibrium for the special case when all bidders have the same value.

Theorem 1.2.3. *If there are N bidders with identical value v, then in the Nash equilibrium, all bidders bid*

$$b = \frac{(N-1)\alpha}{(N-1)\alpha + N} v$$

Proof. Besides the bidder in question, there are $(N-1)$ identical bidders, each with value v. Thus,

$$b_{-i} = \sum_{j \neq i} b_j^\alpha = (N-1)b^\alpha$$

Substitute this into Equation 1.1:

$$-b^{\alpha+1} - (N-1)b^\alpha(\alpha+1)b + (N-1)b^\alpha \alpha v = 0$$

$$\Leftrightarrow b + b(N-1)(\alpha+1) - \alpha(N-1)v = 0$$

$$\Leftrightarrow b = \tfrac{(N-1)\alpha}{(N-1)\alpha+N} v$$

Note that as the number of bidders N gets large, the optimal bid b approaches $\dfrac{\alpha}{\alpha+1}v$, the bound shown in Theorem 1.2.1. $\square$

Next, I prove that in the Nash equilibrium, bids are ordered by value and that bidders with higher value shade more. In Figure 1.3, I give an example of bids in a Nash equilibrium with linear proportional allocation ($\alpha = 1$) where 5 bidders have value 8, 6, 5, 4, and 2. Notice that the equilibrium bids are ordered by value and the shading factors are in reverse order from value. Also note that the shading factors are all less than $\frac{\alpha}{\alpha+1} = 0.5$, the upper bound from Theorem 1.2.1.

Value	Nash bid	Shading factor
8	3.288	0.411
6	2.594	0.432
5	2.216	0.443
4	1.818	0.454
2	0.954	0.477

Figure 1.3: Nash equilibrium bids and shading factors for 5 bidders with $\alpha = 1$

Theorem 1.2.4. *In the FPP Nash equilibrium, bids are ordered by value. That is, for N bidders, given values $\{v_1, \ldots, v_N\}$, Nash bid profile $\{b_1, \ldots, b_N\}$, and any $\alpha > 1$,*

$$v_i > v_j \Leftrightarrow b_i > b_j \ \ \forall i \neq j$$

Proof. Without loss of generality, assume that bidder indices are in order by value, i.e. $v_1 > v_2 > \ldots > v_N$. Assume that $\{b_1, \ldots, b_N\}$ is the Nash equilibrium bid profile in which bids for bidders m and n are not in order with $m > n$, but all other bids are in order. That is, $b_m < b_n$ and $b_i > b_{i+1}$ for all i not equal to m or n.

By the definition of Nash equilibrium, for all i, b_i solves the best response function:

$$-b_i^{\alpha+1} + b_{-i} \left[\alpha v_i - (\alpha + 1)b_i\right] = 0$$

$$\text{Let } B = \sum_{j \neq m,n} b_j^\alpha$$

$$\epsilon v_m = v_n \text{ with } 0 < \epsilon < 1$$

$$b_m = \delta b_n \text{ with } 0 < \delta < 1$$

The best response function for bidder m is:

$$-b_m^{\alpha+1} + (b_n^\alpha + B)(\alpha v_m - (\alpha+1)b_m) = 0$$

Substitute δb_n for b_m:

$$-(\delta b_n)^{\alpha+1} + (b_n^\alpha + B)(\alpha v_m - (\alpha+1)\delta b_n) = 0 \tag{1.12}$$

The best response function for bidder n is:

$$-b_n^{\alpha+1} + (b_m^\alpha + B)(\alpha v_n - (\alpha+1)b_n) = 0$$

Substitute δb_n for b_m and ϵv_m for v_n:

$$-b_n^{\alpha+1} + ((\delta b_n)^\alpha + B)(\alpha \epsilon v_m - (\alpha+1)b_n) = 0 \tag{1.13}$$

If Equation 1.13 is subtracted from Equation 1.12, the left-hand side is greater than 0, yielding a contradiction. Thus, the original assumption that b_m and b_n are not in order must be impossible. Since m and n were picked arbitrarily, it must be the case that all bids are in order. □

Theorem 1.2.5. *In the FPP Nash equilibrium, shading factors are in reverse order from value. That is, for N bidders, and any $\alpha > 1$, if $\{b_1, \ldots, b_N\}$ is the Nash equilibrium bid profile with shading factor c_i defined as $c_i = \frac{b_i}{v_i}$ for all i, then*

$$v_i > v_j \Leftrightarrow c_i < c_j \ \ \forall i \neq j$$

Proof. Without loss of generality, assume that bidder indices are in order by value, i.e. $v_1 > v_2 > \ldots > v_N$. Assume that $\{b_1, \ldots, b_N\}$ is the Nash equilibrium bid profile in which shading factors for bidders m and n are in order with $m > n$, but all other shading factors are in reverse order. That is, $c_m > c_n$ and $c_i < c_{i+1}$ for all i not equal to m or n.

By the definition of Nash equilibrium, for all i, b_i solves the best response function:

$$-b_i^{\alpha+1} - b_{-i}(\alpha+1)b_i + b_{-i}\alpha v_i = 0$$

$$\text{Let } B = \sum_{j \neq m,n} b_j^{\alpha}$$

$$\delta c_m = c_n \text{ with } 0 < \delta < 1$$

$$\epsilon v_m = v_n \text{ with } 0 < \epsilon < 1$$

By definition, $b_m = c_m v_m$ and $b_n = c_n v_n$, so $b_n = \delta\epsilon c_m v_m$. I introduce a new variable $\gamma = \delta\epsilon$ with $0 < \gamma < 1$, so $b_n = \gamma c_m v_m$. The best response function for bidder m is:

$$-b_m^{\alpha+1} + (b_n^{\alpha} + B)(\alpha v_m - (\alpha+1)b_m) = 0$$

Substituting $c_m v_m$ for b_m and $\gamma c_m v_m$ for b_n:

$$-(c_m v_m)^{\alpha+1} + ((\gamma c_m v_m)^{\alpha} + B)(\alpha v_m - (\alpha+1)c_m v_m) = 0 \tag{1.14}$$

The best response function for bidder n is:

$$-b_n^{\alpha+1} + (b_m^{\alpha} + B)(\alpha v_n - (\alpha+1)b_n) = 0$$

Make the same substitutions as above and substitute ϵv_m for v_n:

$$-(\gamma c_m v_m)^{\alpha+1} + ((c_m v_m)^{\alpha} + B)(\alpha\epsilon v_m - (\alpha+1)\gamma c_m v_m) = 0 \tag{1.15}$$

If Equation 1.15 is subtracted from Equation 1.14, the left-hand side is less than 0, yielding a contradiction. Thus, the original assumption that c_m and c_n are in order must be impossible. Since m and n were picked arbitrarily, it must be the case that shading factors are in reverse order. $\square$

Note the generality of the previous two theorems, as they hold independent of the choice of α.

1.3 Second-Price Rule

The main result of this section is that there are no pure-strategy Nash equilibria when the second-price rule is used in the proportional share auction. I briefly outline the formal model before proving this result.

1.3.1 Second-Price Model

Values, bids, and the allocation rule are all identical to the First-Price Model. Bids are normalized by quality (CTR): $b'_i = b_i * q_i$. The ranking function σ maps from bidder index i to rank with smaller-valued ranks indicating better ranking. Bidder payments per click correspond to a second-price rule: normally, a bidder pays a price equal to the bid of the bidder ranked directly below him. In case of a tie, the bidders that tie each pay their bid. The lowest-ranked bidder pays a reserve price p_r. To be precise,

$$p_i = \begin{cases} b'_i & \text{if } i\text{'s bid ties with any other bidder} \\ p_r & \text{if } \sigma(i) = N \\ b'_{\sigma^{-1}(\sigma(i)+1)} & \text{otherwise} \end{cases}$$

1.3.2 Nash Equilibrium, Second-Price

In this section, I prove that with a second-price rule, the proportional share auction has no Nash equilibrium. I begin with a lemma demonstrating that if there were a Nash equilibrium, there could be no gaps in bids. I then prove the main result.

Lemma 4. *In a Nash equilibrium for the second-price proportional share auction, there can be no gaps in bids. Thus, all bidders must submit the same bid.*

Proof. Assume the existence of a Nash equilibrium with bid vector $\mathbf{b}$. Consider two bidders i and j with a bid gap where i is ranked directly above j: $\sigma(i) = \sigma(j) - 1$. Bidder i is bidding ϵ more than j: $b_i = b_j + \epsilon$, with $\epsilon > 0$. If bidder j increases his bid less than ϵ, his allocation $x_i(b_j)$ will increase while his price remains unchanged. This means that his utility u_j rises, contradicting the initial statement that $\mathbf{b}$ is a Nash equilibrium. Thus, it must be the case that in a Nash equilibrium, there cannot be a gap in bids between any two bidders. □

Theorem 1.3.1. *There are no pure-strategy Nash equilibria in the second-price proportional share auction.*

Proof. By Lemma 4, in a Nash equilibrium, there can be no gaps and thus all bidders must submit the same bid. Assume the existence of a Nash equilibrium where all bidders bid b. Now if any bidder i submits a bid higher than b, then his allocation $x_i(b_i)$ will rise while his price remains unchanged. Thus his utility u_i rises, meaning that $\mathbf{b}$ could not have been a Nash equilibrium. Since Lemma 4 says that any Nash equilibrium must have all bidders submitting the same bid and since I have shown that any bidder can then improve his utility from his "equilibrium" utility, it must be the case that there are no pure-strategy Nash equilibria. $\square$

1.4 Ad Fatigue Extension

A natural next question to ask is whether the proportional allocation framework can accommodate other concerns such as ad fatigue. I model ad fatigue by decreasing an ad's CTR over time as users become fatigued with the ad. In the one-shot setting, I can use a scaling factor of between 0 and 1, which decreases as the bidder's share of the impressions grows. The exact fatigue scaling function $f(x)$ can be interpreted as the average decrease in CTR over a time period.

More precisely, under the first-price rule, bidder utility is now modeled as:

$$\begin{aligned} u_i(b_i) &= x_i(b_i)q_i * f(x_i(b_i))(v_i - b_i) \\ &= \frac{b_i^\alpha}{b_i^\alpha + b_{-i}} q_i * f\left(\frac{b_i^\alpha}{b_i^\alpha + b_{-i}}\right) * (v_i - b_i) \end{aligned}$$

where f is the fatigue function.

I assume that f can be any non-negative and monotone decreasing function. In the one-shot game, f is common knowledge to all the bidders. However, in a repeated setting, bidders do not need to know the specific form of f as they can learn about it by bidding in the auction over time. For example, a bidder might initially submit a bid that shades down his value by some small amount. He gets allocated a large share, but the accompanying ad fatigue is also large. The bidder then experiments with dropping his bid a little more. His share is reduced by a little, but ad fatigue is also reduced and his price drops as well. The bidder does not have information about his particular allocation or ad fatigue–all he experiences is a rise in clicks. The bidder can keep revising his bid downwards until he reaches the Nash equilibrium described later in this section.

1.4.1 Best Response Strategies, Ad Fatigue Extension

Ad fatigue analogues to Theorems 1.2.2, 1.2.4, and 1.2.5 can be shown with similar analyses to the original proofs; outlines of these proofs are left for the appendix. In this section, I focus on proving that the bound from Theorem 1.2.1 still holds with the ad fatigue extension, and examine two concrete ad fatigue functions.

Theorem 1.4.1. *In the FPP auction with ad fatigue, a utility-maximizing bidder i will always pick bid $b_i > 0$ such that*

$$b_i < \frac{\alpha}{\alpha + 1} v_i$$

for any $\alpha > 0$ and for any $b_{-i} > 0$.

Proof. As before, I write the best response correspondence as the roots of the partial derivative of u_i with respect to b_i:

$$\frac{\partial u_i}{\partial b_i} = \frac{q_i b_i^{\alpha-1}}{(b_i^\alpha + b_{-i})^3} \left(f\left(\frac{b_i^\alpha}{b_i^\alpha + b_{-i}} \right) (b_i^\alpha + b_{-i}) \right. \tag{1.16}$$

$$* [-b_i^{\alpha+1} - b_{-i}(\alpha + 1) b_i + b_{-i} \alpha v_i] \tag{1.17}$$

$$\left. + \alpha b_i^\alpha b_{-i} (v_i - b_i) f'\left(\frac{b_i^\alpha}{b_i^\alpha + b_{-i}} \right) \right) = 0 \tag{1.18}$$

To make the argument, I am only interested in signs. Every term in (1.16) is non-negative since f is non-negative, and the expression in (1.18) is non-positive since f' is non-positive (f is monotone decreasing). In order to find a value of b_i so that the entire expression is equal to 0, it must be the case that the expression in (1.17) is positive.

Notice that (1.17) is the exact left-hand side of the best response correspondence for FPP without ad fatigue (Equation 1.1), so I can apply the same trick as before. I rewrite (1.17) as the following:

$$-b_i^{\alpha+1} - b_{-i}[(\alpha+1)b_i - \alpha v_i] > 0$$

As before, the expression in brackets must be negative, yielding the desired result. $\square$

Next, I analyze two specific functional forms for the ad fatigue scaling function: $f(x) = \frac{1}{\beta x+1}$ and $f(x) = e^{-\beta x}$. Both functions decline more quickly as β increases, indicating a greater rate of fatigue.

If I repeat the analysis from Sections 1.2.2 and 1.2.3, I can find the best response functions and explicitly write out the symmetric Nash equilibrium for the case when all bidders have the same value.

For the fatigue function $f(x) = \frac{1}{\beta x+1}$, I find that the best response function can be written as the roots of the following equation:

$$-(\beta+1)b_i^{\alpha+1} - b_{-i}(\alpha+1)b_i + b_{-i}\alpha v_i = 0$$

Notice that this is identical to Equation 1.1 when $\beta = 0$.

For the fatigue function $f(x) = e^{-\beta x}$, I find that the best response function can be written as the roots of the following equation:

$$-b_i^{2\alpha+1} - (\alpha+1)b_i b_{-i}^2 + b_i^{\alpha+1} b_{-i}(-2-\alpha+\alpha\beta) - \alpha b_i^{\alpha} b_{-i}(\beta-1)v_i + \alpha b_{-i}^2 v_i = 0$$

Again, this is identical to Equation 1.1 when $\beta = 0$ (substitute $\beta = 0$ and divide by $(b_i^{\alpha} + b_{-i})$).

1.4.2 Symmetric Nash Equilibrium, Ad Fatigue Extension

Here, I examine the same two ad fatigue functions as above, beginning with $f(x) = \frac{1}{\beta x+1}$. Using the same analysis as in the proof of Theorem 1.2.3, I find that in the symmetric Nash equilibrium with equal bidder values, each bidder bids:

$$b = \frac{(N-1)\alpha}{(N-1)\alpha + N + \beta} v$$

Note that this is the same result as Theorem 1.2.3 when $\beta = 0$. As the ad fatigue function becomes more severe (β increases), bidders bid less in the Nash equilibrium.

For the fatigue function $f(x) = e^{-\beta x}$, in the symmetric Nash equilibrium with equal bidder values, each bidder bids:

$$b = \frac{(N-1)(N-\beta)\alpha}{(N-1)(N-\beta)\alpha + N^2} v$$

Again, observe that this is the same result as Theorem 1.2.3 when $\beta = 0$, and bidders bid less in the Nash equilibrium as ad fatigue becomes more severe.

1.5 Empirical Evaluation

In this section, I evaluate FPP's efficiency and revenue properties and examine how value is distributed between the auctioneer and the bidders. I also study the dynamics by which I would expect bidders to reach the Nash equilibrium in practice.

I find in Section 1.5.1 that in equilibrium, FPP is quite efficient when compared to the benchmark Vickrey-Clarke-Groves (VCG) auction, which is maximally efficient. In Section 1.5.3, I find that there is a choice of exponent α that maximizes social utility, and that as α increases, the seller extracts an increasing share of the rents. Finally, in Section 1.5.4, I find that in a repeated FPP setting, bidders very quickly find the Nash equilibrium simply by myopically best-responding.

1.5.1 Efficiency of FPP

To explore the efficiency of the first-price proportional share auction, I make comparisons with the VCG allocation, which is maximally efficient[2]. I generated 100 instances of 5 bidders, all with value uniformly distributed between 0 and 1. I then computed the VCG allocation and the Nash equilibrium allocation of FPP, and examine efficiency, defined as the aggregate bidder utility. Figure 1.4 plots the ratio of FPP efficiency to VCG efficiency for 3 different values of β over varying values of α; this will be referred to as the **efficiency ratio**. For all the simulations, β refers to the parameter for the fatigue function $f(x) = \frac{1}{\beta x+1}$.

There are two observations to make in Figure 1.4. First, for a given curve (holding β constant), the efficiency ratio rises and then falls. Second, for smaller values of α ($\alpha < 7$), the curves tend to rise as β is increased.

For the first observation, as α initially increases from $\alpha = 1$, higher-valued bid-

[2] In this setting, the VCG mechanism picks the allocation $\mathbf{x}$ that maximizes $\sum_i x_i * f(x_i) * v_i$ where f is the fatigue function. Note that this means that in order for VCG to pick an allocation other than the trivial one where the highest-value bidder receives all the impressions, the fatigue function must be something other than $f(x) = 1$.

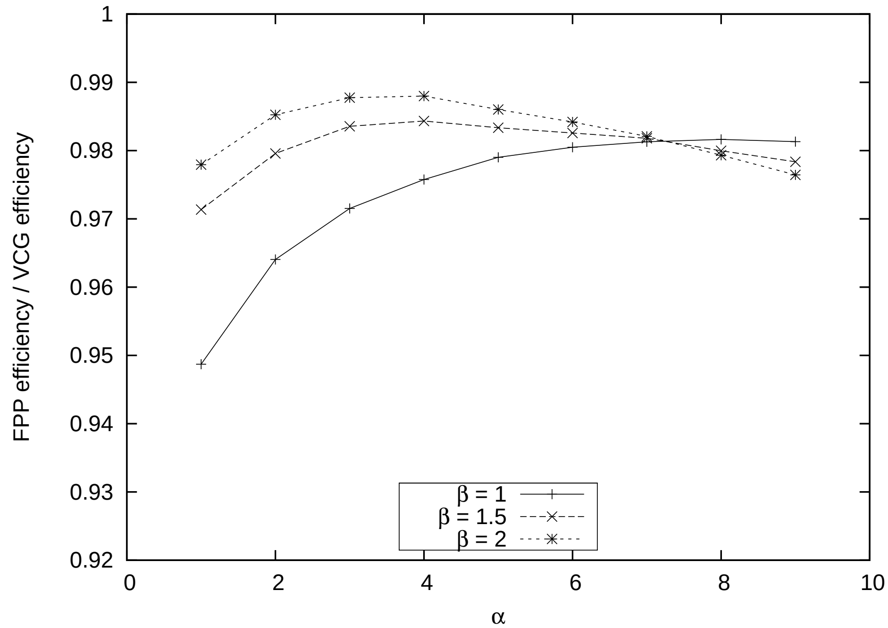


Figure 1.4: FPP efficiency versus VCG efficiency, varying α and β

ders are more heavily favored in the FPP allocation, resulting in higher total utility; meanwhile, the VCG allocation is independent of α and is thus unchanged, implying that the ratio of FPP utility to VCG utility increases as α increases. However, after a certain point, higher-valued bidders receive too much of the allocation and ad fatigue lowers the number of clicks they receive; this results in lower aggregate utility. As can be seen from Figure 1.4, this point of maximum efficiency is reached earlier when ad fatigue is more severe (when β is higher).

Regarding the second observation, for smaller values of α, the intuition behind why the efficiency ratio rises as ad fatigue increases can be seen from the example in Figure 1.5. Here, I look at a simple 3 bidder example with $(v_1, v_2, v_3) = (0.8, 0.6, 0.4)$. Figure 1.5 shows the Nash equilibrium allocations (denoted x_i) for FPP with $\alpha = 1$ and the VCG allocation along with the total utility of the system (prices paid are simply a transfer between bidders to the seller) and the efficiency ratio.

		x_1	x_2	x_3	utility	eff. ratio
$\beta = 1$	FPP	0.406	0.388	0.255	0.464	0.975
	VCG	0.555	0.346	0.099	0.476	
$\beta = 2$	FPP	0.4	0.337	0.263	0.367	0.992
	VCG	0.472	0.341	0.187	0.37	
$\beta = 3$	FPP	0.396	0.336	0.268	0.305	0.996
	VCG	0.444	0.34	0.216	0.306	

Figure 1.5: For small α, high-value bidders have smaller allocations in FPP than in VCG. As β rises, allocations shift towards lower-value bidders for both FPP and VCG, but more dramatically with VCG.

Since bidders with higher values shade more in FPP (Theorem 1.2.5), they end up with less than their socially-optimal allocation. As β rises, both FPP and VCG allocate more weight to lower-value bidders, but VCG does so more dramatically, as is evident in Figure 1.5. This occurs because high-value bidders start with lower allocations in FPP and do not have as much weight to be re-distributed to lower-value bidders as in VCG.

On the other hand, for high α (in our case, $\alpha > 7$), higher-value bidders in FPP start with higher allocations than in VCG, so more weight can be re-distributed to lower-value bidders. This means that as ad fatigue increases, FPP has re-distributes more from high-value to low-value bidders than VCG, resulting in lower aggregate utility for FPP.

1.5.2 FPP Revenue Versus VCG Revenue

While VCG is maximally-efficient, in many settings it does not have the best revenue properties. Here, I compare FPP and VCG revenue for a simple example and find that their relative magnitudes depend on a number of factors including the choice of α, the fatigue factor, and the bidder valuations.

For the 3-bidder case where bidders have value $v = [60, 40, 30]$, I compute the VCG and FPP allocations with $\alpha \in [1, 5]$. I use the fatigue function $f(x) = \frac{1}{\beta x+1}$ and vary β from 1 to 3. In Figure 1.6, I plot the ratio of FPP revenue to VCG revenue in each of these examples; note that VCG revenue does not vary with α since VCG returns the efficient allocation (which is independent of α).

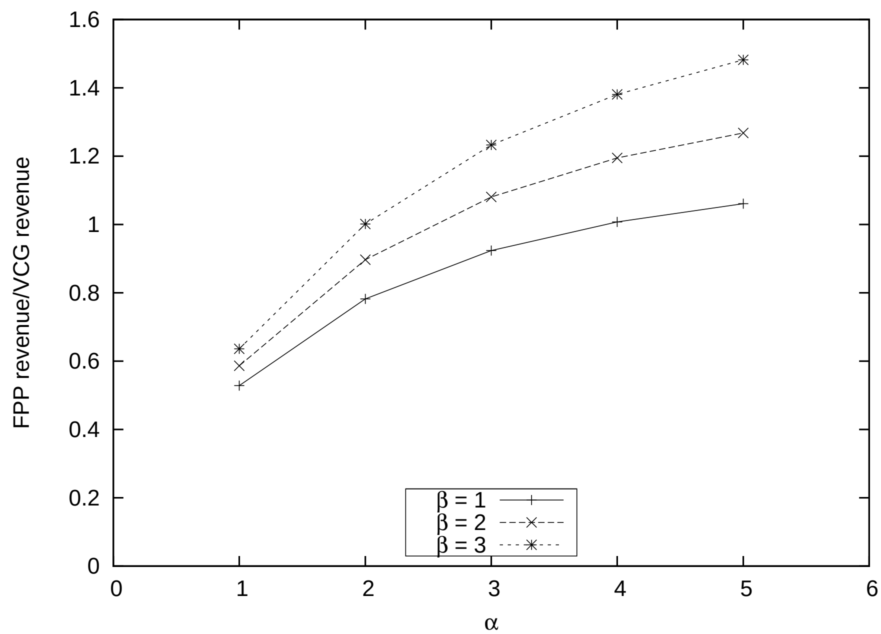


Figure 1.6: FPP revenue is lower than VCG revenue for small values of α, but is higher for greater values of α.

For the particular β coefficients here, I find that FPP revenue is lower than VCG revenue for small values of α and at some point, FPP revenue switches over to being higher than VCG revenue. Recall that as α increases, FPP favors the higher-value bidders in the allocation; since they pay a higher per-unit price, this increases seller revenue. I view this graph as a neutral result: FPP revenue can be higher or lower than VCG revenue depending on auction environment.

1.5.3 FPP Value: Magnitude and Distribution

The next two questions I address are: 1) how does the magnitude of value created (total bidder value[3] plus total seller revenue) vary with α, and 2) how are the rents distributed between the seller and the bidders? In Figure 1.7, I plot utility (measured in dollars[4]) as a function of α given a fixed fatigue $\beta = 1$. The 3 curves represent the utility captured by the auctioneer (**seller revenue**), the value captured by all bidders (**bidder value**), and the total utility in the system (**total utility**). The total utility curve is the sum of the seller revenue and bidder value curves.

The data in Figure 1.7 are generated from 100 trials of 10 bidders with value uniformly distributed between 0 and 1 with a constant fatigue β of 1. The shapes of these curves are robust to the number of bidders and β. Figure 1.7 demonstrates that there exists a socially-optimal α, which in this particular case happens around $\alpha = 6.5$. The value of the socially-optimal α depends on all the parameters.

Intuitively, as α increases, FPP will more heavily favor higher bids, meaning higher-value bidders will be favored (recall from Theorem 1.2.4 that higher-value bidders bid higher). When higher-value bidders receive more of the allocation, total social utility rises. However, there is a limit to this rise in utility because ad fatigue penalizes over-weighting of high-value bidders; without ad fatigue, total utility rises without bound as α increases.

[3]Here, total bidder value refers to total bidder utility not including payments.

[4]The scale of utility on the vertical axis depends heavily on the particular fatigue function. In the current setting with 10 bidders, $\beta = 1$, $\alpha = 1$, the maximum possible bidder utility for any allocation is 0.5, and the maximum seller revenue is $\frac{9}{22} = 0.41$.

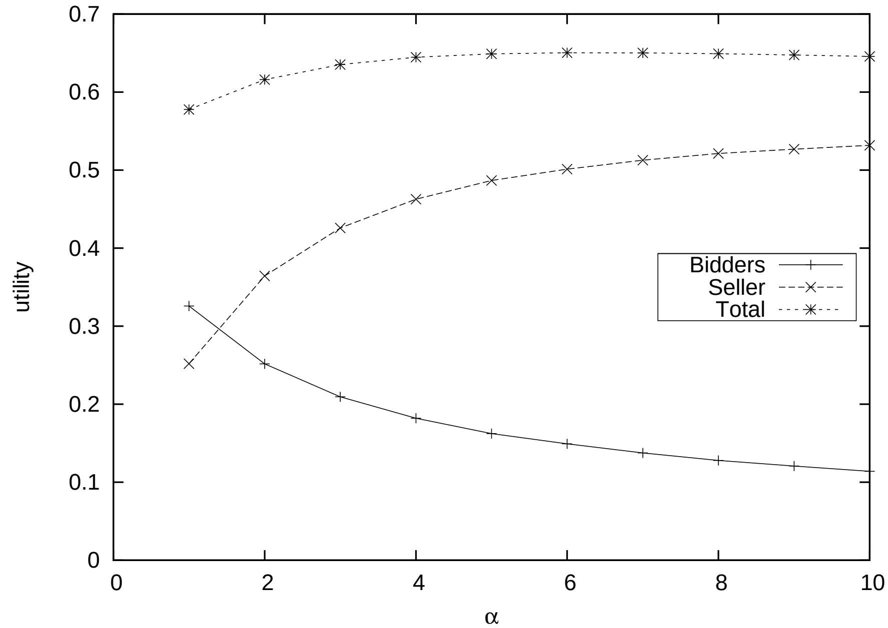


Figure 1.7: Bidder value is a convex, decreasing function of α while seller revenue is a concave, increasing function of α. Total efficiency achieves a local maximum at the socially-optimal α.

Note that the prices paid by bidders are simply transferred to the seller. In Figure 1.8, I re-scale the bottom two curves from Figure 1.7 to show the proportion of value captured by the seller versus the bidders. At $\alpha = 1$, the bidders receive just over half the total value; as α increases, the seller takes an increasing share of the value created.

1.5.4 Dynamics of Realizing the FPP Nash Equilibrium

In this section, I demonstrate that bidders can indeed reach the unique Nash equilibrium in a repeated setting by a simple process of best-responding. This strengthens the theoretical result from Theorem 1.2.2, which only demonstrates the existence of a unique Nash equilibrium. I test both synchronous updating, where all bidders si-

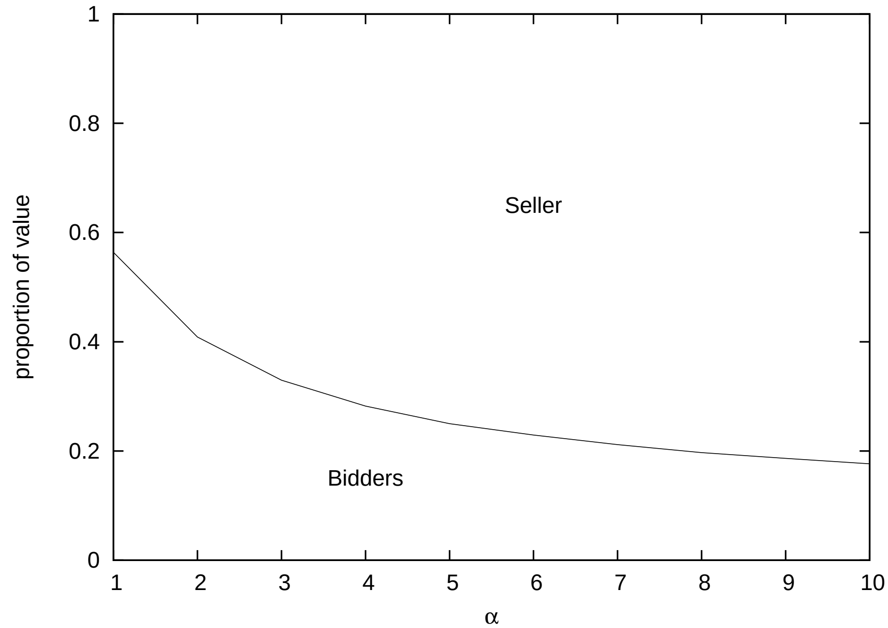


Figure 1.8: As α rises, the seller captures an increasing share of the rents.

multaneously update their bids, and asynchronous updating, where only one bidder updates his bid in a given time period.

For Figure 1.9, I simulate a dynamic game of 10 bidders with values $1, 2, \ldots, 10$ competing in the FPP auction with fatigue. Each bidder starts by bidding his true value and during every iteration, one bidder is chosen at random to myopically best-respond to the bids in the previous iteration. The graph shows that the Nash equilibrium is reached in about 25 rounds, corresponding to only around 2 to 3 bid changes per bidder. The dynamics depicted in this figure are typical, as these results are robust to choice of fatigue function, choice of α, number of bidders, and starting bids.

Simulations reveal that convergence to the Nash equilibrium also happens within 2 to 3 rounds if all bidders update their bids synchronously, even for as many as 200 bidders.

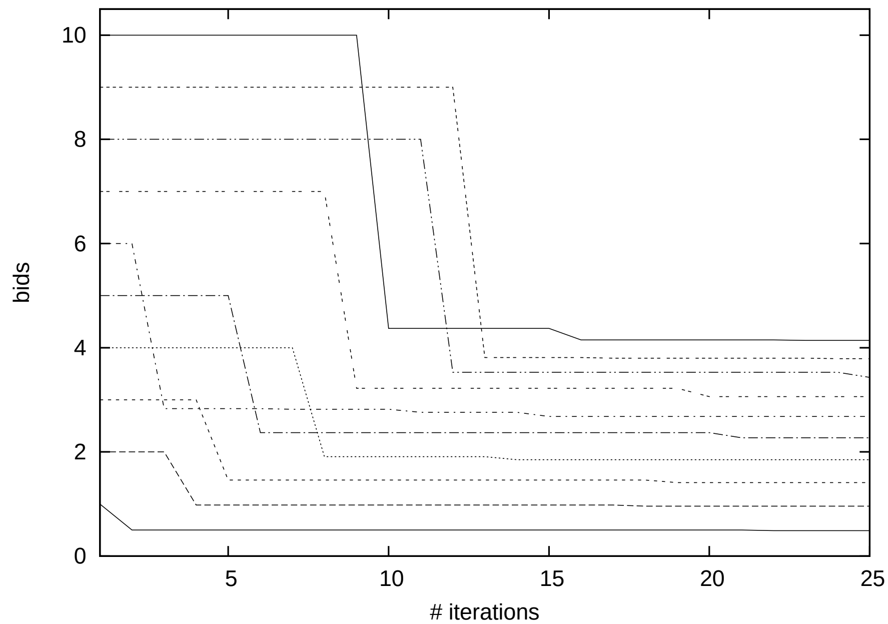


Figure 1.9: In the dynamic game of 10 bidders, bids stabilize to the Nash equilibrium within roughly 25 iterations of asynchronous myopic best response.

Finally, based on the best response behavior seen in Figure 1.9, I would expect bidders to converge to the new Nash equilibrium quickly if a bidder's true value changed. After initially shading his bid to around half his value, each bidder revises his bid only slightly in every iteration; from the others' point of view, this effectively looks like the bidder's value has changed slightly.

I would also expect the system to reach the new Nash equilibrium quickly if bidders entered or left the auction. Recall that a bidder's best response is a smooth function of the sum of all the other exponentiated bids. Given even a small number of bidders, this would mean that the loss or addition of a bidder would have minimal impact on this sum, and would not affect the best response much.

1.6 Auction Design Discussion

In this section, I consider what principles might be desirable in the design of a pay-per-click advertising auction for the mobile web. From a bidder's point of view, an ideal auction would make the bidding decision process simple: 1) bid selection would be computationally easy, 2) an unsophisticated bidder would understand the rules of the auction, and 3) a bidder would have minimal regret over his bid after the auction.

A strategy-proof mechanism like Vickrey-Clarke-Groves satisfies (1) and (3) by allowing the bidder to simply bid his true value, but both the allocation and pricing rules can be opaque to an unsophisticated bidder. To address (2), proportional share allocation rules coupled with either a first-price rule or a second-price rule seem like good candidates to consider.

In this chapter, I have demonstrated that there is a unique Nash equilibrium in the single-shot first-price proportional share auction and that bids converge quickly to this equilibrium in a dynamic setting. During the process of convergence, bidders start by arbitrarily shading down from their true value and make slight adjustments thereafter. This does not require any complex computation, satisfying (1), and at all times, bidder utility is close to what is achievable at the Nash equilibrium, satisfying (3).

In contrast, the second-price proportional share auction currently used in real-world auctions does not necessarily satisfy (1) or (3). I have shown that there are no pure-strategy Nash equilibria with this auction. A bidder needs to anticipate how the others will bid, and after every round, there will always be at least one bidder who would have favored a different bid. Over time, bids cycle over a range of values and the system never reaches a resting point. This results in higher operational costs for the auctioneer as bidders are frequently updating their bids. For all these reasons, I recommend the first-price rule over the second-price rule in the proportional share auction.

It is also interesting to consider the notion of stability when comparing the proportional share and sponsored search settings. In sponsored search auctions, the first-price rule (as originally implemented by Yahoo/Overture) leads to bid cycles (Edelman and Ostrovsky (2007)) whereas the second-price rule admits Nash equilibria (Edelman et al. (2007) and Varian (2007)). I have demonstrated in this chapter that stability properties are reversed in proportional share auctions: the first-price rule admits a unique Nash equilibrium while the second-price rule results in unpredictable bidding. These results are summarized in the following 2x2 matrix:

	proportional share	sponsored search
1st price	stable	unstable
2nd price	unstable	stable

Figure 1.10: Stability of different pricing rules in proportional share versus sponsored search auctions

Observe that stability results when payments and allocations are both continuous or both discrete. That is, in the proportional share auction, as a bidder varies his bid, both his allocated share and the price he pays vary continuously. In GSP for sponsored search, as a bidder varies his bid, the allocation and price only change when the bidder changes slots; this results in discrete changes in allocation and price. If continuity occurs only in payments (e.g. first-price sponsored search) or only in allocations (e.g. second-price proportional share), then instability results. Table 1.11 presents the same information as Table 1.10 with the axes re-labeled.

	continuous pricing	discrete pricing
continuous allocation	stable	unstable
discrete allocation	unstable	stable

Figure 1.11: Stability of different pricing and allocation rules

1.7 Conclusion

In this chapter, I have studied the proportional allocation auction under the first-price and second-price rules. I have demonstrated that there are no pure-strategy Nash equilibria in the second-price auction, and that there is a unique pure-strategy Nash equilibrium for the first-price auction in a one-shot complete information game. My simulations show that bids in the dynamic game quickly converge to this Nash equilibrium.

I also show that by tuning a single parameter in the FPP allocation rule, the auctioneer can make trade-offs between revenue and efficiency. I compare FPP with the maximally-efficient Vickrey-Clarke-Groves (VCG) auction, and find in simulation that FPP efficiency is quite efficient relative to VCG.

1.8 Appendix to Chapter 1

I outline the proofs for the ad fatigue analogues of Theorems 1.2.2, 1.2.4, and 1.2.5.

Theorem 1.8.1. *For all v_i, there exists one unique pure-strategy Nash equilibrium in the FPP auction* ***with ad fatigue*** *if a minimum bid is enforced of no more than M for all i, where*

$$M = \frac{(N-1)\alpha v_i}{(N-1)\alpha + N}$$

Proof. As in the original proof of Theorem 1.2.2, I will apply Brouwer's fixed point theorem to show the existence of a pure-strategy Nash equilibrium. I will once again assume the presence of a minimum bid m.

First, I prove the lemmas regarding how the ad fatigue best response function maps from the space $R = [m, v_1] \times \ldots \times [m, v_N]$ to itself and is a continuous function.

Lemma 5. *The best response function maps from the space $R = [m, v_1] \times \ldots \times [m, v_N]$ to itself.*

Proof. The v_i upper bound of the interval follows from Theorem 1.4.1 since every bidder i will always bid below $\frac{\alpha}{\alpha+1} v_i$, which is less than v_i for any $\alpha > 0$.

As before, I characterize the conditions under which best responses will be at least m assuming that bidders all bid the minimum. As shown in Section 1.4.1, the ad fatigue best response function is the Solution to the following equation:

$$\begin{aligned} f\left(\frac{b_i^\alpha}{b_i^\alpha + b_{-i}}\right)(b_i^\alpha + b_{-i}) \\ *[-b_i^{\alpha+1} - b_{-i}(\alpha+1)b_i + b_{-i}\alpha v_i] \\ +\alpha b_i^\alpha b_{-i}(v_i - b_i) f'\left(\frac{b_i^\alpha}{b_i^\alpha + b_{-i}}\right) = 0 \end{aligned}$$

I know that $b_{-i} = \sum_{j \neq i} m^\alpha = (N-1)m^\alpha$ since every bidder is bidding m and can substitute this into the above expression. I also substitute m for b_i and find when the expression is at least 0 to determine when the best response is at least m:

$$f\left(\frac{1}{N}\right) Nm^{\alpha} \tag{1.19}$$

$$* \left(-m^{\alpha+1} - (N-1)m^{\alpha}(-(\alpha+1)m + \alpha v_i)\right) \tag{1.20}$$

$$+\alpha(N-1)m^{2\alpha}(v_i - m)f'\left(\frac{1}{N}\right) \geq 0 \tag{1.21}$$

(1.19) is non-negative since f is non-negative and (1.21) is non-positive since f' is non-positive (f is monotone decreasing). This means that the (1.20) must be non-negative and the above expression reduces to:

$$-m + (N-1)[-(\alpha+1)m + \alpha v_i] \geq 0$$
$$\Leftrightarrow m \leq \frac{(N-1)\alpha v_i}{(N-1)\alpha + N} = M$$

Thus, if the minimum bid is less than or equal to M, the best response function will map each bidder's initial minimum bid to a higher bid.

I can show that the best response function is increasing in b_{-i} with the same proof by contradiction as before. □

Lemma 6. *The best response function is continuous.*

The proof of this lemma for the ad fatigue case is identical to the original proof. Having established these 2 lemmas, the preconditions of Brouwer's fixed point theorem are met, demonstrating the existence of a Nash equilibrium.

Finally, I can demonstrate that there can only be 1 unique Nash equilibrium by assuming the existence of 2 Nash equilibria and writing one in terms of the other. I reach the same contradiction as before. □

Theorem 1.8.2. *In the Nash equilibrium for FPP* ***with ad fatigue****, bids are ordered by value. That is, for N bidders, given values $\{v_1, \dots, v_N\}$, Nash bid profile $\{b_1, \dots, b_N\}$, and any $\alpha > 1$,*

$$v_i > v_j \Leftrightarrow b_i > b_j \ \ \forall i \neq j$$

Theorem 1.8.3. *In the Nash equilibrium for FPP* ***with ad fatigue****, shading factors are in reverse order from value. That is, for N bidders, and any $\alpha > 1$, if $\{b_1, \dots, b_N\}$ is the Nash equilibrium bid profile with shading factor c_i defined as $c_i = \frac{b_i}{v_i}$ for all i, then*

$$v_i > v_j \Leftrightarrow c_i < c_j \ \ \forall i \neq j$$

The proofs of Theorems 1.8.2 and 1.8.3 follow the original proofs exactly as long as I make the additional assumption that the fatigue function f is weakly concave, i.e. $\frac{\partial f^2}{\partial^2 x} \leq 0$.

Chapter 2

Maximum Entropy Banner Allocation

I examine an optimize-and-dispatch scheme for delivering pay-per-impression advertisements in online and mobile advertising. Using traffic predictions based on historical traffic patterns, the platform provider seeks to allocate future inventory to advertisers such that commitments are fulfilled in expectation, and no single advertiser bears too much of the burden if actual traffic diverges from predicted traffic. I propose a maximum entropy approach and provide theoretical analysis and simulation to show how it accomplishes these goals.

2.1 Introduction

Currently, advertising in the online and mobile arenas employs a number of pricing models including cost-per-click (CPC) and cost-per-impression (CPM). Much of the academic literature on online advertising has concentrated on allocating advertisements to web-surfers through real-time auctions, such as Google's Generalized Second Price auction (Edelman et al. (2007)). In practice, auctions are often used to make allocation decisions for both CPC and CPM price schemes. While CPM or banner advertising can be implemented with an auction, in some cases brand advertisers might prefer to enter into a long-term contract with the platform provider[1] to ensure that their advertisements are shown a certain number of times over a specified time period. For example, a movie advertiser might wish to guarantee a minimum number of ad views to coincide with an upcoming theater release.

This chapter will focus on the question of how an advertising platform provider should make allocation decisions when it has entered into a number of long-term CPM commitments. Because these commitments are usually known far in advance of ad-serving time, an "optimize-and-dispatch" approach (Parkes and Sandholm (2005)) can be used, in which the winner-determination problem is solved in an offline optimization and the solution is used to parameterize a simple online dispatcher.

In particular, I take the point of view of the ad-serving platform provider that has made a number of CPM commitments with various advertisers. Each commitment specifies a price that will be paid per impression[2] and a set of partner websites that the advertiser wishes to target. The platform provider can forecast future advertising inventory based on historical user behavior. A unit of inventory becomes available

[1] Borrowing from the language of the literature on two-sided markets (see Eisenmann et al. (2006) and Rochet and Tirole (2006) for an overview), two distinct sets of network users transact with each other and affiliate with a platform provider. Here, the users are advertisers transacting with "publishers" or the "content network" mediated by the platform providers AdMob or Google.

[2] In this chapter, I will use "impressions," "inventory," and "visits" (as in, website visits) interchangeably.

when a user makes a request for an advertisement by visiting a partner website.

The goal of the optimization problem is to pick a set of probabilistic weights used by the ad server to decide which ad to serve when a user arrives at a partner website. For each advertiser-site pair, I pick a weight representing the probability that the advertiser's ad will be shown if a user goes to the site.

Inventory forecasts are by nature imprecise, so I seek a method for allocating future inventory to advertisers such that all commitments are fulfilled in expectation. In general, many allocation rules are possible, so I am interested in a decision rule that is forgiving when the forecasting engine's predictions turn out to be incorrect. I describe such a method using maximum entropy and prove that it is robust to imprecise predictions by the forecasting engine.

This particular scheme is agnostic to advertisers and sites, treating them all equally. While it may be the case that the platform provider wishes to favor certain high-value advertisers over others, it seems plausible that there would be classes of advertisers for which the platform provider would want to treat advertisers within each class equally. In that case, this scheme can be used for allocation decisions within a class.

2.1.1 Related Work

Some early work on online banner allocation is presented in Chickering and Heckerman (2000). The authors describe a setting in which a single website operator makes commitments to various advertisers to deliver a fixed number of impressions. They employ a linear program to maximize expected click-through rate subject to inventory management constraints. Their delivery system computes clusters for targeting based on user click behavior, but the advertisers have no ability to target. My work addresses the more generalized problem of banner allocation faced by a platform provider connecting multiple advertisers with multiple sites, instead of just one site. In addition, I allow for advertisers to specify targeting parameters, a feature especially valuable for large brand advertisers interested in reach or the demographic mix of users who see their ads.

Tomlin (2000) presents a maximum entropy approach for a similar setting. As with Chickering and Heckerman (2000), "targeting" in Tomlin (2000) refers to user-based clustering. Tomlin demonstrates that a maximum entropy allocation shows each advertiser's banners to a wider mix of users than other allocation strategies that maximize the clickthrough rates. He also assumes that the number of impressions available at each site is known. The work presented in this chapter rigorously proves that the maximum entropy allocation not only shows the banners to a wider mix, but also provides a more robust allocation policy in the face of uncertainty of arrival rates.

This work is an instance of an "optimize-and-dispatch" architecture (Parkes and Sandholm (2005)). In particular, I allow advertisers to express targeting preferences, which are used in an offline optimization to enable real-time ad-serving. Optimize-and-dispatch is also implicit in the approaches of Chickering and Heckerman (2000) and Tomlin (2000).

2.2 Banner Allocation Problem

The banner allocation problem I study is composed of two different optimization problems: the delivery problem and the commitment problem. The delivery problem returns an allocation of probabilistic weights to be used by the ad server while the commitment problem returns the amount of inventory available for commitment to a new advertiser. The commitment problem is solved once for each new advertiser and these commitments are used to solve the delivery problem once per day. Both optimization problems are subject to the same constraints, referred to as the **banner allocation problem**, and they make use of different objective functions. I begin with an example before introducing notation.

2.2.1 Example

I conceptualize the banner allocation problem as a bipartite graph with advertisers on the left, sites on the right, and edges indicating which sites are targeted by the advertisers:

$$
\begin{tikzcd}
1 \arrow[dr] & \\
2 \arrow[r] \arrow[dr] & A \\
3 \arrow[r] & B \\
4 \arrow[ur] &
\end{tikzcd}
$$

The delivery model requires that that optimizer has been exogenously given commitments[3] for each advertiser $C_1, \ldots, C_4$ and predicted inventories for each site $\hat{I}_A, \hat{I}_B$.

[3] The maximum amount of inventory that can be committed to an advertiser is computed in the banner commitment problem. This value is then used in negotiations with the advertiser as a ceiling

The delivery optimization problem picks probabilistic weights for each of the edges in the graph representing the probability that the given advertiser's ad will be shown on the given site.

2.2.2 Notation

The banner allocation problem is formally defined as follows:

Indices

$t \in \mathcal{T}$	The targets (websites)
$a \in \mathcal{A}$	The advertisers

Variables

$w_{t,a}$	Probability of serving advertiser a's ad on target t;
	$w_{t,a}$ only exists if advertiser a chooses to target site t

Data

$\hat{I}_t$	Estimated number of impressions served for target t
C_a	Number of impressions committed to advertiser a

Constraints

$$\sum_{a \in \mathcal{A}} w_{t,a} = 1 \quad \forall t \in \mathcal{T} \qquad \text{(Delivery)} \tag{2.1}$$

$$\sum_{t \in \mathcal{T}} \hat{I}_t w_{t,a} = C_a \quad \forall a \in \mathcal{A} \qquad \text{(Commitments)} \tag{2.2}$$

$$w_{t,a} > 0 \quad \forall t \in \mathcal{T},\ \forall a \in \mathcal{A} \qquad \text{(Positive weights)} \tag{2.3}$$

The commitment constraint makes use of a forecasting engine for estimating inventories $\hat{I}_t = E[I_t]$ where random variable I_t is the actual inventory for target t. The

figure for how much inventory can be committed to the advertiser.

commitment constraint also takes the number of committed impressions C_a as exogenously given.

Note that the goal of fulfilling commitments in expectation need not be a hard constraint. One might imagine that in order to build in some slack, the platform provider could select weights so that the ad-serving engine plans to serve more than the promised commitment; this could be implemented by adding a scaling factor greater than 1 to the right-hand side of the commitment constraint.

Any selection of weights $\{w_{t,a}\}$ that satisfies the preceding constraints constitutes a feasible solution to the **banner allocation problem**.

2.2.3 Banner Commitment Problem

The banner commitment problem is an optimization problem that returns the maximum number of impressions that can be allocated to a new advertiser without violating any existing commitments (in expectation). For new advertiser $\tilde{a}$, this value $\overline{C}_{\tilde{a}}$ is:

$$\begin{aligned}
\overline{C}_{\tilde{a}} = \max_{\mathbf{w}} \quad & \sum_{t \in \mathcal{T}} w_{t,\tilde{a}} \hat{I}_t \\
\text{subject to} \quad & \sum_{a \in \mathcal{A}} w_{t,a} = 1, \qquad \forall t \in \mathcal{T} \\
& \sum_{t \in \mathcal{T}} \hat{I}_t w_{t,a} = C_a, \qquad \forall a \in \mathcal{A} \cup \{\tilde{a}\} \\
& w_{t,a} > 0, \qquad \forall t \in \mathcal{T},\ \forall a \in \mathcal{A} \cup \{\tilde{a}\}
\end{aligned}$$

If the optimizer chooses the actual commitment $C_{\tilde{a}}$ so that it is less than or equal to $\overline{C}_{\tilde{a}}$, then it must be the case that every instance of the banner commitment problem and the banner delivery problem is feasible. As new advertisers enter the system, the banner commitment problem is solved and informs the platform provider how much inventory is available for the new advertisers. These commitments are then used as inputs to the banner delivery problem.

2.2.4 Banner Delivery Problem: Maximum Entropy Formulation

Any selection of weights $w_{t,a}$ that satisfies the banner allocation problem constraints (2.1), (2.2), and (2.3) is a valid solution to the banner delivery problem. In practice, when a user surfs to site t, the online dispatcher looks at the set of all the advertisers targeting t and randomly selects advertiser a based on the weights $w_{t,a}$. If all commitments to all advertisers targeting t have already been fulfilled, the dispatcher selects an ad from outside the banner allocation system; this ad is usually chosen through a CPC auction[4].

In this section, I describe the maximum entropy solution, beginning with a definition of entropy.

Definition The Shannon entropy of a variable X is

$$H(X) \equiv -\sum_{x} P(x) log[P(x)]$$

where $P(x)$ is the probability that is in the state x.

[4]Note that long-term CPM contracts are fulfilled first since they are more profitable than CPC ads and since under-serving these commitments leads to penalties.

In this setting, a state x corresponds to an edge in the advertiser-site graph and $P(x)$ corresponds to the weight of an edge $w_{t,a}$. In the absence of other constraints, entropy maximization returns the uniform distribution. In this section, I show that the maximum entropy solution has useful properties when constrained by (2.1), (2.2), and (2.3) (see Theorem 2.2.1).

More formally, I maximize Shannon entropy subject to the constraints of the banner allocation problem:

$$
\begin{aligned}
\max_{\mathbf{w}} \quad & \sum_{t \in \mathcal{T}} \sum_{a \in \mathcal{A}} -w_{t,a} \ln w_{t,a} \\
\text{subject to} \quad & \sum_{a \in \mathcal{A}} w_{t,a} = 1, \qquad \forall t \in \mathcal{T} \\
& \sum_{t \in \mathcal{T}} \hat{I}_t w_{t,a} = C_a, \qquad \forall a \in \mathcal{A} \\
& w_{t,a} > 0, \qquad \forall t \in \mathcal{T},\ \forall a \in \mathcal{A}
\end{aligned}
\tag{2.4}
$$

This concave maximization is equivalent to a convex minimization, so the solution can be computed efficiently using convex optimization techniques (Boyd and Vandenberghe (2004)).

Next, I will show in Theorem 2.2.1 that when the maximum entropy problem is solved, the objective value provides an upper bound on the maximum shadow price for the commitment constraint. To interpret this result, recall that a shadow price, or the value of the Lagrangian dual variable at the optimal solution, reflects the change in the objective value of the optimal solution when relaxing the constraint by one unit. Here, the objective function is the information entropy of the weight vector, and the shadow price of interest is the Lagrangian dual variable for the commitment constraint. The entropy of the weight vector is a measure of randomness, or the evenness of the allocation: for a higher entropy allocation, the impressions are spread more evenly among all the advertisers (i.e. closer to the uniform distribution).

Since inventory commitments are based on estimated traffic patterns, it is possible for there to be violations to commitments made in (2.2). By bounding the shadow price on the commitment constraint, I limit the change in the objective value if commitments are violated.

Theorem 2.2.1. *The maximum entropy solution to the banner allocation problem provides an upper bound on the maximum shadow price for the commitment constraint. Formally, the concave maximization in* (2.4) *is equivalent to the following convex minimization:*

$$
\begin{aligned}
\min_{\mathbf{w}} \quad & \sum_{t\in\mathcal{T}}\sum_{a\in\mathcal{A}} w_{t,a}\ln w_{t,a} && (2.5)\\
\textit{subject to} \quad & \sum_{a\in\mathcal{A}} w_{t,a} = 1, \qquad \forall t\in\mathcal{T} \\
& \sum_{t\in\mathcal{T}} \hat{I}_t w_{t,a} = C_a, \qquad \forall a\in\mathcal{A} \\
& w_{t,a} > 0, \qquad \forall t\in\mathcal{T},\ \forall a\in\mathcal{A}
\end{aligned}
$$

If the vector λ^* *denotes the optimal values of the dual variables corresponding to the commitment constraints and the vector* $\mathbf{w}^*$ *denotes the optimal values of the primal variables, then:*

$$\max_{a\in\mathcal{A}} \lambda_a^* \le \frac{-\sum_{t\in\mathcal{T}}\sum_{a\in\mathcal{A}} w_{t,a}^*\ln w_{t,a}^*}{\sum_{t\in\mathcal{T}}\hat{I}_t - \sum_{a\in\mathcal{A}} C_a}$$

Proof. The dual problem to (2.5) with dual variables ν and λ for the equality constraints and μ for the inequality constraints is given by:

$$\max_{\lambda,\nu,\mu} \quad -\sum_{t\in\mathcal{T}} \nu_t + \sum_{a\in\mathcal{A}} \lambda_a C_a + \sum_{t\in\mathcal{T}} e^{-\nu_t - 1} \sum_{a\in\mathcal{A}} e^{\hat{I}_t \lambda_a + \mu_{t,a}}$$

$$\text{subject to} \quad \lambda \succeq 0,\ \nu \succeq 0,\ \mu \succeq 0$$

(See Appendix 2.5 for a derivation of the dual).

The dual objective function is maximized when the derivative with respect to ν_t is zero:

$$\nu_t = \ln\left(\sum_{a\in\mathcal{A}} e^{\hat{I}_t \lambda_a + \mu_{t,a}}\right) - 1$$

Substituting ν_t into the objective and multiplying by -1 to switch to a minimization, the dual problem is now:

$$\min_{\lambda,\mu} \quad \sum_{t\in\mathcal{T}} \ln\left(\sum_{a\in\mathcal{A}} e^{\hat{I}_t \lambda_a + \mu_{t,a}}\right) - \sum_{a\in\mathcal{A}} \lambda_a C_a \tag{2.6}$$

$$\text{subject to} \quad \lambda \succeq 0,\ \mu \succeq 0$$

Slater's conditions are satisfied for the problem in (2.6) and hence, the duality gap is zero.

If the vector λ^* denotes the optimal values of the dual variables corresponding to the commitment constraints, μ^* denotes the optimal values of the dual variables corresponding to the positive weight constraints, and $\mathbf{w}^*$ denotes the optimal values of the primal variables, I can conclude the following:

$$
\begin{aligned}
-\sum_{t\in\mathcal{T}}\sum_{a\in\mathcal{A}} w^*_{t,a}\ln w^*_{t,a} &= \sum_{t\in\mathcal{T}}\ln\left(\sum_{a\in\mathcal{A}} e^{\hat{I}_t\lambda^*_a+\mu^*_{t,a}}\right)-\sum_{a\in\mathcal{A}}\lambda^*_a C_a && (2.7)\\
&\geq \sum_{t\in\mathcal{T}}\max_{a\in\mathcal{A}}(\hat{I}_t\lambda^*_a+\mu^*_{t,a})-\left(\max_{a\in\mathcal{A}}\lambda^*_a\right)\sum_{a\in\mathcal{A}} C_a && (2.8)\\
&\geq \sum_{t\in\mathcal{T}}\max_{a\in\mathcal{A}}(\hat{I}_t\lambda^*_a)-\left(\max_{a\in\mathcal{A}}\lambda^*_a\right)\sum_{a\in\mathcal{A}} C_a\\
&= \sum_{t\in\mathcal{T}}\hat{I}_t\left(\max_{a\in\mathcal{A}}\lambda^*_a\right)-\left(\max_{a\in\mathcal{A}}\lambda^*_a\right)\sum_{a\in\mathcal{A}} C_a\\
&= \left(\max_{a\in\mathcal{A}}\lambda^*_a\right)\left[\sum_{t\in\mathcal{T}}\hat{I}_t-\sum_{a\in\mathcal{A}} C_a\right] && (2.9)
\end{aligned}
$$

Equation (2.7) follows from the fact that the duality gap is zero, so the objective value from maximizing entropy in (2.5) is equivalent to the objective value from minimizing the dual in (2.6). In (2.8), I use the softmax approximation for the max function:

$$\max\{x_1,\ldots,x_n\}\leq\ln\sum_i e^{x_i}$$

In (2.9), the expression in brackets is a scalar representing the total amount of predicted traffic minus the total commitments.

Re-arranging, I arrive at the result:

$$\max_{a\in\mathcal{A}}\lambda^*_a\leq\frac{-\sum_{t\in\mathcal{T}}\sum_{a\in\mathcal{A}} w^*_{t,a}\ln w^*_{t,a}}{\sum_{t\in\mathcal{T}}\hat{I}_t-\sum_{a\in\mathcal{A}} C_a} \tag{2.10}$$

In (2.10), observe that the entropy expression in the numerator on the right-hand side is always positive since the weights $w^*_{t,a}$ are between 0 and 1; also, the denominator is always positive since the commitments $\sum_{a\in\mathcal{A}} C_a$ are chosen in order to avoid exceeding the predicted inventory $\sum_{t\in\mathcal{T}}\hat{I}_t$. To summarize, the maximum entropy value provides an upper bound on the maximum shadow price for the commitment constraint. $\square$

2.3 Empirical Evaluation

2.3.1 Commitment Violations

The goal of this current design is to choose a feasible solution to the banner allocation problem that is robust to imprecise inventory estimates. To evaluate the maximum entropy solution empirically, I solve the offline optimization and simulate use of the maximum entropy weights in the online ad-serving mechanism. I measure robustness by counting the number of commitment violations that would be made for each advertiser. In particular, for each advertiser a, I sum the number of impressions over-committed and under-committed:

$$v_a = \left| \sum_{t \in \mathcal{T}} (\hat{I}_t - I_t) * w_{t,a} \right|$$

where I_t is the realized number of impressions for site t. Since I take the point of view of the platform provider, I am interested in both over- and under-commitments. Over-commitments result in direct revenue loss and reputation damage to the platform provider[5] while under-commitments result in lost opportunities. There is no particular motivation for treating over- and under-commitment differently, so for this model I treat them symmetrically. The goal is to spread commitment violations uniformly across all the advertisers; in practice, the penalties for missing commitments are marginally increasing with the size of the missed commitment.

[5] Depending on the specifics of the contract, under-serving can lead to linear costs in terms of direct revenue loss, but damage to the platform provider's reputation can have long-term repercussions that are more difficult to measure.

2.3.2 Methodology

In order to test the maximum entropy (ME) approach to the banner allocation problem, I measured its performance against two other objective functions: least-squares (LS) and least-norm (LN), i.e., minimizing the ℓ_2 and ℓ_1 norms. I generated problem instances randomly and solved the banner allocation for each problem instance with each of the three objectives. Each problem instance consisted of 4 components:

1. advertiser/target graph
2. actual traffic I_t drawn from a stochastic distribution
3. predicted traffic $\hat{I}_t$ for each target t
4. commitments C_a for each advertiser a

I randomly created advertiser/target graphs with varying numbers of advertisers and targets, and with varying numbers of edges. I created commitments for each advertiser by sequentially allotting inventory to each advertiser. For each advertiser a, I took every site t targeted by a and allot to a a value drawn from the uniform distribution from 0 to the amount of inventory still available from site t. This mimicked the actual process of creating commitments whereby commitments to new advertisers enter the system sequentially. For each problem instance, I generated 50 days of traffic realizations. I drew traffic from three distributions: Poisson, multinomial, and multivariate normal. I found the results to be robust to choice in distribution, so I only present results for traffic drawn from the multivariate normal distribution.

For each objective function, I computed the commitment violation v_a for every advertiser a, resulting in a vector of violations $\mathbf{v}$. I then normalized the vector $\mathbf{v}$ into the unit simplex $\hat{\mathbf{v}}$ ($\sum_i \hat{v}_i = 1$). This new vector $\hat{\mathbf{v}}$ could then be compared to the uniform distribution $\mathbf{u} = (\frac{1}{|\mathcal{A}|}, \ldots, \frac{1}{|\mathcal{A}|})$. If the commitment violations were perfectly spread among all the advertisers, then $\hat{\mathbf{v}}$ would equal $\mathbf{u}$. Thus, I was able to compare the evenness of violations experienced by the advertisers under the 3 objective functions by examining the distances between $\hat{\mathbf{v}}_{obj}$ and $\mathbf{u}$ for $obj = \text{ME}, \text{LS}, \text{LN}$.

In particular, I measured the Kullback-Leibler divergence (KL) of the uniform distribution from $\hat{\mathbf{v}}$ for each of the three objective functions: $KL(\hat{\mathbf{v}}_{ME}||\mathbf{u})$, $KL(\hat{\mathbf{v}}_{LS}||\mathbf{u})$, and $KL(\hat{\mathbf{v}}_{LN}||\mathbf{u})$. Recall that:

$$KL(P||Q) = \sum_i P(i) \log_2 \frac{P(i)}{Q(i)}$$

I define the uniformity score U as a value between 0 and 1 that measures how uniformly violations are spread between the advertisers:

$$U(\hat{\mathbf{v}}) := \frac{H(\hat{\mathbf{v}})}{H(\mathbf{u})} = 1 - \frac{KL(\hat{\mathbf{v}}, \mathbf{u})}{H(\mathbf{u})}$$

where $H(\mathbf{x}) = -\sum_i x_i \log_2 x_i$ is the information entropy of probability vector $\mathbf{x}$.

2.3.3 Results

My simulations reveal that on average, the ME allocation achieves a higher uniformity score than either LS or LN. In Figure 2.1, I show uniformity scores for all 3 objective functions on a graph with equal numbers of advertisers and sites, and traffic drawn from a Poisson distribution; I vary the graph density from 0 (no targeting) to

1 (fully-connected). ME consistently outperforms LN and outperforms LS for graph densities greater than 0.3.

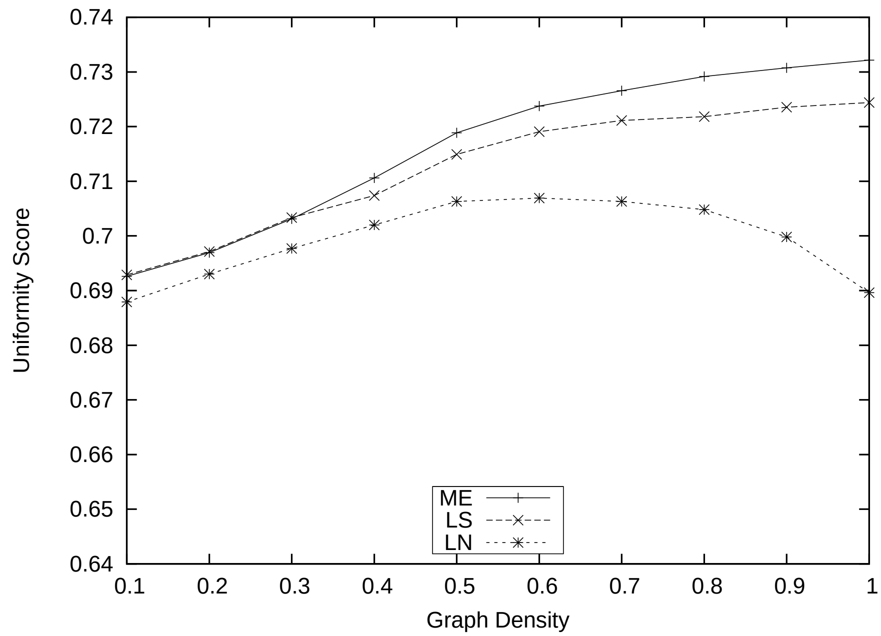


Figure 2.1: Uniformity scores for Poisson-distributed traffic, $|\mathcal{T}| = |\mathcal{A}|$

The shape of the ME and LS curves reflects an increase in uniformity score as graph density increases. The intuition behind this shape is that with fewer edges in the advertiser-site targeting graph, each edge bears a greater burden in carrying the flow of advertisements. Thus, when actual inventory differs from predicted inventory for a particular site, the edges connected to that site carry more weight and violations are not spread as evenly. In aggregate, this lowers the potential uniformity score that can be achieved.

The downward-sloping portion of the LN curve reflects that the LN objective is not effectively spreading commitment violations evenly among the advertisers as graph density increases.

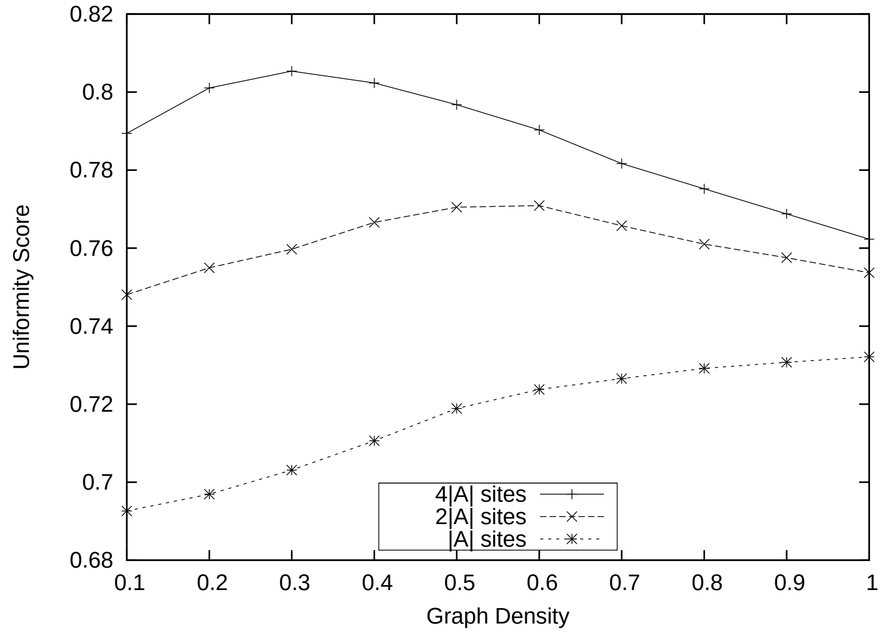


Figure 2.2: Maximum entropy uniformity scores with varying site count

In Figure 2.2, I present the effect of increasing the number of sites relative to the number of advertisers[6]. As the ratio of sites to advertisers increases, the uniformity score rises. The intuition behind this behavior is the same as the intuition for why uniformity scores rise with increasing graph density–with more sites, each edge bears less of the flow and thus the potential for spreading violations evenly increases, thus increasing the uniformity score.

For the top two curves where there are more advertisers than sites, above a certain graph density, the curve starts sloping downwards. The intuition here is that when the graph gets too dense, each advertiser has many targeted sites and is likely to be

[6]I only examined graphs with more sites than advertisers since I am modeling the situation where the platform provider actively enters into contracts with CPM advertisers. Signing up new advertisers does not scale as quickly as signing up new websites, a process which is largely self-serve or medium-serve.

affected whenever there are traffic shocks somewhere in the system.

The policy implications of Figure 2.2 are that the platform provider can improve the uniformity score by influencing the ratio of sites to advertisers (e.g., signing on sites more quickly than advertisers) or by influencing the graph density (e.g., suggesting to the advertisers that they may wish to target at a certain level). With regard to targeting recommendations, the platform provider faces a trade-off: encouraging advertisers to target widely makes it easier for the platform provider to fulfill commitments while potentially decreasing ad relevance; on the flip side, encouraging highly-specific targeting makes it harder for the platform provider to fulfill commitments while increasing ad relevance. This results in a tension since both better performance in fulfilling commitments and increasing ad relevance will likely tend to increase revenue to the platform provider.

2.4 Conclusion

In this chapter, I have described an optimize-and-dispatch approach for delivering pay-per-impression advertisements. Using traffic predictions based on historical traffic patterns, the platform provider seeks to allocate future inventory to advertisers such that commitments are fulfilled in expectation, and no single advertiser bears too much of the burden if actual traffic diverges from predicted traffic. I proposed a maximum entropy approach and provided theoretical analysis and simulation to show how it accomplishes these goals.

2.5 Appendix to Chapter 2

I derive the Lagrangian dual function for the maximum entropy objective function subject to the constraints of the banner allocation problem.

The maximum entropy problem is the following:

$$
\begin{aligned}
\max_{\mathbf{w}} \quad & \sum_{t\in\mathcal{T}}\sum_{a\in\mathcal{A}} -w_{t,a}\ln w_{t,a} \\
\text{subject to} \quad & \sum_{a\in\mathcal{A}} w_{t,a} = 1, \qquad \forall t\in\mathcal{T} \\
& \sum_{t\in\mathcal{T}} \hat{I}_t w_{t,a} = C_a, \qquad \forall a\in\mathcal{A} \\
& w_{t,a} > 0, \qquad \forall t\in\mathcal{T},\ \forall a\in\mathcal{A}
\end{aligned}
\tag{2.11}
$$

The Lagrangian with dual variables ν and λ for the equality constraints and μ for the inequality constraints is:

$$
\begin{aligned}
L(\mathbf{w},\lambda,\mu,\nu) = & \sum_{t\in\mathcal{T}}\sum_{a\in\mathcal{A}} w_{t,a}\ln w_{t,a} + \sum_{t\in\mathcal{T}}\nu_t\left(\sum_{a\in\mathcal{A}} w_{t,a}-1\right) \\
& + \sum_{a\in\mathcal{A}}\lambda_a\left(\sum_{t\in\mathcal{T}}\hat{I}_t w_{t,a}-C_a\right) + \sum_{t\in\mathcal{T}}\sum_{a\in\mathcal{A}}(-\mu_{a,t}w_{t,a})
\end{aligned}
$$

Consider a general convex minimization problem with inequality and equality constraints,

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & f(x) \\
\text{subject to} \quad & Ax \preceq b \\
& Cx = d
\end{aligned}
\tag{2.12}
$$

I write the dual function for (2.12) with equality dual variables λ and inequality dual variables ν using the convex conjugate of f, which I denote f^* (see Boyd and

Vandenberghe (2004)):

$$
\begin{aligned}
g(\lambda,\nu) &= \inf_x \left(f(x) + \lambda^T(Ax-b) + \nu^T(Cx-d)\right) \\
&= -b^T\lambda - d^T\nu + \inf_x (f(x) + (A^T\lambda + C^T\nu)^T x) \\
&= -b^T\lambda - d^T\nu - f^*(-A^T\lambda - C^T\nu)
\end{aligned}
$$

Recall that the convex conjugate is defined as:

$$f^* = \sup_{x\in \mathbf{dom} f} (y^T x - f(x))$$

In (2.11), f is the sum of negative entropy functions: $f(\mathbf{x}) = \sum_i x_i \ln x_i$ with $\mathbf{dom} f = [0,1]$. The function $\sum_i x_i y_i - x_i \ln x_i$ attains its maximum at $x_i = e^{y_i - 1}, \forall i$. By substituting, I find the convex conjugate:

$$f^*(\mathbf{y}) = \sum_i e^{y_i - 1} \tag{2.13}$$

Applying (2.13), I write the dual function for (2.11):

$$
\begin{aligned}
g(\lambda,\nu,\mu) &= \inf_{\mathbf{w}} \Bigg[\sum_{t\in\mathcal{T}} \sum_{a\in\mathcal{A}} w_{t,a} \ln w_{t,a} + \sum_{t\in\mathcal{T}} \nu_t \left(\sum_{a\in\mathcal{A}} w_{t,a} - 1 \right) \\
&\quad + \sum_{a\in\mathcal{A}} \lambda_a \left(- \sum_{t\in\mathcal{T}} \hat{I}_t w_{t,a} + C_a \right) + \sum_{t\in\mathcal{T}} \sum_{a\in\mathcal{A}} (-\mu_{a,t} w_{t,a}) \Bigg] \\
&= -\sum_{t\in\mathcal{T}} \nu_t + \sum_{a\in\mathcal{A}} \lambda_a C_a + \inf_{\mathbf{w}} \Bigg[\sum_{t\in\mathcal{T}} \sum_{a\in\mathcal{A}} w_{t,a} \ln w_{t,a} + \sum_{t\in\mathcal{T}} \nu_t \sum_{a\in\mathcal{A}} w_{t,a} \\
&\quad - \sum_{a\in\mathcal{A}} \lambda_a \sum_{t\in\mathcal{T}} \hat{I}_t w_{t,a} + \sum_{t\in\mathcal{T}} \sum_{a\in\mathcal{A}} (-\mu_{a,t} w_{t,a}) \Bigg] \\
&= -\sum_{t\in\mathcal{T}} \nu_t + \sum_{a\in\mathcal{A}} \lambda_a C_a - f^*\left(-\nu_t + \hat{I}_t \lambda_a + \mu_{t,a}\right) \\
&= -\sum_{t\in\mathcal{T}} \nu_t + \sum_{a\in\mathcal{A}} \lambda_a C_a - \sum_{t\in\mathcal{T}} e^{-\nu_t - 1} \sum_{a\in\mathcal{A}} e^{\hat{I}_t \lambda_a + \mu_{t,a}}
\end{aligned}
$$

Chapter 3

Market Design for Mobile Application Stores

I analyze the markets used in mobile device application stores, the largest of which is currently the Apple iPhone App Store. In these markets, software developers submit their applications to be sold in the store and set the price paid by consumers. The application store takes a fixed percentage of each sale and assigns a ranking based on parameters like download volume, price, and age of the application.

In this chapter, I study how the application store should set its ranking policy to maximize its own revenue. I characterize the conditions under which there are no pure strategy Nash equilibria in product prices. Given that these conditions are likely to exist in real-world mobile application stores, I would like to understand how the choice of ranking function influences the set of prices through which bidders cycle. Using agent-based simulation, I am able to study bidding cycles and compare ranking functions.

I also present empirical data from the iPhone App Store both to illustrate some of the general features of this market and to validate the assumptions used in the simulations.

3.1 Introduction

The development of mobile application stores in 2008 dramatically reshaped the mobile application market. Prior to this point, mobile software application developers seeking to distribute their products at scale to customers had to strike deals with network operators, network by network, and country by country. They often had to develop the same software for tens or even hundreds of mobile devices with differing technical capabilities. All of these factors greatly limited software distribution to mobile devices.

On July 11, 2008, Apple launched the iPhone App Store, and as of this writing, over one billion applications (free and paid) have been downloaded[1]. Apple shipped over 10 million iPhones in 2008 to over 50 countries[2]. Industry leaders that are also starting their own mobile application stores include Google[3], Nokia[4], Microsoft[5], Research in Motion (BlackBerry's manufacturer)[6], T-Mobile[7], and Palm[8]. For the first time, individual software developers have the ability to create mobile software and easily distribute their products to millions of customers in multiple countries.

[1] Apple website, http://www.apple.com/itunes/billion-app-countdown/, retrieved 4/30/2009
[2] Apple earnings call, 10/21/08
[3] see Android Market, http://www.android.com/market/
[4] see Nokia Ovi Store
[5] see Microsoft Skymarket
[6] see BlackBerry App World
[7] see T-Mobile App Store
[8] see Palm Software Store, http://appstore.pocketgear.com/palm/

Successful developers programming in the evenings and weekends for the iPhone have earned enough to quit their day jobs[9], and a number of companies have been formed specifically to develop iPhone applications (e.g., Smule, Pelago).

Of interest to economists is the fact that mobile application stores are a new type of market for study. Software developers submit their applications to be sold in the store and set the price paid by consumers. The application store takes a fixed percentage of each sale[10] and has the freedom to list the application anywhere in the store. This application ranking is generally done algorithmically based on a number of possible parameters including sales volume, price, and age of the application.

An open policy question revolves around how the application store should determine its ranking algorithm to maximize its revenue. The two extremes I will study in this chapter are rank-by-download and rank-by-revenue. In a related mechanism used in online sponsored search advertising, search engines rank by expected revenue (Edelman et al. (2007)), which is the product of bid price and expected clickthrough rate (CTR). The crucial difference between these two settings is that with mobile application stores, the download rate (which is the analog of CTR in this setting) is a function of price since the end consumer is the one paying the price, not the bidder in the auction.

In this chapter, I present a model of the mobile application store and consider how the store sponsor (e.g., Apple) might design its ranking algorithm to maximize

[9]For example, $4.99 iPhone game Trism earned its one developer $250,000 in profits in its first two months, http://www.cnn.com/2008/TECH/11/18/iphone.game.developer/index.html

[10]For each application sold in the iPhone App Store, Apple keeps a revenue share of 30%, http://developer.apple.com/iPhone/program/distribute.html. Google also takes 30% of each sale in the Android Market, http://android-developers.blogspot.com/2008/10/android-market-now-available-for-users.html

its own revenue. I characterize the conditions under which there are no pure strategy Nash equilibria in product prices. When these conditions hold, the usual analytical tools for equilibrium analysis are unavailable; instead, I use agent-based simulation to compare ranking functions. I model the incentives faced by bidders and simulate how they would respond to different ranking functions.

I also present empirical data gathered directly from the Apple iPhone App Store both to illustrate some of the general features of this market and to validate the assumptions used in the simulations.

3.1.1 Related Work

I compare mobile app stores with sponsored search (Edelman et al. (2007)) and dynamic aggregation websites such as digg.com (Wu and Huberman (2008)). There are three dimensions on which comparisons can be made:

1. how applications/ads/websites are ordered,
2. how users respond to the choices with which they are presented, and
3. how advertisers/developers bid or interact with the market.

Mobile app stores currently have "Top 100" lists that order applications based on popularity and may be experimenting with price and other factors; in this chapter, I model ranking based on popularity and revenue. Users purchase applications based on application price, quality, and ranking. Application developers seek to maximize their revenue by investing the appropriate amount of time into application development and by selecting an appropriate application price.

In sponsored search (Edelman et al. (2007)), search engines rank advertisements based on expected revenue, which is the product of expected clicks and price. Users click on ads based on the position of the ad and the quality of the match between their search and the ads displayed. Advertisers submit bids based on expected sales once a user clicks through to their website.

In Wu and Huberman (2008), the authors explore three ranking rules for dynamic aggregation websites: 1) novelty, 2) popularity, and 3) expected clicks (defined as the product of past popularity and a novelty decay factor). The goal is to pick a ranking mechanism that maximizes clicks over time. Users click on links to websites based on quality (which includes novelty) and link position. Unlike the app store and sponsored search settings, the websites to which links point are not clear strategic actors. Wu and Huberman (2008) find that the best click-maximizing ranking rule depends crucially on the rate of decay of novelty. In this chapter, I will find that the revenue-maximizing ranking rule will depend on the consumer demand and developer cost functions.

I summarize this comparison in Table 3.1.

	ranking rule	user behavior	bidder strategy
app stores	popularity or revenue	quality, price, position	maximize revenue, amortize development costs
sponsored search	revenue	quality, position	keep marginal cost below marginal revenue
dynamic websites	popularity or novelty	quality, position	n/a

Figure 3.1: Comparison between app stores, sponsored search, and dynamic websites

3.2 App Store Model

In this section, I present a model of the app store in which developers select application price and effort in response to 1) expected user demand, 2) the app store ranking rule, and 3) the slot effect. I model the choice of price and effort since I am interested in studying how developers price their products and which market designs encourage developers to create high-quality applications.

3.2.1 Applications

There are N software developers, also referred to as bidders, who each submit 1 application for sale in the app store. Each application i has an associated price p_i and quality q_i. The price is set by the bidder and the quality is determined based on his effort, e_i. A bid is the pair (p_i, e_i) composed of bidder i's price and effort level.

Bidder i expends effort e_i to create his application, resulting in an application of quality $q_i = q_i(e_i)$. This quality function is the single parameter that distinguishes between different developers; it represents a developer's ability to create a quality product from hours invested. Quality is increasing in effort, i.e., $\frac{\partial}{\partial e} q_i(e) > 0$.

Each bidder bears a cost $c(e_i)$ for expending effort e_i. I assume that this cost function is identical across bidders and is increasing in effort, i.e., $\frac{\partial}{\partial e} c(e) > 0$. This models the situation in which the per-hour cost of development time is the same across developers; the difference in cost is due only to the number of hours invested. I also assume that the marginal cost function is increasing, which models the phenomenon in which managing larger projects costs more than smaller projects due to increases in complexity; that is, $\frac{\partial^2}{\partial e^2} c(e) > 0$.

3.2.2 Demand Functions

Consumer demand $D(p_i, q_i)$ for a given application i is a function of the price p_i and quality q_i. $D(p_i, q_i)$ is the demand that application i would receive if it were in the top slot. This is scaled down by the slot effect based on the application's actual ranking (see Section 3.2.4).

I assume that demand decreases as price increases, i.e., $\frac{\partial}{\partial p} D(p, q) < 0$, and demand increases as quality increases, i.e., $\frac{\partial}{\partial q} D(p, q) > 0$. There are no interactions between consumer demand for two competing products.

Throughout this chapter, I will refer to an example demand function $D(p, q) = 1 - \frac{p}{q}$.

3.2.3 Ranking Functions

The application store ranks each application based on revenue, download volumes, or a mixture of the two. In reality, there may be other parameters such as product reviews or developer reputation that are also used in ranking, but this model focuses solely on price and quantity sold.

In particular, the application store assigns a ranking score

$$r(p, q) = D(p, q) * p^{\alpha}$$

to each bidder. Let $\sigma(i)$ be application i's rank, which is based on its score relative to other scores (ranks start from 1 and increase). α is a parameter that allows the application store to vary the ranking function smoothly between rank-by-revenue and rank-by-download. $\alpha = 0$ corresponds to rank-by-download and $\alpha = 1$ corresponds to rank-by-revenue. This chapter will explore the conditions under which the application store would prefer certain values of α over others.

I assume that every single application receives a ranking, i.e., there are N slots. This corresponds to a setting in which the consumer can eventually view each application by scrolling through the entire app store.

3.2.4 Slot Effect

A key characteristic of the mobile application store is the existence of a **slot effect**. This captures the phenomenon in which users are less likely to download an application that is further down in the rankings because they need to expend more effort to scroll down to see this application. In the sponsored search literature, the slot effect has often been modeled using the exponential decay model where slot j has its clickthrough rate decreased by a factor of $\gamma(j) = \frac{1}{\delta^{j-1}}$ for $\delta > 1$. Feng et al. (2007) empirically fit actual online clickthrough data to an exponential decay model with $\delta = 1.428$. In Section 3.6, I will present some empirical data estimating the magnitude of the slot effect in the iPhone App Store.

In this model, the slot effect γ multiplicatively scales down demand, so that demand for bidder i in slot $\sigma(i)$ is $D(p_i, q_i) * \gamma(\sigma(i))$. For example, in Section 3.6, I approximate the slot effect to be around $\gamma = 0.971$. This means that the application in slot 1 receives demand $D(p, q)\gamma(1) = D(p, q) * 0.971^0 = D(p, q)$, the application in slot 2 receives demand $D(p, q)\gamma(2) = D(p, q) * 0.971^1$, and the application in slot 3 receives demand $D(p, q)\gamma(3) = D(p, q) * 0.971^2$.

3.2.5 Application Store Revenue Share

The application store gives each developer a fixed revenue share $s \in [0, 1]$. For example, Apple keeps 30% of all revenue from the App Store and the individual developers keep 70%, so for the iPhone App Store, $s = 0.7$.

3.2.6 Bidder Utility

Bidder i's utility is the difference between his revenue and his cost:

$$u_i(p_i, e_i) = D(p_i, q_i(e_i))\gamma(\sigma(i))p_i s - c(e_i)$$

3.2.7 Platform Revenue

The platform earns revenues:

$$\sum_i D(p_i, q_i)\gamma(\sigma(i))p_i(1 - s)$$

3.2.8 Bidder Strategies

In this market, there is no notion of a bidder's "true value," as the marginal cost of distributing an additional application after it has been produced is zero. In selecting his (price, effort) strategy pair, a bidder cares only about maximizing his profits, measured as the difference between revenue and the original cost of development. Even if a developer thinks his application is "worth" a certain amount, a lower price would increase his demand and could potentially increase his revenue.

For example, in my interview with a top iPhone app developer, the developer originally priced his application at \$0.99 and experimented with repricing at \$1.99.

He discovered that doubling his price decreased his sales volume by more than 50%, so he quickly dropped his price back down to \$0.99 again. Even though he believed his application was delivering far more than \$0.99 of value to the consumer, he selected the \$0.99 price to maximize his profits. My model of the app store is consistent with this bidder logic.

3.3 Examples

I present some 2-bidder examples to illustrate sample conditions when Nash equilibria exist or do not exist. In these examples, I will assume the following functional forms and parameters:

- demand function $D(p, q) = 1 - \frac{p}{q}$
- cost function $c(e) = c * e^3$ with $c > 0$; I will set $c = 0.01$
- quality function $q_i(e_i) = a_i * e_i$ with $a_i > 0$; I will vary a_i
- slot effect γ to vary
- revenue share $s = 0.7$
- ranking function exponent $\alpha = 1$ (rank by revenue)

I will look at the effects of varying quality coefficient a_i and slot effect γ on whether there exists a pure-strategy Nash equilibrium. In each example, there will be 2 bidders with bidder 1 having higher quality than bidder 2; the technical condition for having greater quality is that the marginal quality function is always greater.

Due to Theorem 3.4.1 in the next section, if there is a pure-strategy Nash equilibrium, the bidders will be sorted by quality; i.e., bidder 1 is assigned slot 1 and bidder 2 is assigned slot 2. By Corollary 3.4.2, if there is a pure-strategy Nash equilibrium, each bidder submits a bid targeting this assigned slot; i.e., bidder 1 maximizes his utility assuming he will be assigned slot 1, and likewise for bidder 2. If I can find any bid that yields greater utility than this “Nash equilibrium bid,” then there can exist no pure-strategy Nash equilibrium, by Corollary 3.4.3.

I can determine the Nash equilibrium analytically (if one exists) by finding bidder 1’s optimal bid conditional on winning slot 1 and bidder 2’s optimal bid conditional on winning slot 2. Given the parameters above, the utility function for bidder i is:

$$u_i(p_i, e_i) = \left(1 - \frac{p_i}{a_i e_i}\right) \gamma(\sigma(i)) p_i s - ce^3$$

Conditional on winning slot 1, which has a slot effect of $\gamma(1) = 1$, bidder 1’s utility is $u_1(p_1, e_1) = \left(1 - \frac{p_1}{a_1 e_1}\right) p_1 s - ce^3$.

I find the optimal bid (p_1^*, e_1^*) by examining the first-order necessary conditions for a critical point; I set the partial derivatives of u_1 equal to 0 and solve for p_1 and e_1.

$$\begin{aligned} \frac{\partial u_1}{p_1} = s\left(1 - \frac{2p_1}{a_1 e_1}\right) &= 0 \\ \Leftrightarrow p_1 &= \frac{a_1 e_1}{2} \qquad (3.1) \\ \frac{\partial u_1}{e_1} = p_1 s \left(\frac{p_1}{a_1 e_1^2}\right) - 3ce_1^2 &= 0 \\ \Leftrightarrow e_1 &= \sqrt{\frac{a_1 s}{12c}} \qquad (3.2) \end{aligned}$$

I arrive at (3.2) by substituting $p_1 = (a_1 e_1)/2$. By substituting (3.2) back into (3.1), I can write p_1 in terms of a_1, s, and c. I find that bidder 1's optimal bid is

$$(p_1^*, e_1^*) = \left(\sqrt{\frac{a_1^3 s}{48c}}, \sqrt{\frac{a_1 s}{12c}}\right)$$

In the appendix, I verify that this bid is indeed utility-maximizing for this utility function combined with a more general class of functions q and e.

Similarly, I compute bidder 2's optimal bid (p_2^*, e_2^*) by finding the maximal point for bidder 2's utility function $u_2(p_2, e_2) = \left(1 - \frac{p_2}{a_2 e_2}\right) \gamma p_2 s - ce^3$ where the slot effect for slot 2 is $\gamma(2) = \gamma$. I find that the optimal bid is

$$(p_2^*, e_2^*) = \left(\sqrt{\frac{a_1^3 s\gamma}{48c}}, \sqrt{\frac{a_1 s\gamma}{12c}}\right)$$

To check whether these bids do indeed represent a Nash equilibrium, I will look for bids targeting other slots that yield greater utility. For example, let (p_2', e_2') be bidder 2's optimal bid when targeting slot 1 (instead of slot 2). If bidder 2's utility from bidding (p_2', e_2') is greater than his utility from (p_2^*, e_2^*), then there cannot exist a pure-strategy Nash equilibrium.

In Figure 3.2, (p_i^*, e_i^*) is bidder i's optimal bid when targeting his "appropriate" Nash equilibrium slot; i.e., bidder 1 targets slot 1 and bidder 2 targets slot 2. u_i^* is bidder i's utility when assigned his Nash equilibrium slot. (p_2', e_2') is bidder 2's optimal bid when targeting slot 1 assuming that bidder 1 submits his optimal slot 1 bid (p_1^*, e_1^*). u_2' is bidder 2's utility when assigned to slot 1.

If $u_2' < u_2^*$, then bidder 2 prefers slot 2 to slot 1 and $((p_1^*, e_1^*), (p_2^*, e_2^*))$ is the Nash equilibrium. If $u_2' > u_2^*$, then bidder 2 prefers to target slot 1. By Corollary 3.4.3, this is a situation where there can be no pure-strategy Nash equilibrium.

Example	bidder i	a_i	γ	p_i^*	e_i^*	u_i^*	p_2'	e_2'	u_2'
1	1	1.2	0.9	1.59	2.65	0.370	-	-	-
	2	1		1.15	2.29	**0.241**	1.59	3.17	**0.236**
2	1	1.2	0.85	1.59	2.65	0.370	-	-	-
	2	1		1.11	2.23	**0.220**	1.59	3.17	**0.236**

Figure 3.2: Bids and utilities when quality functions and slot effect are varied

In Example 1, bidder 2 prefers his utility of 0.241 from targeting slot 2 to his utility of 0.236 from targeting slot 1. This means that I have identified a Nash equilibrium. Going from Example 1 to Example 2, the slot effect is changed from 0.9 to 0.85, so the bidder assigned to slot 2 is now penalized more and has a greater incentive to consider targeting slot 1. In this particular example, bidder 2 prefers his utility of 0.236 from targeting slot 1 to his utility of 0.220 from targeting slot 2, so he will target slot 1. Bidder 1 will also fight for slot 1, and cycling ensues, resulting in no pure-strategy Nash equilibrium (by Corollary 3.4.3).

3.4 Characterizing Nash Equilibria

Since the app store market employs a first-price rule, there are cases where bidder strategies would cycle in a dynamic setting and never reach equilibrium. In this section, I examine properties of a Nash equilibrium and the conditions under which a pure-strategy Nash equilibrium exists in a one-shot game of complete information.

For all the theorems in this section, I assume that bidders can be ordered by quality functions $q(e)$. Bidder i has higher quality than bidder j if $\frac{\partial}{\partial e}q_i(e) > \frac{\partial}{\partial e}q_j(e)\ \ \forall e > 0$. In this case, I say that $q_i \succ q_j$. Note that since $q_i(0) = 0$ and $q_j(0) = 0$, this assumption also implies that $q_i(e) > q_j(e)\ \ \forall e > 0$.

First, I show that in any Nash equilibrium, bidders are ordered by quality with the highest-quality bidder assigned to the best slot. Intuitively, since bidders in this model are only differentiated by their quality function, a bidder with higher quality can more easily convert effort into quality so his optimal bid will be at a higher effort level, resulting in higher quality. Since demand is increasing in quality, a higher-quality bidder will be receive a better ranking from the app store.

For all the theorems in this section, I assume that bidders can be ordered by quality functions $q(e)$. Bidder i has higher quality than bidder j if $\frac{\partial}{\partial e} q_i(e) > \frac{\partial}{\partial e} q_j(e) \;\; \forall e > 0$. In this case, I say that $q_i \succ q_j$. Note that since $q_i(0) = 0$ and $q_j(0) = 0$, this assumption also implies that $q_i(e) > q_j(e) \;\; \forall e > 0$.

Theorem 3.4.1. *If a pure-strategy Nash equilibrium exists, then bidders are sorted by quality. That is, the bidder with the i^{th} highest quality is assigned slot i.*

Proof. For this proof, I assume that in addition to having the form $u(p, e) = D(p, q(e))ps - c(e)$, the utility function has exactly one critical point, which is the global maximum; in the appendix, I give an example of functions D, q, and c such that this is true.

Without loss of generality, consider the case of 2 bidders where bidder 1 has higher quality than bidder 2, i.e., $q_1 \succ q_2$. Assume that there exists a Nash equilibrium where the bidders are reverse sorted by quality. Bidder 1 is assigned to slot 2 and bidder 2 is assigned to slot 1.

Let the following dotted variables be unconstrained utility-maximizing bids where the first subscript refers to the bidder and the second subscript refers to the targeted slot:

$$
\begin{aligned}
\dot{b_{11}} = (\dot{p_{11}}, \dot{e_{11}}) &= \arg\max_{(p,e)} D(p, q_1(e))ps - c(e) \\
\dot{b_{12}} = (\dot{p_{12}}, \dot{e_{12}}) &= \arg\max_{(p,e)} D(p, q_1(e))\delta ps - c(e) \\
\dot{b_{21}} = (\dot{p_{21}}, \dot{e_{21}}) &= \arg\max_{(p,e)} D(p, q_2(e))ps - c(e) \\
\dot{b_{22}} = (\dot{p_{22}}, \dot{e_{22}}) &= \arg\max_{(p,e)} D(p, q_2(e))\delta ps - c(e)
\end{aligned}
$$

Let the following be utility-maximizing bids contingent on winning the appropriate slot:

$$
\begin{aligned}
b_{11} = (p_{11}, e_{11}) &= \arg\max_{(p,e)} D(p, q_1(e))ps - c(e) \\
&\quad \texttt{s.t.} \quad D(p, q_1(e))p^{\alpha} \geq D(p_2, q_2(e_2))p_2^{\alpha} \qquad (3.3) \\
b_{12} = (p_{12}, e_{12}) &= \arg\max_{(p,e)} D(p, q_1(e))\delta ps - c(e) \\
&\quad \texttt{s.t.} \quad D(p, q_1(e))p^{\alpha} < D(p_2, q_2(e_2))p_2^{\alpha} \qquad (3.4) \\
b_{21} = (p_{21}, e_{21}) &= \arg\max_{(p,e)} D(p, q_2(e))ps - c(e) \\
&\quad \texttt{s.t.} \quad D(p, q_2(e))p^{\alpha} \geq D(p_1, q_1(e_1))p_1^{\alpha} \qquad (3.5) \\
b_{22} = (p_{22}, e_{22}) &= \arg\max_{(p,e)} D(p, q_2(e))\delta ps - c(e) \\
&\quad \texttt{s.t.} \quad D(p, q_2(e))p^{\alpha} < D(p_1, q_1(e_1))p_1^{\alpha} \qquad (3.6)
\end{aligned}
$$

The reverse-quality Nash equilibrium is (b_{12}, b_{21}).

To simplify notation, I will write utility as

$$u_i(b_{ij}) = u_i(p_{ij}, e_{ij}) = D(p_{ij}, q_i(e_{ij}))p_{ij}\delta^{j-1}s - c(e_{ij})$$

and rank score as

$$r_i(b_{ij}) = r_i(p_{ij}, e_{ij}) = D(p_{ij}, q_i(e_{ij}))p_{ij}^{\alpha}$$

where i is the bidder index and j is the targeted slot.

I can re-write constraints (3.3)-(3.6) using the new notation for the reverse-quality Nash equilibrium as (3.7)-(3.10):

$$r_1(b_{11}) \geq r_2(b_{21}) \tag{3.7}$$

$$r_1(b_{12}) < r_2(b_{21}) \tag{3.8}$$

$$r_2(b_{21}) \geq r_1(b_{12}) \tag{3.9}$$

$$r_2(b_{22}) < r_1(b_{12}) \tag{3.10}$$

First, I show that $\dot{b}_{11}$ and $\dot{b}_{21}$ are infeasible as solutions to the constrained optimizations. Observe that

$$u_1(\dot{b}_{11}) > u_1(\dot{b}_{12})$$

since $\delta < 1$. By the definition of Nash equilibrium, bidder 1 prefers slot 2 to slot 1:

$$u_1(b_{12}) > u_1(b_{11}) \tag{3.11}$$

The only way that (3.11) can be true is if $\dot{b}_{11}$ is infeasible in the constrained optimization, i.e.,

$$r_1(\dot{b}_{11}) < r_2(b_{21}) \tag{3.12}$$

Next, I establish in the following lemma that bidder 1's unconstrained slot 1 bid is ranked above bidder 2's unconstrained slot 1 bid:

Lemma 7.

$$r_1(\dot{b}_{11}) > r_2(\dot{b}_{21}) \tag{3.13}$$

Proof. I expand and re-write the rank score expressions in (3.13) as the following:

$$\begin{aligned} D(\dot{p_{11}}, q_1(\dot{e_{11}}))\dot{p_{11}}^{\alpha} &> D(\dot{p_{21}}, q_2(\dot{e_{21}}))\dot{p_{21}}^{\alpha} \\ \Leftrightarrow \left(u_1(\dot{b_{11}}) + c(\dot{e_{11}})\right)\dot{p_{11}}^{\alpha-1} &> \left(u_2(\dot{b_{21}}) + c(\dot{e_{21}})\right)\dot{p_{21}}^{\alpha-1} \end{aligned} \tag{3.14}$$

Since $q_1 \succ q_2$, bidder 1's utility from bidder 2's unconstrained bid is strictly greater than bidder 2's own utility for this bid: $u_1(\dot{b_{21}}) > u_2(\dot{b_{21}})$. Bidder 1's unconstrained utility-maximizing bid must yield at least this much utility, so $u_1(\dot{b_{11}}) > u_2(\dot{b_{21}})$.

Next, I show that $\dot{e_{11}} > \dot{e_{21}}$ and $\dot{p_{11}} > \dot{p_{21}}$ for the specific demand function $D(p, q) = (1 - p/q)$ and general cost and quality functions $c(e)$, $q_1(e)$ and $q_2(e)$. As before, I assume that marginal cost is increasing, and marginal quality for bidder 1 is greater than marginal quality for bidder 2.

Given this demand function, the utility functions are:

$$\begin{aligned} u_1(p_1, e_1) &= (1 - p_1/q_1(e_1))p_1 s - c(e_1) \\ u_2(p_2, e_2) &= (1 - p_2/q_2(e_2))p_2 s - c(e_2) \end{aligned}$$

To find bidder 1's utility-maximizing bid, I set the first partial derivative with respect to both p and e equal to 0:

$$\begin{aligned} \left.\frac{\partial u_1}{\partial p_1}\right|_{(\dot{p_{11}},\dot{e_{11}})} &= 1 - 2\frac{\dot{p_{11}}}{q_1(\dot{e_{11}})} = 0 \\ \Leftrightarrow \dot{p_{11}} &= \frac{q_1(\dot{e_{11}})}{2} \end{aligned} \tag{3.15}$$

$$\begin{aligned} \left.\frac{\partial u_1}{\partial e_1}\right|_{(\dot{p_{11}},\dot{e_{11}})} &= \frac{\dot{p_{11}}^2}{[q_1(\dot{e_{11}})]^2}\frac{\partial q_1(\dot{e_{11}})}{\partial e} - \frac{\partial c(\dot{e_{11}})}{\partial e} = 0 \\ \Leftrightarrow \left.\frac{\partial q_1}{\partial e}\right|_{\dot{e_{11}}} &= 4\left.\frac{\partial c}{\partial e}\right|_{\dot{e_{11}}} \end{aligned} \tag{3.16}$$

Similarly, for bidder 2, I have:

$$\begin{aligned} \dot{p}_{21} &= \frac{q_2(\dot{e}_{21})}{2} && (3.17) \\ \left.\frac{\partial q_2}{\partial e}\right|_{\dot{e}_{21}} &= 4 \left.\frac{\partial c}{\partial e}\right|_{\dot{e}_{21}} && (3.18) \end{aligned}$$

Since $\frac{\partial}{\partial e} q_1(e) > \frac{\partial}{\partial e} q_2(e)$ and marginal cost is increasing, (3.16) and (3.18) imply that $\dot{e}_{11} > \dot{e}_{21}$. Since cost is increasing, $c(\dot{e}_{11}) > c(\dot{e}_{21})$. Finally, combining (3.15), (3.17), and the fact that $q_1(e) > q_2(e)\ \ \forall e > 0$, I conclude that $\dot{p}_{11} > \dot{p}_{21}$.

I have now established that every term on the left-hand side of (3.14) is greater than the corresponding term on the right-hand side, proving the lemma. □

Combining (3.12) and (3.13), I have that:

$$r_2(b_{21}) > r_2(\dot{b}_{21}) \tag{3.19}$$

Re-stated, $\dot{b}_{21}$ is infeasible in the constrained optimization.

Next, I use the fact that the utility function has exactly one critical point, which is a global maximum. This means that when I start from the unconstrained utility-maximizing bid $\dot{b}_{21}$ and pick bids with greater rank score, these new bids must have lower utility. This implies that the constrained utility-maximizing b_{21} has as low a rank score as possible, equal to the minimum rank score that is feasible, $r_1(b_{12})$.

I prove this formally in the next lemma:

Lemma 8.

$$r_2(b_{21}) = r_1(b_{12}) \tag{3.20}$$

Proof. Assuming that (3.20) is not true, there exists a bid b^* with the following two properties:

1. $r_2(b^*) > r_1(b_{12})$

2. $u_2(b^*) > u_2(b)$ for any b where $r_2(b) \geq r_1(b_{12})$

Since u has only one critical point, which is the global maximum, the arc A from $u_2(\dot{b}_{21})$ to $u_2(b^*)$ is monotonically decreasing. By continuity of r, there exists a bid b' in the neighborhood of b^* on arc A such that $r_2(b') \geq r_1(b_{12})$. However, since arc A is monotonically decreasing from $u_2(\dot{b}_{21})$ to $u_2(b^*)$, b' must have higher utility than b^*: $u_2(b') > u_2(b^*)$. This contradicts property 2 regarding b^*, so b^* cannot exist. □

Since (3.8) and (3.20) cannot both be true, a contradiction has been reached, and bidders cannot be sorted by reverse-quality in a Nash equilibrium. □

Corollary 3.4.2. *Let bidders be numbered so that higher-quality bidders have lower indices (e.g., bidder 1 has the highest quality). If a pure-strategy Nash equilibrium exists, then the equilibrium bid vector is $(\dot{b}_{11}, \ldots, \dot{b}_{NN})$, where*

$$\dot{b}_{ii} = \arg\max_{(p,e)} D(p, q_i(e))p\delta^{i-1} - c(e)$$

That is, bidder i targets slot i and bids his unconstrained utility-maximizing bid.

Proof. From Theorem 3.4.1, bidders must be sorted by quality in a pure-strategy Nash equilibrium. This means that the bidder i is assigned slot i. Contingent on being assigned slot i, bidder i's optimal bid is to maximize utility assuming that he will be assigned slot i: $max\ \ u_i(b_{ii}) = D(p_{ii}, q_i(e_{ii}))p_{ii}\delta^{i-1}s - c(e_{ii})$. This is exactly the unconstrained utility-maximizing bid $\dot{b}_{ii}$. □

Corollary 3.4.3 is the contrapositive of the Corollary 3.4.2.

Corollary 3.4.3. *If there exists a bid for bidder i that gives the bidder greater utility than his unconstrained utility-maximizing bid* $\dot{b_{ii}} = \arg\max_{(p,e)} D(p, q_i(e))p\delta^{i-1} - c(e)$, *then there can exist no pure-strategy Nash equilibrium.*

In cases where bidders can be ordered by quality, there are two implications of Theorem 3.4.1, Corollary 3.4.2, and Corollary 3.4.3:

1. Whether or not a pure-strategy Nash equilibrium exists can be computed easily.

2. If a pure-strategy Nash equilibrium exists, it can be computed easily.

The following steps can be used to compute the Nash equilibrium or determine that no pure-strategy Nash equilibrium exists:

1. Assuming bidders are numbered so that higher-quality bidders have lower indices, for each bidder i, compute the unconstrained utility-maximizing bid targeting slot i:

$$\dot{b_{ii}} = \arg\max_{(p,e)} D(p, q_i(e))p\delta^{i-1} - c(e)$$

2. For each bidder i, compute the unconstrained utility-maximizing bid targeting slot j for all $j \neq i$:

$$\dot{b_{ij}} = \arg\max_{(p,e)} D(p, q_i(e))p\delta^{j-1} - c(e)$$

3. If $\dot{b_{ij}} > \dot{b_{ii}}$ for any (i, j) pair (i.e., bidder i prefers targeting slot j over targeting his appropriate slot i), then there exists no pure-strategy Nash equilibrium. Otherwise, the pure-strategy Nash equilibrium is $(\dot{b_{11}}, \ldots, \dot{b_{NN}})$.

3.5 Simulation

In this section, I simulate bidders participating in a dynamic app store game. As shown in the previous section, there are conditions under which no pure strategy Nash equilibria exist, but I would still like to be able to analyze how bidders respond to differing market designs. In particular, I allow bidders to best-respond to each other and record their bids and app store revenue over the course of a bidding cycle.

In Simulation 1, I plot app store revenue, and price/effort levels as the ranking function α is varied in Examples 1 and 2 from Section 3.3. In cases where there is no Nash equilibrium, I record average values over the entire bidding cycle.

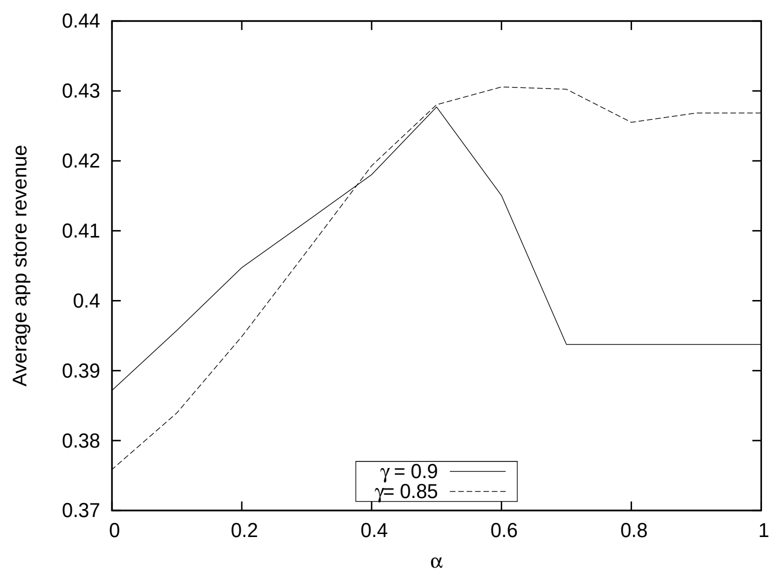
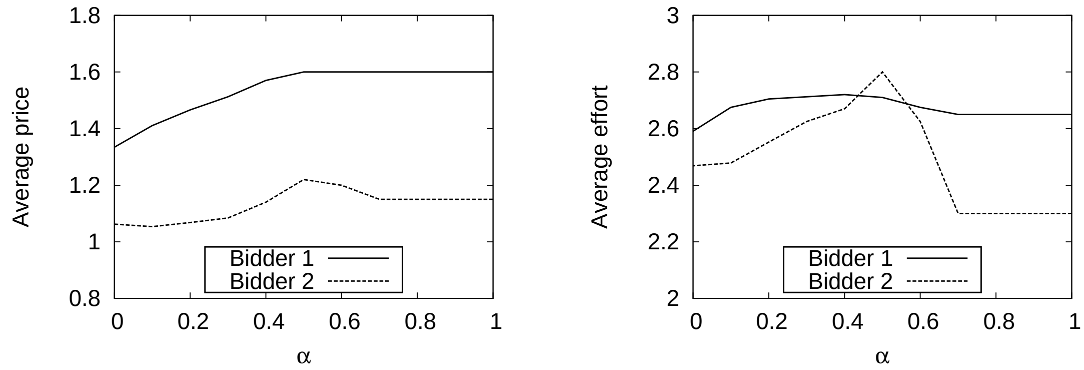


Figure 3.3: Simulation 1: average app store revenue for Examples 1 and 2, $\gamma = 0.9$, 0.85

In Example 1 ($\gamma = 0.9$), for α between 0.7 and 1, I find that there is a Nash equilibrium that is independent of α. For high values of α, there are Nash equilibria where bids are stable, but below $\alpha = 0.7$, bidding becomes unstable and there are no



(a) Average price

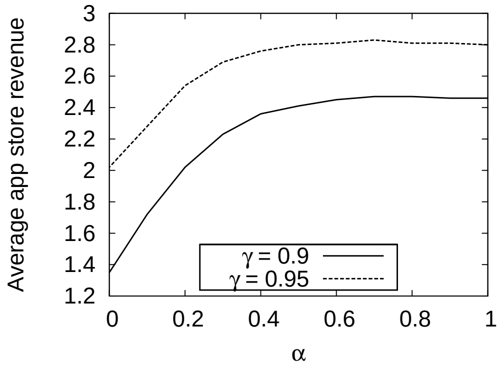
(b) Average effort

Figure 3.4: Average price and average effort for Simulation 1: Example 1, $\gamma = 0.9$

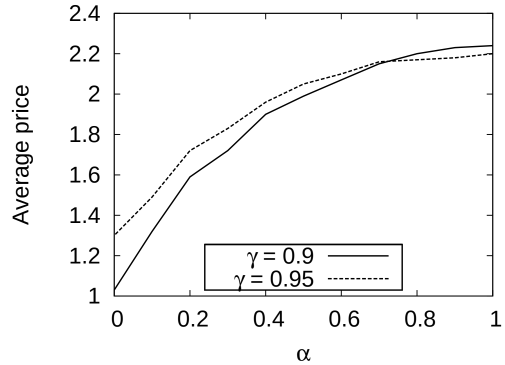
Nash equilibria.

In Figure 3.3, for values of α between 0.7 and 1, the Nash equilibrium yields constant revenue for the app store. When α drops below 0.7, app store revenue starts rising until the peak at $\alpha = 0.5$. The price and effort graphs in Figures 3.4(a) and 3.4(b) demonstrate that this increase in revenue is largely due to a large increase in effort by bidder 2. For α near 0.5, bidder 2 expends more effort in order to fight for the top slot over the course of the bid cycle.

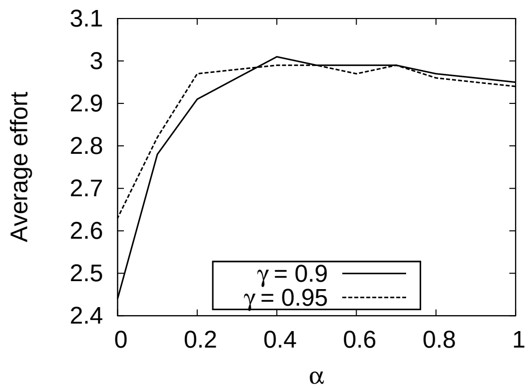
From the point of view of the market designer, it is unclear from Figure 3.3 how to choose α. There appears to be an optimal α greater than 0 and less than 1, but this peak depends on the particular conditions (such as the magnitude of the slot effect and the relative abilities of the bidders). However, these examples do suggest that rank-by-revenue ($\alpha = 1$) and rank-by-download ($\alpha = 0$) are not revenue-maximizing for the app store. These simulations suggest that a market designer may opt to pick α to be high so that a Nash equilibrium is more likely, but not as high as 1 since it



(a) Average app store revenue



(b) Average price



(c) Average effort

Figure 3.5: Average revenue, average price, and average effort for Simulation 2

appears that revenue may be increased by picking smaller values of α.

Next, I present some larger scale simulations with 10 bidders. In Simulation 2, the 10 bidders have quality coefficients $a = [1.1, 1.2, \ldots, 1.9]$. I examine slot effects $\gamma = .9$ and $\gamma = .95$.

In Figure 3.5(a), the app store maximizes revenue by picking an α of around 0.7 in each case. The exact value of α that maximizes revenue appears to depend on the particular distribution of abilities, but these simulations show that revenue for $\alpha = 1$

seems to be higher than revenue for $\alpha = 0$, with a maximum somewhere in between.

In Figure 3.5(b), average prices rise monotonically as α increases. At the extremes, this means that I would expect to see higher average prices when the app store ranks by revenue than when it ranks by popularity. This coincides with intuition since as α rises, the app store favors higher prices, and I would expect to see average prices rise.

In Figure 3.5(c), effort levels are lower for low values of α and higher for $\alpha > .3$. The specific shape of this curve depends heavily on the demand and cost functions; the functions I have studied tend to result in an effort curve that reaches a maximum away from the extremes of $\alpha = 0$ and $\alpha = 1$. This suggests that the app store may want to pick an α value such as 0.5 to create incentives for higher-quality application development, although more research would be necessary to support such a recommendation.

3.6 Empirical Work

In this section, I present empirical data from the iPhone App Store both to illustrate some of the general features of this market and to validate the assumptions used in the simulations.

3.6.1 iPhone App Store Background

My iPhone App Store data was gathered starting late December 2008, so I describe the user experience during this period. On November 21, 2008, Apple released its version 2.2 firmware for the iPhone and iPod Touch; this update affected the App

Store by allowing users to sort applications by "Top Paid," "Top Free," and "Release Date."

When a user launches the iPhone App Store from his iPhone, he can view applications by category (e.g., Entertainment, Games, or Productivity) or by selecting the "Top 25" list. For each application, the user sees cover art, application title, category, price, and release date, and can click through to read a longer description of the application and see other users' reviews.

3.6.2 App Store Data

The raw data for this analysis comes from the XML used by Apple's desktop iTunes client. For 9 days in the period from December 23, 2008 to January 7, 2009, I collected data once a day for every application available in the iPhone App Store (free or paid). For each application, I recorded 1) price, 2) popularity, and 3) days since release. "Popularity" is a field in the iTunes XML between 0 and 1 and for this analysis, I use it as a proxy for sales volume. I assume that the popularity scores are correct up to a scaling factor; that is, if Application A has a popularity score twice that of Application B, then Application A has double the sales.

I separately recorded Apple's overall Top 50 list on each of these days. When viewing the App Store from an iPhone device, the Top 25 list is prominently featured with a link to the Top 50 list, meaning that the Top 50 applications receive much visibility. The Top 50 list changes daily and many developers have observed that an application's specific rank in the list has a large effect on sales volume, so the Top 50 list is a natural place to begin studying the magnitude of the slot effect.

3.6.3 App Store Prices

Here, I present summary statistics for overall prices set by developers in the iPhone App Store. Note that Apple constrains pricing for paid applications so that applications must cost \$0.99, \$1.99, \$2.99, and so on. This means that the lowest possible price that can be charged for a paid application is \$0.99. Figure 3.6 is a histogram of all applications that have prices lower than \$20 available on January 2, 2009. This sample represents 98.5% of the total 12,477 applications.

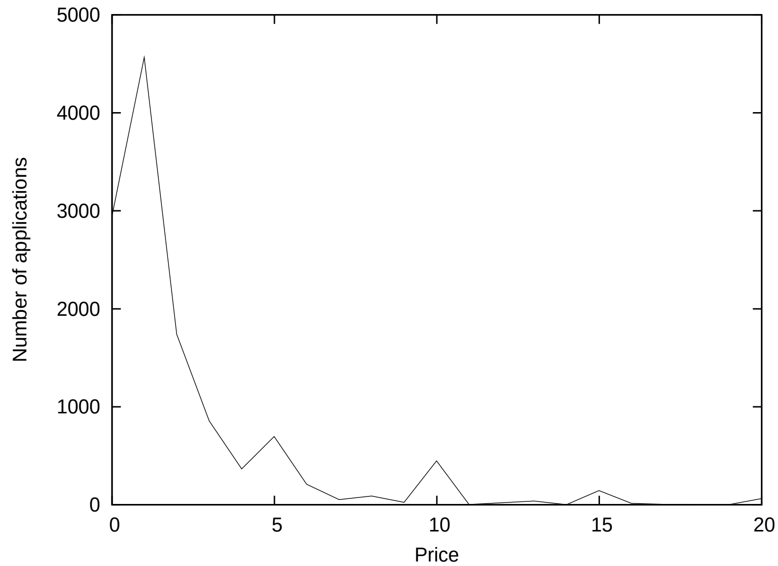


Figure 3.6: Distribution of applications over price for free applications and applications less than \$20, iPhone App Store, 1/2/09.

On this day, free applications accounted for 23.6% of all applications, the average price overall was \$2.21 and the average price of paid applications was \$2.90.

Figure 3.7 shows the distribution of applications over prices for applications under \$10, representing 95% of all paid applications.

As can be seen from Figures 3.6 and 3.7, many developers are setting low prices,

Price	Number of apps	%
\$0.99	4567	47.9
\$1.99	1741	18.3
\$2.99	856	9.0
\$3.99	366	3.8
\$4.99	696	7.3
\$5.99	209	2.2
\$6.99	53	0.6
\$7.99	89	0.9
\$8.99	25	0.3
\$9.99	448	4.7
\$10+	1093	5.0
Total	9528	100.0

Figure 3.7: Distribution of applications over price for paid applications, iPhone App Store, 1/2/09.

with almost half of all paid applications selling for the minimum paid price of \$0.99. There are also spikes at multiples of \$5, probably because these are natural anchor points for consumers.

3.6.4 Application Popularity and the Slot Effect

In this section, I examine the key drivers behind application popularity. Using the iTunes XML dataset, the most readily available data are: 1) price, 2) rank, and 3) days since release. In Figure 3.8, I plot the natural logarithm of popularity against rank for all applications in the Top 50 during the 9-day period, resulting in 450 data points. Using a linear fit, I show the average values as the red line and the 95% prediction interval as the green lines.

In Figure 3.9, I present the results from three linear regressions predicting log-popularity where I pooled data from all 9 days. The results from a fixed effects OLS regression with date dummy variables are very similar. Model 1 regresses log-

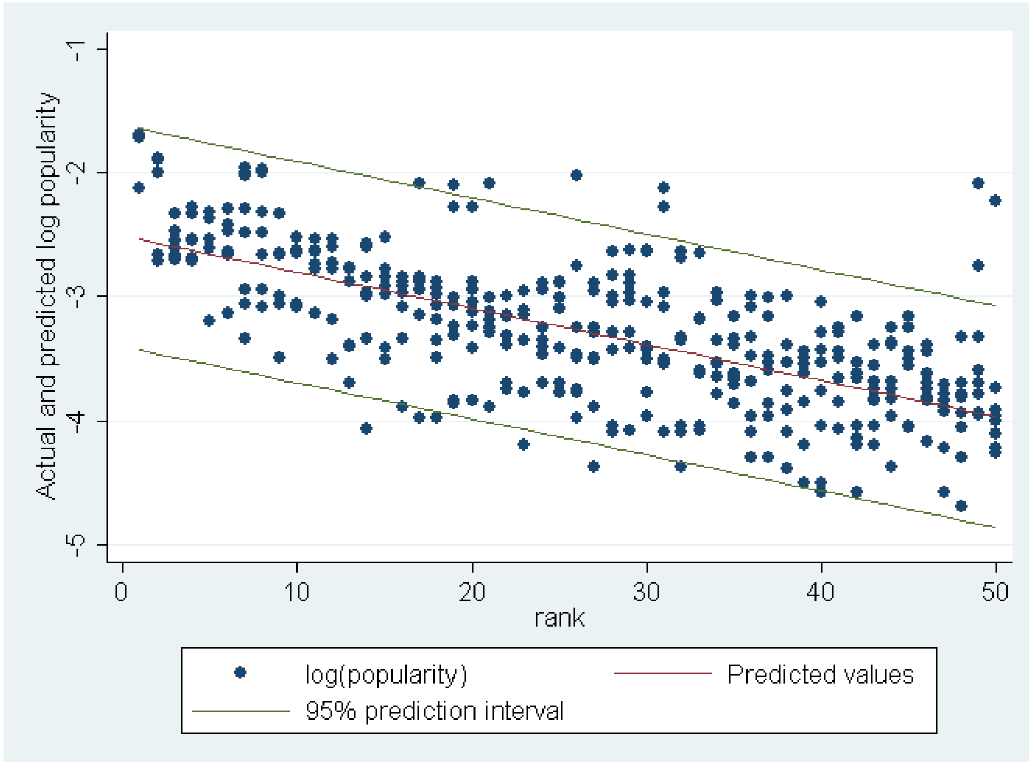


Figure 3.8: Point and interval predictions demonstrating relationship between log-popularity and rank, iPhone App Store, 12/23/08-1/7/09.

popularity on rank, which is the scatter plot shown in Figure 3.8. Model 2 adds log-price as an independent variable, and Model 3 adds the number of days since release. Every coefficient is significant at the 99.9% level.

I find that in the iPhone App Store, application ranking is the single most important factor in predicting popularity. This is unsurprising given the constraints of shopping for applications on a mobile device–it seems reasonable that if users need to scroll down a couple pages to see an application, they will have a lower likelihood of purchasing the application. What may be surprising is how important rank is; in particular, I can see from the R^2 value of Model 1 in Figure 3.9 that rank alone explains 46.3% of the variation in log-popularity.

Variable	(1)	(2)	(3)
Rank	-0.029***	-0.028***	-0.027***
	(0.001)	(0.001)	(0.001)
Ln Price		-0.133***	-0.139***
		(0.024)	(0.024)
Days			0.0013***
			(0.0004)
Constant	-2.512***	-1.500***	-1.562***
	(0.043)	(0.189)	(0.187)
R^2	0.463	0.497	0.511
N	450	450	450

***p < .001
Standard errors are in parentheses.

Figure 3.9: OLS models of log-popularity, iPhone App Store Top 50 List, 12/23/08-1/7/09

Since the dependent variable is log-popularity, I can interpret the coefficient -0.029 on the rank variable in Model 1 as the percentage change in popularity for each unit increase in rank[11]. Re-stated, if an application increases in rank by 1 (i.e., its rank becomes worse by 1), its expected popularity will decrease by 2.9%. This corresponds exactly to the slot effect γ_j described in Section 3.2.4.

Similarly, in Model 2, a 1% increase in price will result in a 13.3% drop in expected popularity.

3.6.5 Data limitations

There may be potential concerns about simultaneity in these regressions since Apple states it takes popularity into consideration when determining ranking for the

[11]Taking the Taylor expansion of $\ln(1+x)$ near $x=0$, $\ln(1+x) \approx x$ for small x. For regression $\ln y = \beta x$, if x is increased by 1, the left-hand side becomes $\ln y + \beta \approx \ln y + \ln(1+\beta) = \ln[y(1+\beta)]$. Thus, a unit increase in x corresponds approximately to a percentage increase in y.

App Store. There is a positive feedback loop wherein an application that sells well gets picked up by the ranking algorithm and placed in a highly visible rank, leading to even further sales.

Next, the "popularity" field in the iTunes XML is undocumented so I cannot be certain that it is correct to a scaling factor as assumed in the above analysis. To address some of these concerns, I present some download versus ranking data made publicly available by the creator of the top-selling iFart Mobile application[12]. In Figure 3.10, I show the number of units sold accompanied by ranking for each available data point when iFart made it in the Top 100 list.

Ranking	Units sold
1	13274
2	9760
4	5497
9	3117
10	3086
15	2836
22	1797
39	1510
76	841

Figure 3.10: Downloads and ranking for iFart application, iPhone App Store, 12/14/08-12/22/08.

If I repeat the OLS regression of log-popularity on rank from Section 3.6.4 for this single application, I find the following:

$$\ln(\texttt{popularity}) = -0.031 * \texttt{rank} + 8.71$$

These coefficients are significant at the 99% level and the R^2 is 71.4%. The coefficient on rank of -0.031 is very close to the coefficient from the pooled OLS

[12]Data available at: http://www.joelcomm.com/updated_iphone_app_sales_the_f.html. The price of this application remained constant at $0.99 throughout this period

regression of -0.029 in Section 3.6.4, lending confidence to my interpretation of the "popularity" field in the iTunes XML.

3.7 Conclusion

In this chapter, I have studied market design in mobile app stores. I have presented a model of this market where developers select application price and effort in response to expected user demand and the app store ranking rule. I analytically characterize the conditions under which there are no pure-strategy Nash equilibria, and prove that bidders are sorted by quality if there exists a Nash equilibrium.

Using agent-based simulation, I find that the optimal choice of ranking rule lies somewhere between the extremes of ranking by revenue and ranking by popularity, and depends on characteristics of the developer population and consumer demand.

I also present empirical data from the iPhone App Store to illustrate some of the general features of this market and to validate the assumptions used in the simulations.

3.8 Appendix to Chapter 3

I give an example of a utility function of the form $u(p, e) = D(p, q(e))ps - c(e)$ that has exactly one critical point, which is the global maximum.

Let the demand function be $D(p, q) = 1 - p/q$. I will characterize the conditions for functions q and c such that the utility function has one critical point, the global maximum.

The first-order necessary conditions for a critical point are:

1. $\frac{\partial u}{\partial p} = 0$

2. $\frac{\partial u}{\partial e} = 0$

The second-order sufficient conditions for the point (p^*, e^*) to be a local maximum are:

3. $\frac{\partial^2 u}{\partial p^2} < 0$

4. $\frac{\partial^2 u}{\partial p^2}\frac{\partial^2 u}{\partial e^2} - \left(\frac{\partial^2 u}{\partial p \partial e}\right)^2 > 0$

where (3) and (4) are both evaluated at (p^*, e^*).

For my particular utility function, (1) becomes:

$$\begin{aligned} \frac{\partial u}{\partial p} = s\left(1 - \frac{2p^*}{q(e^*)}\right) &= 0 \\ \Leftrightarrow p^* &= \frac{q(e^*)}{2} \end{aligned} \tag{3.21}$$

(2) becomes:

$$\begin{aligned} \frac{\partial u}{\partial e} = -c'(e^*) + \frac{(p^*)^2 s q'(e^*)}{(q(e^*))^2} &= 0 \\ \Leftrightarrow c'(e^*) &= \frac{s}{4}q'(e^*) \end{aligned}$$

(3) becomes:

$$\frac{\partial^2 u}{\partial p^2} = \frac{-2s}{q(e^*)}$$

This is always less than 0 since q is always positive.

(4) becomes:

$$\frac{\partial^2 u}{\partial p^2}\frac{\partial^2 u}{\partial e^2} - \left(\frac{\partial^2 u}{\partial p \partial e}\right)^2 = \frac{2s\left([q(e^*)]^2 c''(e^*) - (p^*)^2 s q''(e^*)\right)}{[q(e^*)]^3}$$

This expression is positive if the numerator is positive. That is, if:

$$\begin{aligned}
[q(e^*)]^2\, c''(e^*) - (p^*)^2 s q''(e^*) &> 0 \\
\Leftrightarrow [q(e^*)]^2\, c''(e^*) - \left(\frac{q(e^*)}{4}\right)^2 s q''(e^*) &> 0 \qquad (3.22) \\
\Leftrightarrow c''(e^*) &> \frac{s}{4} q''(e^*) \qquad (3.23)
\end{aligned}$$

In (3.22), I substitute (3.21).

If c'' and q'' are such that (3.23) is true, then condition (4) is met and (p^*, e^*) is a local maximum.

Next, (3.21) has only one solution, so this utility function has one critical point. This means that the local maximum I identified must also be the global maximum.

To summarize, when $D(p, q) = 1 - p/q$, utility function $u(p, e) = D(p, q(e))ps - c(e)$ has one critical point, which is the global maximum, and functions q and c satisfy (3.23).

Bibliography

Abrams, Z., O. Mendelevitch, and J. Tomlin. 2007. Optimal delivery of sponsored search advertisements subject to budget constraints. In *ACM Conference on Electronic Commerce*.

Abrams, Z., and E. Vee. 2007. Personalized ad delivery when ads fatigue: An approximation algorithm. In *WINE 2007*, 535–540.

Boyd, S., and L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press.

Chickering, D. M., and D. Heckerman. 2000. Targeted advertising with inventory management. In *ACM Conference on Electronic Commerce*, 145–149.

Edelman, B., and M. Ostrovsky. 2007. Strategic bidder behavior in sponsored search auctions. *Decision Support Systems* 43(1):192–198.

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

Eisenmann, T., G. Parker, and M. V. Alstyne. 2006. Strategies for two-sided markets. *Harvard Business Review* 84(10).

Feldman, J., S. Muthukrishnan, M. Pal, and C. Stein. 2007. Budget optimization in search-based advertising auctions. In *ACM Conference on Electronic Commerce*.

Feng, J., H. K. Bhargava, and D. M. Pennock. 2007. Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms. *Informs Journal on Computing* 19:137–148.

Gonen, R., and E. Pavlov. 2007. An incentive-compatible multi-armed bandit mechanism. In *WWW*.

Johari, R., and J. N. Tsitsiklis. 2004. Efficiency loss in a network resource allocation game. *Mathematics of Operations Research* 29(3):407–435.

———. 2007. Efficiency of scalar-parameterized mechanisms. *Operations Research* Submitted.

Kelly, F. P. 1997. Charging and rate control for elastic traffic. *European Transactions on Telecommunications* 8:33–37.

Lahaie, S., D. Pennock, A. Saberi, and R. Vohra. 2007. Sponsored search. In *Algorithmic Game Theory*, ed. N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, chap. 28.

Mas-Colell, A., M. Whinston, and J. Green. 1995. *Microeconomic Theory*. Oxford University Press.

Pandey, S., and C. Olston. 2006. Handling advertisements of unknown quality in search advertising. In *Neural Information Processing Systems*.

Parkes, D. C., and T. Sandholm. 2005. Optimize-and-dispatch architecture for expressive ad auctions. In *Proceedings of First Workshop on Sponsored Search Auctions*.

Rochet, J.-C., and J. Tirole. 2006. Two-sided markets: A progress report. *RAND Journal of Economics* 37(3):645–667.

Sanghavi, S., and B. Hajek. 2004. Optimal allocation of a divisible good to strategic buyers. In *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2748–2753.

Shenker, S. 1990. A theoretical analysis of feedback flow control. In *Proceedings of the ACM SIGCOMM*, 151–160.

Tomlin, J. 2000. An entropy approach to unintrusive targeted advertising on the web. In *9th International World Wide Web Conference (WWW9)*, 767–774.

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

Wu, F., and B. Huberman. 2008. Popularity, novelty, and attention. In *ACM Conference on Electronic Commerce*.