

The Impact of Search Engine Optimization on Online Advertising Market

Bo Xing

The Rawls College of Business Administration
Texas Tech University
Lubbock, TX 79409-2101
+1-806-742-2397
bo.xing@ttu.edu

Zhangxi Lin

The Rawls College of Business Administration
Texas Tech University
Lubbock, TX 79409-2101
+1-806-742-1926
zhangxi.lin@ttu.edu

ABSTRACT

Online advertising market is becoming a popular area of academic research. Among other types of advertising, search engine advertising is leading the growth in terms of revenue. In general, there are two types of search engine advertising: paid placement and search engine optimization (SEO). This study aims to analyze the condition under which SEO exist and further, its impact on the advertising market. With an analytical model, several interesting insights are generated. The results of the study fill the gap of SEO in academic research and help managers in online advertising make informed advertising decisions.

Categories and Subject Descriptors

General Terms

Keywords

search engine, online advertising, search engine marketing, search engine optimization, sponsored links, paid placement

1. INTRODUCTION

In the last decade, there has been a tremendous surge of interest in Internet search engines. Fueling this trend is the fact that Internet search engines have become popular both as information-seeking vehicles and as an online advertising media. The 10th WWW User Survey by Georgia Tech University[10] found that Internet search engines are used by 85 percent of the web users. In 2005, online advertising revenue reached \$16.5 billion. Leading search engines have turned profit from advertising income. *Google*, for example, reported record revenues of \$1.578 billion for the

quarter ended September 30, 2005, up 96 percent compared to the third quarter of 2004. Of the \$1.578 billion, \$1.559 billion are related to online advertising.

In general, there are two types of online advertisement associated with Internet search engines: paid placement and Search Engine Optimization (SEO). Paid placement is operated by search engines in the form of sponsored or paid results, where an advertisement is displayed in a pre-specified region of a search result page along with web search results. Search engines charge placement fees tied to the price of the relevant keywords, which is primarily determined by auction and measured by CPC (cost per click), and the number of click-throughs the advertisement receives. SEO, on the other hand, is the practice of optimizing web pages in a way that improves their ranking in the web search results, which are also known as natural or organic results because they are supposed to reflect relevancy in searchers' standard. In this type of advertisement, advertisers pay SEO firms which specialize in this practice.

Recently, SEO is gaining momentum primarily for two reasons. First, CPC has increased tremendously over years. According to a Fathom Online report, keyword cost has risen 19% in one year since September 2004[8]. Second, it has been realized that organic results are more appealing to searchers because these results are considered more objective and unbiased than sponsored results. According to an online survey by Georgia Tech University[10], over 70% of the search engine users prefer clicking organic results to sponsored results. The SEMPO survey[17] concurs with this finding, showing that organic listings are chosen first by 70% of the people viewing search results, while sponsored listings receive about 24.6% of clicks.

As a new phenomenon, while SEO is drawing much of the attention in online advertising industry, there has been very few published academic research in this area. To address this gap, this study aims to understand the impact of SEO on online advertising market. Two research questions naturally arise. Under what condition will SEO firms survive? How do SEO and other factors impact search engine profit? These questions are important to both search engines and SEO firms, because they share the revenue of the online advertising market. Insights in answering these questions will help managers of both entities make informed strategic decisions. Since SEO is a practice specific to search

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '04, Month 1–2, 2004, City, State, Country.
Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.

results, other search engine feature, such as page design, reputation and personalization are beyond the scope of this study.

With an analytical model, this study approaches both questions from the perspective of advertisers' choice of advertising vehicle, the value of which depends on search engine quality in a market where indirect network externality exists. The rest of the study is structured as follows. In chapter 2, we will introduce the background of Internet search engine and review related literature on search engine quality. Chapter 3 features an analytical model of search engine advertising market, its related propositions, and results. In the subsequent chapter, we will discuss the implications of these findings and conclude the article with limitations and future suggestions.

2. BACKGROUND AND RELATED WORKS

As both information retrieval vehicles and an advertising media, Internet search engines have several unique characteristics. In the first place, Internet search engines are generally free to use, implying that the use of one search engine does not by itself exclude the use of other search engines. This feature has been observed in a study of search engine market structure by Telang *et al.*[23], who posit that this feature inflates the demand of search market by allowing users to sample multiple engines. This multiple sampling of product is not only feasible but also rational because user satisfaction with a particular search session is stochastic[23] due to the difference in search engine design and heterogeneous user preferences. Empirical studies has found that surfers consult a second search engines 22% of search sessions[22]. Meta-search engines such as Dogpile are becoming popular by including, combining, and re-ranking top results of major search engines. Defining the quality of a search engine as the probability that it satisfied searchers in a particular search session, Telang *et al.* [23] further explain why lower-quality search engines could enter the market and survive, contrary to the conventional wisdom that lower quality producers have to charge lower prices to compensate for inferior quality.

Given zero access fee, search engines strive to improve user satisfaction with search experience through quality search results. The ability of a search engine in doing so largely depends on the page ranking page inclusion. It has been found that page ranking impacts user satisfaction[25]; each search engine only contains a fraction of the index-able information on the Internet[15]. Page ranking affects search engine quality because it has been observed that in real world, most of the users consume only top ranked results due to cognitive limitations, time constraints and other factors[14]. This finding has also been reproduced in laboratory setting[26]. Furthermore, this diminishing attention of search engine users appears to follow an exponential decay over rank[3]. User satisfaction, therefore, is primarily determined by the quality of top results. At present, there are considerable gaps among different search engines in terms of top results, because there is very little overlap among major search engines in terms of their first two result pages[7].

The presence of SEO is a distinct characteristic of Internet advertising market. However, its impact on user satisfaction is ambiguous. For a low quality search engine, SEO firms may

actually boost the ranking of a link which improves overall user satisfaction. For a high quality search engine, on the other hand, SEO firms are often regarded as spam, because they could boost the ranking of a link which decreases user satisfaction at large. Overall, SEO introduces additional "noises" to the Internet content and challenge to content inclusion and page ranking of search engines. Besides, it is a deliberate attempt at manipulating the page ranking the search engine. In this study, instead of modeling the direct effect of SEO on user satisfaction, we analyze the impact of SEO as an advertising choice from advertisers' perspective. The economic model in the Chapter 4 will expound the direction and properties of this impact.

The impact of SEO has rarely been addressed in published research, except in a theoretical study by Sen[18]. In his work, Sen addresses search engine marketing strategy of an online seller, who, besides setting prices, either chooses paid placement, SEO, or no advertisement all together. Surprisingly, he finds that, in equilibrium, SEO is not an optimal choice, even if SEO fees are not higher than paid placement. Different from his work which concentrates on the transaction of online sellers, this work focuses on advertisers' net payoff from online advertising in the presence of SEO. Interestingly, some findings of this study concur with those of Sen's work.

3. AN INTEGRATED VIEW OF SEARCH ENGINE QUALITY

Existing literature on search engines advertising market is insufficient in explaining the impact of SEO because search engine quality has been exclusively associated with user satisfaction, without regard to the impact of SEO. Traditionally, quality has been defined in many disciplines. According to Garvin[9], quality is a multifaceted concept defined from several competing perspectives each communicating in its own terminology and favoring its own framework. The transcendent approach posits that quality is universally recognizable, but defies a precise definition. The product-based approach, on the contrary, views quality as a precise and measurable variable. Widely embraced by economists, this view introduces horizontal and vertical dimensions of quality and associates quality with cost. Apart from the product-based approach, the user-based approach is highly subjective, perception-based. The practical appeal of this approach, however, is seriously hampered by the problem of how product characteristics are aggregated and distinguished in a meaningful way to reflect quality. From an opposite perspective to the user-based approach, the manufacturing-based approach defines quality on the supply side which involves the engineering and manufacturing practices. In this approach, conformance to requirement is the most standard definition of quality. The sub-domains of operation management research such as quality assurance primarily adopt this approach. The value-based approach incorporates both demand and supply into the framework and defines quality as "performance at an acceptable cost or conformance at an acceptable cost". For lack of well-defined limit, its application is difficult in practice.

So far, studies in search engine market as represented by literature reviewed above have uniformly adopted the user-based approach

to model quality and its impacts. This dimension of quality is a major domain expertise of search engines in satisfying the information need of searchers, reflecting “crawling and indexing algorithms, the database index, and search and retrieval algorithms”[3]. In this study, we call this quality dimension *algorithm effectiveness*, an aggregated, reduced form a user-based quality. The higher the algorithm effectiveness, the more likely searchers will be satisfied. With zero access price, the demand in the search market increases with algorithm effectiveness[23].

In the online advertising market, Internet search engines attempts to make its page ranking unbiased with regard to its own relevancy standard[5]. Meanwhile, SEO firms generate “noises” that bias page ranking. Therefore, search engines differ in their ability to exclude these “noises”; the greater this ability, the less likely SEO firms are able to improve page ranking of search results of the search engine. we call this ability *algorithm robustness*. This dimension of quality a distinct attribute of Internet search engines. On one hand, an engine with the highest algorithm robustness is not vulnerable to SEO; SEO firms have no chance to improve the page ranking of any website. On the other hand, when the algorithm robustness is low, SEO firms have a chance of improving page ranking and cannibalize the search engine’s advertising revenue, to the extent that the algorithm robustness allows.

Algorithm robustness is a component of Internet search engine quality from the manufacture’s perspective in Garvin’s[9] framework. Higher algorithm robustness implies a greater conformance to the search engine’s ranking specification and lower vulnerability to “noises”, while lower algorithm robustness renders the information “manufactured” less predictable, more contaminated by “noises”. In the case of Internet search engines, the effects of algorithm robustness are twofold. First, since users’ satisfaction with search results is stochastic, so is the SEO practice in improving page ranking. No SEO firm knows the ranking algorithm of the search engine, and therefore, SEO practice only improves the chance of ranking improvement, rather than guarantees top ranking. Given an advertiser and advertising requirement, algorithm robustness denotes the effectiveness of SEO with the search engine.

Second, because SEO requires constant learning and adjustment of inter- and intra-website structures on the side of SEO firms, the efforts of optimization are expected to increase as algorithm robustness rises. It follows that algorithm robustness of a search engine alters the marginal cost of SEO with the search engine in question. As the algorithm robustness of a search engine rises, it becomes more expensive to optimize results in it. In other words, algorithm robustness is a “counter-quality” for SEO firms. As will be shown in the analytical model, algorithm robustness impacts a search engine’s advertising revenue in a competitive market. Overall, the higher the algorithm robustness, the less likely and more expensive is SEO. Figure 1 illustrates the search engine positions according to these two quality dimensions. In the figure, search engine 2 is more likely to satisfy searchers than search engine 1, which, however, is less vulnerable to SEO in comparison.

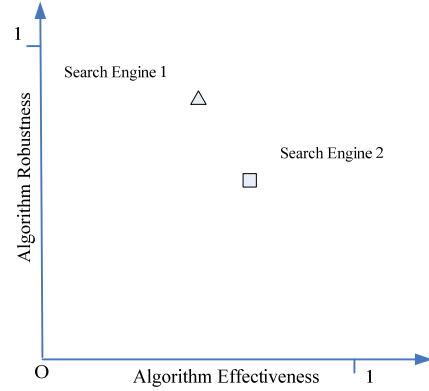


Figure 1: Algorithm Effectiveness and Algorithm Robustness

Internet search engine quality, therefore, is comprised of two dimensions: algorithm effectiveness and algorithm robustness. To a certain extent, while algorithm robustness can be improved over time primary through learning on the search engine’s side, algorithm effectiveness is highly sensitive to initial investment and less flexible. For example, *Google’s* core ranking algorithm is based on *PageRank*[1; 4], a patent of information retrieval filed when *Google* was founded. In this study, it is assumed that these two components are completely independent. It should be noted that algorithm robustness is affect by both search engines and SEO firms, because both parties engage in constant learning and improvement. Lower algorithm robustness, therefore, reflects a lack of learning and investment on the search engine’s side, or a lack of accumulated experience on the SEO firms’ side, or both.

This definition of search engine quality expands Telang *et al.*’s[23] seminal work on Internet search engine market. By doing so, this work differs from their treatment of per-click advertising revenue, which is assumed constant. According to Hoffman and Novak[12] and Dewan *et al.*[6], the advertising revenue of an advertising website is based on the number of hits the website receives. A search engine, however, differs from an advertising website in that it is not the number of hits on organic results but the number of hits on sponsored advertisements that directly impacts the revenue of the search engine. In addition, CPC also differs across keywords. In order to relate search engine quality with the advertising market, the analytical model in chapter 4 begins by deriving equilibrium price of paid advertisement as a function of search engine quality.

4. AN ANALYTICAL MODEL OF SEARCH ENGINE MARKET

4.1 Model Setting

Table 1 summarizes most notations used in this study.

Table 1: Summary of Notations

q	Algorithm effectiveness, a user-based quality
h	Algorithm robustness, a manufacturing-based

v	An advertiser's willingness-to-pay for a referred customer
$D(.)$	The demand function for searcher
s	The average ratio of searchers who click sponsored links
r	The impression-click conversion rate within a search result page
$g(.)$	Market clearing price of the per-click fee the search engine charges. It is a function of v and other factors related to keyword auction.
f	Per-period fee SEO firms charge
u	Net payoff for an advertiser
C	Cost function of the search engine

Suppose there are one search engine and many SEO firms in the market in a period of time, in which they both offer advertising in search market. Let q ($0 < q < 1$) denote the algorithm effectiveness and h ($0 < h < 1$) the algorithm robustness of the search engine. Besides the search engine and SEO firms, the market also contains advertisers and searchers. Advertisers are formally defined as merchants paying advertising fee to either the search engine or SEO firms for referral.

It is assumed that advertisers are heterogeneous only in their valuation of online advertisement. This difference can be attributed to conversion rate, industry as identified by advertising keywords, and other idiosyncratic, advertiser-specific factors. Given a keyword, advertisers differ in their conversion rate, or the ratio of the number of sales to the number of distinct advertising clicks, and their valuation of the keyword, both of which can be reduced into willingness-to-pay for online advertisement. Across industry, one would normally expect that a referred customer from "attorney service" or "Caribbean travel" to worth a lot more than that of "mp3", which, nevertheless, might be more popular.

In this study, we use the random variable v to denote the advertiser type in terms of its willingness-to-pay for a referred customer from the search engine. For simplicity, v is assumed to be uniformly distributed over the interval $[0, V]$ and used to denote a particular advertiser's willing-to-pay in a certain industry. Compared to classical definition of type in economics, the advertiser type in this study is broader, incorporating within industry and across-industry willingness-to-pay.

The next variable, D , denotes the demand for the search engine in terms of the total number of searchers in a period of time. These searchers are potential referred customer from the search engine through either organic results or sponsored results. Since searchers do not pay a fee to the search engine, searcher demand

	quality
--	---------

is a function of q [23]. As defined, q has a direct effect on the probability that a searcher will be satisfied with the search experience. We assume that D is a linear function of q , with $D(q) = a + bq$, where $a < 0$ and $b > 0$. Here we assume $a < 0$ because q needs to reach a certain threshold ($-a/b$) to drive demand. That is, when $D = 0$, there exist q such that $q > 0$.

Consistent with search engine users' preference of organic results over sponsored results, let s ($0 < s < 1$) denote the proportion of searchers who click sponsored results. Since there are multiple results in a search result page, not all impressions are translated into actual clicks. Let r ($0 < r < 1$) denote this rate of conversion within the search engine. Further, for simplicity, ranking *within* top ranked results is assumed not affect the probability of being clicked by searchers. This is reasonable because it has found that ranking within top-ranked results is not a significant factor of overall user satisfaction[19].

Let g be per-click fee the search engine charges for each click, and f be the per-period SEO fee charged by a SEO firm. Consistent with the keyword auction setting, the search engine does not set g .

Instead, it is determined by keyword auction. Further, based on the definition of h , in average, the probability that an optimized link will appear in organic results is $(1 - h)$. As previously assumed, SEO does not directly affect q because the impact of an optimized link on q may be positive, or negative, depending on the relevance of the link.

4.2 Advertisers' Problem

In order to analyze advertisers' problem, we make a classical assumption that advertisers have perfect knowledge of the payoff of either type of advertisement through learning and past experience. This is reasonable because in reality, advance web technologies allow advertisers to track and count link referrals periodically and calculate profit per referred customer. Let u denote the payoff for advertisers. It increases as the search engine attracts more searchers. This is in line with the widely accepted theory of indirect network externality [2; 11; 16] which determines the value of a network. In this case, each search engines is a network of searchers. The payoff that an advertiser gets from sponsored links per period depends on the size of the searcher pool of the engine. In specific, it equals the total advertising value from sponsored links net of the total advertising cost. Advertisers' problem is to choose the advertising vehicle that maximizes its net payoff. The net payoff from sponsored links, u_1 , is

$$u_1 = sr(a + bq)(v - g) \quad (1)$$

Denote the net payoff from SEO by u_2 . It equals the payoff that the advertiser gets from organic results per period net of the SEO fee. In equation:

$$u_2 = (1 - s)r(a + bq)v(1 - h) - f \quad (2)$$

Since there are many SEO firms, it is assumed that the SEO market is perfectly competitive. In addition, optimization of

search results, the product of SEO firms, is homogeneous regardless of the specific SEO firm, the industry, or the keyword chosen. Therefore, the price of SEO firms equals their marginal cost, which is assumed to be strictly increasing and strictly concave with respect to h . This is reasonable because SEO is a complicated and non-linear process resembling the cost of quality. With zero fixed cost, the equilibrium price of the SEO firms is $f = dh^2$ ($d > 0$). Plug into (2),

$$u_2 = (1-s)r(a+bq)v(1-h) - dh^2 \quad (2')$$

where dh^2 is the marginal cost of SEO firms. According to net payoff, the advertiser now chooses one of the three alternatives: no advertisement, paid placement, and SEO. In equation, the problem is

$$\max \begin{cases} \text{No Advertisement} \\ \text{Paid Placement} \\ \text{SEO} \end{cases} = \max \begin{cases} 0 \\ sr(a+bq)(v-g) \\ rv(1-s)(a+bq)(1-h) - dh^2 \end{cases}$$

4.3 Equilibrium Price of Paid Advertisement

The objective of this section is to partition online advertising based on v , the advertiser type. First, let $g_1(v)$ be the market clearing price for the indifferent advertiser of type between no advertisement and paid placement. With $u_1 = 0$, it can be solved as:

$$g_1(v) = v \quad (3)$$

Since prices of paid advertisement are determined by auction, $g_2(v)$, the market clearing price for the indifferent advertiser of type v between paid place and SEO, can be solved with $u_1 = u_2$:

$$g_2(v) = v - \frac{r(1-s)(a+bq)(1-h)v - dh^2}{rs(a+bq)} \quad (4)$$

v has two critical values: v_0 and v_1 . v_0 is the solution of $g_1(v) = g_2(v)$:

$$v_0 = \frac{dh^2}{r(a+bq)(1-s)(1-h)}$$

and v_0 is the solution of $g_2(v) = 0$:

$$v_1 = \frac{dh^2}{r(a+bq)(1-2s+sh-h)}$$

Notice $g_1'(v) > g_2'(v)$. However, the sign of $g_2'(v)$ is indeterminate. In general, three scenarios may arise:

4.3.1 Scenario 1

This scenario is characterized by $g_1'(v) > g_2'(v) \geq 0$. Figure 2 shows the market clearing price according to advertisers' choice.

Indifferent advertisers, who will, by assumption, choose search engine's paid advertisement, reside on the bold line, which is the equilibrium price of paid advertisement. In area A, the search engine dominates with SEO firms exiting the market. In area B, the auction price adjusts to SEO firms' marginal cost and SEO firms will also exit the market. Area C is the surplus gain for higher type advertisers ($v > v_0$) due to the existence of SEO firms. As advertiser type increases, this surplus grows as well. Overall, SEO firms do not survive in this scenario.

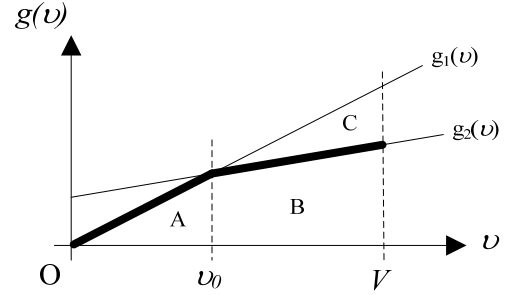


Figure 2

4.3.2 Scenario 2

This scenario is characterized by $g_2'(v) < 0$ and $v_1 \geq V$. Figure 3 shows the market partition where indifferent advertisers reside on the bold line, the equilibrium price of paid advertisement. Likewise, both area A and area B are dominated by the search engine with SEO firms exiting the market. Area C is the surplus gain for advertisers of types higher than v_0 . Compared to scenario 1, the growth of advertiser surplus with respect to v in this scenario is faster. No SEO firms survive in this scenario, either.

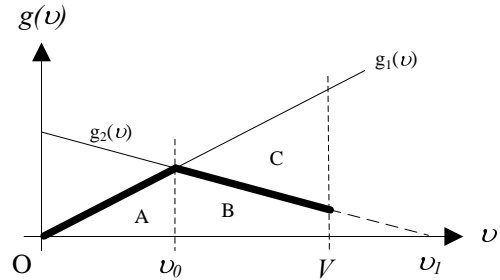


Figure 3

4.3.3 Scenario 3

This scenario is characterized by $g_2'(v) < 0$ and $v_1 < V$. Similar to scenario 2, in Figure 4 both area A and area B are dominated by the search engine. The bold line represents the equilibrium price of paid advertisement. However, the clearing price curve $g_2(v)$ stops at v_1 , because the auction price can not be negative. Like scenario 2, area C illustrates that the growth of advertiser surplus in this scenario is faster than scenario 1. More

importantly, SEO firms survive in area D, where higher type advertisers ($v_l < v < V$) reside.

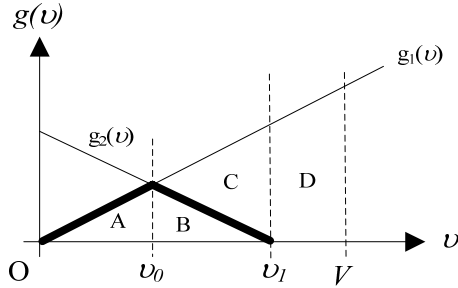


Figure 4

4.4 Sustainability conditions of SEO firms

In order to study the sustainability of SEO firms, the model has to allow SEO firms to exist in the boundary condition as they do in reality. Therefore, it is assumed that $v_0 < V$ holds for all scenarios. It can be observed that in all scenarios, advertisers gain surplus at the expense of the search engine's loss, due to the existence of competitive SEO firms. Therefore SEO firms are beneficial for higher type advertisers because of the additional surplus advertisers gain. SEO firms, in return, exist in market for higher type advertisers ($v_l < v < V$). In proposition:

Proposition 1: *There exists a threshold of advertiser type (v_l) above which SEO firms exist.*

This finding is consistent with trade experts' comments on SEO firms, that SEO is more expensive than sponsored results for many advertisers[13]. The model formally shows why SEO only appeals to higher type advertisers.

The above analysis depicts three scenarios of advertising market where SEO firms survive only in one scenario. From scenario 3, the condition under which SEO firms exist is:

$$0 < v_l < V$$

Rearrange,

$$0 < \frac{dh^2}{1-2s+sh-h} < Vr(a+bq)$$

Notice the sign of the denominator depends on s and h . A necessary condition for the sustainability of SEO firms in the market, therefore, is $1-2s+sh-h > 0$. Figure 5 demonstrates the "acceptance area" in which SEO firms may sustain in the market. On the other hand, in the rejection area, SEO firms will be driven out of the market.

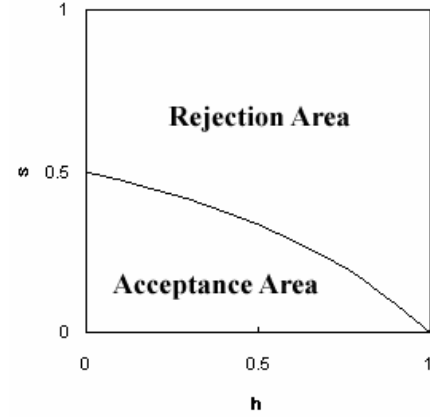


Figure 5

Given $0 < s < 0.5$, Figure 6 indicates that the search engine could drive the SEO firms out of the market if its algorithm effectiveness is sufficiently high. In proposition:

Proposition 2: *The sustainability of SEO firms is negatively associated with the algorithm robustness of the search engine.*

It should be realized that even in the acceptance area, SEO firms may not sustain. A sufficient condition requires, in addition,

$$\frac{dh^2}{1-2s+sh-h} < Vr(a+bq)$$

Therefore, it can be concluded that in the shard area, SEO firms tend to sustain when $V\theta(a+bq)$ is large. In propositions:

Proposition 3: *If the search engine operates in the "acceptance area", the sustainability of SEO firms is negatively associated with the marginal cost of SEO firms.*

Proposition 4: *If the search engine operates in the "acceptance area", the sustainability of SEO firms is positively associated with advertisers' maximum willingness-to-pay for a referred customer.*

This proposition should be interpreted comprehensively, because variance in v may come from difference sources. One reasonable interpretation is that SEO firms may sustain in some industries in which advertisers' willing to pay is sufficiently high, but not in others. An equally valid interpretation is that, within the same industry, SEO firms may sustain with some advertisers but not others. This comprehensive interpretation incorporates both within-industry and across-industry differences as factors of the sustainability of SEO firms. More importantly, in reality, scenarios 1 through 3 may co-exist in the market, where SEO firms sustain wherever v is sufficiently high.

Proposition 5: *If the search engine operates in the "acceptance area", the sustainability of SEO firms is positively associated with the algorithm effectiveness of the search engine.*

The sustainability of SEO firms is also subject to the market concentration of the search engine market. Now consider two search engines, SE_1 and SE_2 in the market with all other settings unchanged. Consistent with the scope of this study, we assume the two engines are identical except in their algorithm

effectiveness and algorithm robustness. That is, $SE_1 = (q_1, h_1)$, $SE_2 = (q_2, h_2)$.

With two search engines, market demand on searchers' side and that advertisers' side have to be reconsidered. On the advertisers' side, the demand for advertisement no longer depends on keyword auction alone, because both search engines use auction. Since Internet search is stochastic and no single engine can satisfy all users for all searches, we assume advertisers consider *both* search engines *separately* in order to achieve fuller coverage. To this end, the structure of the decision rule used in 4.2 still applies to either search engine. On the searchers' side, we adopt the concept of "residual demand" proposed by Telang *et al.*[23] to model the market demand for searches. Different from their study which features a two-stage, sequential-move, game theoretical model, we focus on a static, one-stage setting and analyze the impact on the sustainability of SEO firms.

Now suppose $q_2 > q_1$. The searcher demand of the SE_2 is $D_2(q_2) = a + bq_2$, while the searcher demand of the SE_1 is $D_1(q_1) = (1 - q_2)(a + bq_1)$. The opposite case ($q_1 > q_2$) can be analyzed symmetrically. Advertisers adopt separate decision rules for the two engines. For SE_1 :

$$\max \left\{ \begin{array}{c} 0 \\ sr(1 - q_2)(a + bq_1)(v - g_1) \\ rv(1 - s)(1 - q_2)(a + bq_1)(1 - h_1) - dh_1^2 \end{array} \right\},$$

For SE_2 :

$$\max \left\{ \begin{array}{c} 0 \\ sr(a + bq_2)(v - g_1) \\ rv(1 - s)(a + bq_2)(1 - h_1) - dh_1^2 \end{array} \right\},$$

Notice the decision rule for SE_2 is the same as in the case of one search engine. In other words, SE_1 would have absolutely no impact on the searcher demand of SE_2 in this model, because SE_2 has generated new demand not previously available to SE_1 . Hence, the sustainability condition for SEO firms with SE_2 stays the same:

$$0 < \frac{dh_2^2}{1 - 2s + sh_2 - h_2} < Vr(a + bq_2)$$

However, the reverse is not true: the existence of SE_2 alters the searcher demand and thus the sustainability condition for SEO firms with SE_1 . In the same manner, this condition can be derived:

$$0 < \frac{dh_1^2}{1 - 2s + sh_1 - h_1} < Vr(a + bq_1)(1 - q_2)$$

Now SEO firms could potentially sustain with two search engines. These two conditions are otherwise independent, except for q_2 . In particular, compared to the condition in the one search engine case, SEO firms are less likely to sustain with the SE_1 , because the entrance of SE_2 decreases the searcher demand of SE_1 and makes the sustainability condition with SE_1 more stringent. Formally,

Proposition 6: *When there are two search engines of different algorithm effectiveness in the market, the sustainability of SEO firms with the lower algorithm effectiveness engine is negatively influenced by the algorithm effectiveness of the higher algorithm effectiveness engine.*

The situation of *Yahoo* after *Google* entered the market closely matches this proposition. After *Google* became a clear leader in algorithm effectiveness, it has been a major target of SEO firms. *Google Dance Syndromes*[20; 21], or events that *Google* drastically revises its ranking algorithm and updates its index, are explicit attempts at counteracting SEO practice. In contrast, such events have seldom occurred in *Yahoo* since *Google* became the leader. According to the model, this is because *Google* has decreased the searcher demand of *Yahoo* and made SEO with *Yahoo* less sustainable.

4.5 Profit of the search engine

Next, we analyze search engine's profit. First, let $C(q)$ be the cost function of the search engine. This cost is a quadratic function of q , denoted by $C(q) = kq^2$. It is also assumed that h , the algorithm robustness of the search engine is a long-term investment decision and thus does not affect $C(q)$. By definition, the revenue of the search engine from a given keyword is the equilibrium price of paid advertisement multiplied by the number of clicks received by the sponsored results of the keyword. Suppose no SEO firms exist, the search engine earns the monopolistic profit:

$$\pi_M(q, h) = \frac{1}{V} \int_{0 < v < V} sr(a + bq_M)g_1(v)dv - kq_M^2$$

by setting its algorithm effectiveness, $q = q_M$, that satisfies

$$\text{F.O.C.: } \frac{\partial \pi(q, h)}{\partial q} = 0$$

and

$$\text{S.O.C.: } \frac{\partial^2 \pi(q, h)}{\partial q^2} < 0.$$

It can be derived that:

$$q_M = \frac{srbV}{4k}$$

$$\pi_M = sraV + \frac{s^2r^2b^2V^2}{8k}$$

In the presence of SEO firms, the search engine's profit in scenarios 1 and 2 can be expressed as:

$$\begin{aligned} \pi_1(q, h) &= \frac{1}{V} \int_{\psi} sr(a + bq)g(v)dv - kq^2 \\ &= \frac{sr(a + bq)}{V} \int_{0 < v < v_0} g_1(v)dv + \frac{sr(a + bq)}{V} \int_{v_0 < v < V} g_2(v)dv - kq^2 \end{aligned}$$

and the profit of the search engine in scenarios 3 is:

$$\begin{aligned} \pi_2(q, \rho) &= \frac{1}{V} \int_{\Psi} sr(a+bq)g(v)dv - kq^2 \\ &= \frac{sr(a+bq)}{V} \int_{0 < v < t_0} g_1(v)dv + \frac{sr(a+bq)}{V} \int_{t_0 < v < t_1} g_2(v)dv - kq^2 \end{aligned}$$

where Ψ is the area of the search engine's surplus in Figure 3 through Figure 5. It can be observed from these figures that, compared to the monopolistic profit, the search engine's maximum profit is less in Scenario 1 through 3, due to its loss of areas C and D. Derivation of the search engine's profit is included in Appendix.

To study the impact on the search engine's profit in scenario 1 and 2, we use numerical analysis and initially choose the parameters listed in Table 2. Parameters of interest to this study are h and V .

Table 2: Initial values of parameters

Parameter	Value	Source
s	0.3	According to searchers' preference
h^*	0.8	Algorithm robustness of the search engine
a	-100	The maximum demand in terms of number of searchers per period is $(a+b)$
b	1000	
d	10	dh^2 is the marginal cost of SEO firms per period. In order for SEO firms to cover certain portion of the market, $dh^2 < V$ should hold.
r	0.1	In <i>Google</i> , there are ten organic results and about eight paid results per page.
k	50	The search engine's marginal cost in terms of quality is kq^2 per period.
V^*	20	Maximum keyword bidding price for the given keyword, such as "real estate". Actual prices follow $U(\theta, V)$.

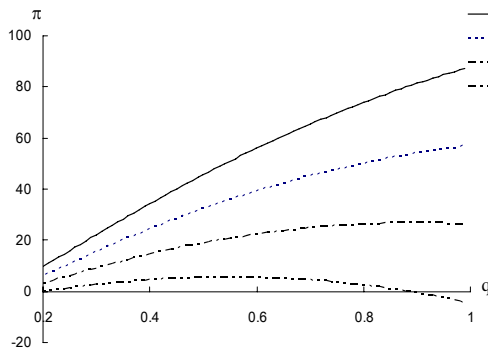


Figure 6: Effect of h

From Figure 6, it can be observed that as h increases, the maximum quality also increases. When h is above 0.7, quality improvement has a monotonic effect on profit. When h is below 0.7, an optimal quality emerges where the search engine achieves maximum, albeit lower profit. It can be concluded that h has an effect of protecting the search engine's investment in q . When h is low, investment in algorithm effectiveness will not be as effective, because of the "free-riding" effect: a high algorithm effectiveness search engine attracts both SEO firms and advertisers. The overall effect of SEO firms, therefore, is prohibiting the search engine from setting monopoly algorithm effectiveness and gain monopoly profit. In propositions:

Proposition 7: Overall, algorithm robustness positively influences the search engine's profit.

Proposition 8: Algorithm robustness positively influences the search engine's optimal algorithm effectiveness and maximum profit.

To a certain extent, the market scenario before *Google* was launched in 1998 can be explained with these findings. Before 1998, *Yahoo* was a barely profitable search engine in the online advertising market. Meanwhile, SEO was a prevalent practice available to advertisers. Algorithm robustness of *Yahoo* then was low. This explains why the quality of *Yahoo* had not been improving dramatically, until recently.

Another effect on the profit of the search engine is the willingness-to-pay for online advertisement at large. This effect is interesting because over time, advertisers begin to realize the value of search engine advertising and therefore bear higher valuation on online advertisement. In reality, the increase in CPC in recent years reflects this trend. In the model, V captures this effect, which is shown in Figure 7.

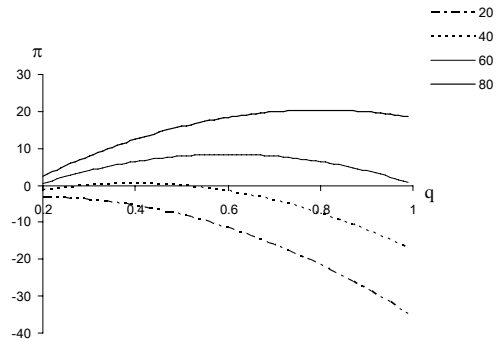


Figure 7: Effect of V

Consistent with intuition, from the graph, it can be observed that V has a positive effect on the overall profit of the search engine. As V increases, the online advertising market expands and the search engine is able to reap more profit. Additionally, growth in V also increases the search engine's optimal algorithm effectiveness. Formally:

Proposition 8: Overall, advertisers' willingness-to-pay for online advertisement positively influences the search engine's profit.

Proposition 9: Advertisers' willingness-to-pay for online advertisement positively influences the search engine's optimal algorithm effectiveness.

These findings further explain what facilitates algorithm effectiveness improvement of a search engine when its algorithm robustness is low ($h = 0.6$ in this case). As online advertisement has been valued higher, more advertising dollars flow into the market. Again, because of the "free-riding" effect, if the search engine is not able to improve its algorithm robustness, it is "forced" to improve its algorithm effectiveness at the expense of giving up more advertising revenue, albeit less in proportion, to SEO firms.

Next, by observing the profit function in scenario 3, it can be concluded that, given a positive market demand, the effect of quality on profit is either positive or negative, depending on the parameter configuration of s and h . In particular:

$$\text{sign}\left[\frac{\partial \pi_2(q, h)}{\partial q}\right] = \text{sign}\left[\frac{1}{(1-s)(1-h)} - \frac{1}{(1-s)(1-h)-s}\right]$$

Combining areas separated by signs in this equation with Figure 6, this condition supplements the operation areas of the search engine with the impact of SEO firms on the search engine's decision on quality investment, as shown in Figure 8.

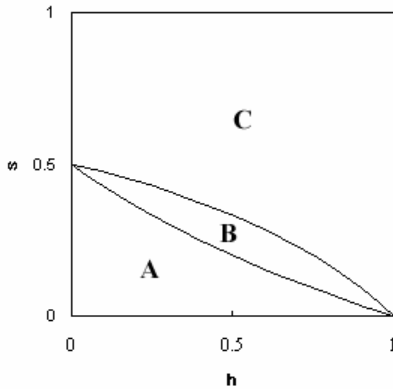


Figure 8: Search Engine Operation Areas

The acceptance area in Figure 5 is now sub-divided into Area A and Area B, leaving Area C, the rejection area unchanged. Area A and Area B have different implications to the search engine's choice of optimal quality. In A, the search engine's improvement in algorithm effectiveness works against its profit because of the free-riding effect of SEO firms. On the contrary, in B, the improvement in algorithm effectiveness has a positive return. Since Figure 9 is based on scenario 3, whereas the actual profit of the search engine is composed of profits from advertisers of different willingness-to-pay, the distinction of Area A and Area B can not be used to draw an unambiguous conclusion. However, it is reasonable to conclude that the search engine's incentive to invest in quality depends on the specific configuration of s and h .

5. Discussion

Several interesting insights emerge from the result of the analysis. First, the sustainability of SEO firms depends, in the first place, on the advertisers' willingness-to-pay for online advertising. As this valuation rises over time, SEO firms offer an advantage over paid placement. This result is primarily due to the different pricing policies adopted by the search engine and SEO firms. The net payoff for higher type advertisers using paid placement decreases because the marginal cost from CPC does not keep up with the marginal benefit from advertising. On the contrary, in the case of SEO, the marginal benefit increases due to the constant SEO fee. The practical implication is that search engines could increase its profit by adopting period-based pricing policy, rather than CPC, for higher-type advertisers.

The sustainability of SEO firms also depends on s , the proportion of sponsored results returned, and h , the algorithm robustness. Intuitively, decreasing the proportion of organic results could pose threat to SEO firms. However, there is a limit with regard to s , because the majority of searchers prefer organic results over sponsored results. Algorithm robustness, on the other hand, has a monotonic negative effect on the sustainability of SEO because it directly confines the practice of SEO. The practical implication, therefore, is that search engines could improve its profit through constant learning and "outsmart" SEO firms, so that its results are less vulnerable to SEO practice in general.

More importantly, a search engine is potentially subject to "free-riding" effect from SEO firms, because of the parasitic nature of these firms. As the search engine invest in algorithm effectiveness improvement, SEO firms may also benefit from this investment. In order to reap a fuller benefit from investment, the search engine has the incentive to improve its algorithm robustness at the same time. This phenomenon has been frequently observed in *Google Dance Syndrome*[20; 21], a deliberate attempt at improving its algorithm robustness.

Interestingly, in the case where two search engines exist, namely a leader with higher algorithm effectiveness and a follower with lower algorithm effectiveness, the pressure from SEO firms is lower for the follower. This is so because the follower only covers the "residual demand" left over from the leader. In particular, this pressure is negatively influenced by the algorithm effectiveness of the leader. This finding closely resembles the relationship between *Google* and *Yahoo*: massive updates of ranking algorithm have seldom been observed with *Yahoo*. To certain degree, this follower position offers an additional benefit to the follower in that SEO firms are less likely to sustain. This benefits partly justifies why *Yahoo*[24] recently announced that it gives up search dominance strategy in the presence of *Google*.

6. Conclusion

Search Engine Optimization is an interesting but less understood area in online advertising industry. In this study, we attempt to analyze the sustainability of SEO firms and the impact of SEO and other factors on search engine profit. Several interesting findings have been derived that concur with other SEO research and explain phenomena in the online advertising industry. First, a search engine could optimize its pricing policies for higher-type advertisers to reap higher profit. Second, investment in algorithm robustness has the effect of protecting the investment in algorithm

effectiveness. Third, the second market position endows the follower additional benefits due to low sustainability of SEO firms. In summary, the contribution of the study is threefold. First, this is the first study that decomposes search engine quality into two components. This distinction allows search engines to make informed decision in quality investment. Second, the study builds an analytical model that reveals the impact of SEO and helps understand its future trend. Third, the results generated offer practical insights into search engine's profitability.

Due to the scope constraint and the type of the study, there are several limitations to this study. First, the uniform distribution assumption of advertisers' willingness-to-pay is simplistic in reality. One possible extension is to model industry difference and advertiser difference with a hierarchical distribution and thereby, separate the two effects. The result could yield managerial insights in terms of market segmentation. Alternatively, horizontal differentiation model could be used to address the advertiser heterogeneity in keyword preferences. This differentiation echoes the product differentiation of the online marketplace in reality because keywords, like consumer products, could also be differentiated. This latter alternative shifts the focus of the study to the advertisers' profit and strategies as Sen's[18] work. Second, the profit analysis of two search engine could further be expanded into a dynamic, game-theoretical model as in Telang *et al.*'s[23] study. Future inquiries will benefit from these limitations and suggestions and explore the area of SEO and online advertising with more comprehensive treatments.

7. REFERENCES

- [1] A. Acharya, M. Cutts, J. Dean, P. Haahr, M. Henzinger, U. Hoelzle, S. Lawrence, K. Pflieger, O. Sercinoglu and S. Tong "Information retrieval based on historical data," US, 2005,
- [2] A. Basu, T. Mazumdar and S.P. Raj "Indirect network externality effects on product attributes," *Marketing Science* (22:2), 2003, pp. p. 209.
- [3] H.K. Bhargava and J. Feng "Recommendation Bias in Information Gatekeepers," 2005,
- [4] S. Brin and L. Page "The Anatomy of a Large-Scale Hypertextual Web Search Engine," *Proceedings of the The seventh international conference on World Wide Web 7*, Brisbane, Australia, 1998.
- [5] J. Caufield "Where Did Google Get Its Value?" *Libraries and the Academy* (5:4), 2005, pp. 555-572.
- [6] R. Dewan, M. Freimer and J. Zhang "Managing web sites for profitability: Balancing content and advertising," *Proceedings of the The 35th Hawaii International Conference on System Sciences*, Hawaii, 2002.
- [7] DogPile.com "Difference Engines, Different Results," 2005.
- [8] FathomOnline.com "Search Keyword Prices Rose 19 Percent In One Year," October 5, 2005.
- [9] D.A. Garvin "What Does "Product Quality" Really Mean?" *Sloan Management Review* (26:1), 1984, pp. 25-43.
- [10] Georgia_Tech "How Users Find out About WWW Pages," 1998, last accessed: at: http://www.cc.gatech.edu/gvu/user_surveys/survey-1998-10/graphs/use/q52.htm.
- [11] S. Gupta, D.C. Jain and M.S. Sawhney "Modeling the Evolution of Markets with Indirect Network Externalities: An Application to Digital Television," *Marketing Science* (18:3), 1999, pp. p. 396.
- [12] D. Hoffman and T. Novak "A new paradigm for electronics commerce," *Information Society* (13:January-March), 1997, pp. 43-54.
- [13] G. Jarboe "Why does search engine marketing look like a penny-farthing bicycle?", 2005,
- [14] E.W. Johnson, W. Moe, P. Fader, S. Bellman and J. Lohse "On the Depth and Dynamics of Online Search Behavior," *Management Science* (50), 2004.
- [15] S.C. Lawrence and L. Giles "Searching the World Wide Web," *Science* (280:3), 1998, pp. pp. 98-100.
- [16] S.J. Liebowitz and S.E. Margolis "Network Externality: An Uncommon Tragedy," *Journal of Economic Perspectives* (8:2), 1994, pp. pp. 133-150.
- [17] SEMPO.org "The State of Search Engine Marketing 2004," 2004.
- [18] R. Sen "Optimal Search Engine Marketing Strategy," *International Journal of Electronic Commerce* (10:1), 2005, pp. 9-25.
- [19] L.T. Su "A Comprehensive and Systematic Model of User Evaluation of Web Search Engines: I. Theory and Background," *Journal of the American Society for Information Science and Technology* (54:13), 2003, pp. 1175-1192.
- [20] D. Sullivan "Coping With GDS, The Google Dance Syndrome," 2003,
- [21] D. Sullivan "Google Dance Syndrome Strikes Again," 2003,
- [22] R. Telang and T. Mukhopadhyay "The drivers of Web portal use," *Electronic Commerce Research and Application* (forthcoming), 2003.
- [23] R. Telang, U. Rajan and T. Mukhopadhyay "The Market Structure for Internet Search Engines," *Journal of Management Information Systems* (21:2), 2004, pp. pp. 137-160.
- [24] J. Thaw "Yahoo! gives up quest for search dominance," 2006,
- [25] I. Witten, A. Mofflt and T.C. Bell *Managing gigabytes: Compressing and indexing documents and images*, Van Nostrand Reinhold, New York, NY, 1994.
- [26] R. Zwick, A. Rapoport, A. Lo and A. Muthukrishnan "Consumer Sequential Search: Not Enough or Too Much?" *Marketing Science* (22), 2003, pp. pp. 503-519.

8. Appendix

8.1 Search engine's profit in scenario 1 and 2

$$\begin{aligned}
& \frac{sr(a+bq)}{V} \int_{0 < v < v_0} g_1(v) dv \\
&= \frac{sr(a+bq)}{V} \int_{0 < v < v_0} v dv \\
&= \frac{sr(a+bq)}{V} \left[\frac{v^2}{2} \right]_0^{v_0} \\
&= \frac{sr(a+bq)}{2V} \left(\frac{dh^2}{r(1-s)(a+bq)(1-h)} \right)^2 \\
&= \frac{sd^2h^4}{2Vr(1-s)^2(a+bq)(1-h)^2}
\end{aligned}$$

$$\begin{aligned}
& \frac{sr(a+bq)}{V} \int_{v_0 < v < v^*} g_2(v) dv \\
&= \frac{sr(a+bq)}{V} \int_{v_0 < v < v^*} \left(v - \frac{r(1-s)(a+bq)(1-h)v - dh^2}{sr(a+bq)} \right) dv \\
&= \frac{sr(a+bq)}{V} \int_{v_0 < v < v^*} \left[v - \frac{r(1-s)(a+bq)(1-h)v + dh^2}{sr(a+bq)} \right] dv \\
&= \frac{sr(a+bq)}{V} \int_{v_0 < v < v^*} \left[v \frac{2s-1+h-sh}{s} - \frac{dh^2}{sr(a+bq)} \right] dv \\
&= \frac{sr(a+bq)}{V} \left[v^2 \frac{2s-1+h-sh}{2s} - v \frac{dh^2}{sr(a+bq)} \right]_{v_0}^{v^*} \\
&= \frac{sr(a+bq)}{V} \left[v^2 \frac{2s-1+h-sh}{2s} - v \frac{dh^2}{sr(a+bq)} \right] - \frac{sr(a+bq)}{V} \left[v_0^2 \frac{2s-1+h-sh}{2s} - v_0 \frac{dh^2}{sr(a+bq)} \right] \\
&= \left[\frac{Vr(a+bq)}{2} (2\alpha - 1 + h - sh) - dh^2 \right] \\
&\quad - \frac{sr(a+bq)}{V} \left[\left(\frac{dh^2}{r(1-s)(a+bq)(1-h)} \right)^2 \frac{2s-1+h-sh}{2s} - \left(\frac{dh^2}{r(1-s)(a+bq)(1-h)} \right) \frac{dh^2}{sr(a+bq)} \right] \\
&= \frac{Vr(a+bq)(2s-1+h-sh)}{2} - dh^2 - \frac{d^2h^4(2s-1+h-sh)}{2Vr(1-s)^2(a+bq)(1-h)^2} + \frac{d^2h^4}{Vr(1-s)(a+bq)(1-h)}
\end{aligned}$$

$$\begin{aligned}
\pi_1(q, h) &= \frac{sd^2h^4}{2Vr(1-s)^2(a+bq)(1-h)^2} + \frac{Vr(a+bq)(2s-1+h-sh)}{2} - dh^2 \\
&\quad - \frac{d^2h^4(2s-1+h-sh)}{2Vr(1-s)^2(a+bq)(1-h)^2} + \frac{d^2h^4}{Vr(1-s)(a+bq)(1-h)} - kq^2 \\
&= \frac{d^2h^4}{2Vr(1-s)(a+bq)(1-h)} + \frac{Vr(a+bq)(2s-1+h-sh)}{2} - dh^2 \\
&\quad + \frac{d^2h^4}{Vr(1-s)(a+bq)(1-h)} - kq^2 \\
&= \frac{3d^2h^4}{2Vr(1-s)(a+bq)(1-h)} + \frac{Vr(a+bq)(2s-1+h-sh)}{2} - dh^2 - kq^2
\end{aligned}$$

8.2 Search engine's profit in scenario 3

$$\begin{aligned}
& \frac{sr(a+bq)}{V} \int_{v_0 < v < v_1} g_2(v) dv \\
&= \frac{sr(a+bq)}{V} \int_{v_0 < v < v_1} \left(v - \frac{r(1-s)(a+bq)(1-h)v - dh^2}{sr(a+bq)} \right) dv \\
&= \frac{sr(a+bq)}{V} \int_{v_0 < v < v_1} \left[v - \frac{r(1-s)(a+bq)(1-h)v + dh^2}{sr(a+bq)} \right] dv \\
&= \frac{sr(a+bq)}{V} \int_{v_0 < v < v_1} \left[v \frac{2s-1+h-sh}{s} - \frac{dh^2}{sr(a+bq)} \right] dv \\
&= \frac{sr(a+bq)}{V} \left[v^2 \frac{2s-1+h-sh}{2s} - v \frac{dh^2}{sr(a+bq)} \right]_{v_0}^{v_1} \\
&= \frac{sr(a+bq)}{V} \left[v_1^2 \frac{2s-1+h-sh}{2s} - v_1 \frac{dh^2}{sr(a+bq)} \right] - \frac{sr(a+bq)}{V} \left[v_0^2 \frac{2s-1+h-sh}{2s} - v_0 \frac{dh^2}{sr(a+bq)} \right] \\
&= \frac{sr(a+bq)}{V} \left[\left(\frac{dh^2}{r(a+bq)(1-2s+sh-h)} \right)^2 \frac{2s-1+h-sh}{2s} - \left(\frac{dh^2}{r(a+bq)(1-2s+sh-h)} \right) \frac{dh^2}{sr(a+bq)} \right] \\
&\quad - \frac{sr(a+bq)}{V} \left[\left(\frac{dh^2}{r(1-s)(a+bq)(1-h)} \right)^2 \frac{2s-1+h-sh}{2s} - \left(\frac{dh^2}{r(1-s)(a+bq)(1-h)} \right) \frac{dh^2}{sr(a+bq)} \right] \\
&= \frac{d^2h^4}{2Vr(a+bq)(1-2s+sh-h)} - \frac{d^2h^4}{Vr(a+bq)(1-2s+sh-h)} \\
&\quad - \frac{d^2h^4(2s-1+h-sh)}{2Vr(1-s)^2(a+bq)(1-h)^2} + \frac{d^2h^4}{Vr(1-s)(a+bq)(1-h)} \\
&= \frac{3d^2h^4}{2Vr(a+bq)(1-2s+sh-h)} - \frac{d^2h^4(2s-1+h-sh)}{2Vr(1-s)^2(a+bq)(1-h)^2} + \frac{d^2h^4}{Vr(1-s)(a+bq)(1-h)}
\end{aligned}$$

$$\begin{aligned}
\pi_2(q, h) &= \frac{sd^2h^4}{2Vr(1-s)^2(a+bq)(1-h)^2} - \frac{3d^2h^4}{2Vr(a+bq)(1-2s+sh-h)} \\
&\quad - \frac{d^2h^4(2s-1+h-sh)}{2Vr(1-s)^2(a+bq)(1-h)^2} + \frac{d^2h^4}{Vr(1-s)(a+bq)(1-h)} \\
&= \frac{d^2h^4}{2Vr(1-s)(a+bq)(1-h)} - \frac{3d^2h^4}{2Vr(a+bq)(1-2s+sh-h)} + \frac{d^2h^4}{Vr(1-s)(a+bq)(1-h)} \\
&= \frac{3d^2h^4}{2Vr(a+bq)(1-s)(1-h)} - \frac{3d^2h^4}{2Vr(a+bq)(1-2s+sh-h)} \\
&= \frac{3d^2h^4}{2Vr(a+bq)} \left[\frac{1}{(1-s)(1-h)} - \frac{1}{(1-s)(1-h)-s} \right]
\end{aligned}$$