Consideration Sets and Competitive Marketing*

Kfir Eliaz† and Ran Spiegler‡

First version: April 2006
This version: December 2007

Abstract

We study a market model in which competing firms use costly marketing devices to influence the set of alternatives that boundedly rational consumers perceive as relevant. Consumers in our model apply well-defined preferences to a "consideration set", which is a subset of the feasible set and subject to manipulation by firms. We examine the implications of this behavioral model on otherwise competitive markets. In our model, the market equilibrium outcome is not competitive, yet firms earn competitive payoffs because the strategic use of costly marketing devices wears off the collusive impact of consumers' bounded rationality. Equilibrium behavior satisfies an "effective marketing property": if a consumer considers a firm only because of a marketing device it employs, he ends up buying from that firm.

KEYWORDS: marketing, advertising, consideration sets, bounded rationality, attention, persuasion, reason-based choice, irrelevant alternatives, internet search, search engine optimization

1 Introduction

We present a model of competitive marketing based on the notion that consumers are boundedly rational and that marketing interferes with their decision process. The standard model of consumer behavior assumes that the consumer has a clear perception of what is desirable (captured by the preference relation) and a clear perception of what is feasible, taking into account informational constraints (captured by the choice set). Our model retains the assumption that consumers have stable preferences, while

---

*We thank Amitav Chakravarti, Eddie Dekel, Barton Lipman and Ariel Rubinstein for useful conversations. Spiegler acknowledges financial support from the ISF, Grant no. 0610114022.

†Department of Economics, Brown University, Providence, RI 02912. E-mail: Kfir_Eliaz@Brown.edu. URL: http://www.econ.brown.edu/fac/Kfir_Eliaz.

‡Department of Economics, University College London, Gower St., London WC1E 6BT, UK. E-mail: r.spiegler@ucl.ac.uk. URL: http://www.homepages.ucl.ac.uk/~uctprsp.

relaxing the assumption that they perfectly perceive what is feasible, thus allowing firms to manipulate that perception. Our objective is to examine the implication of this departure from the standard model on the nature of competition between firms, and on the role that marketing plays in such competitive environments.

The cornerstone of our model is the observation that in the modern marketplace, consumers face an overwhelmingly large variety of products and therefore often use screening criteria (deliberate as well as unconscious) in order to reduce the number of "relevant" alternatives. As a result, consumers apply their preferences not to the set of objectively feasible alternatives, but to a potentially smaller set which they construct at an earlier stage of the decision process. Borrowing a term from the marketing literature, we refer to this set as the "*consideration set*" (see Roberts and Lattin (1997)). For a rational consumer, there is no distinction between the consideration set and the feasible set. However, for a boundedly rational consumer, the consideration set may be a strict subset of the feasible set, either because he is unaware of certain products or because he is unconvinced that they are relevant to his decision problem.

For concreteness, imagine a store manager who wishes to get potential customers into the store. Once a customer is inside, he can survey the products on display and if he finds something better than his outside option, he will purchase it at the store. But how would the manager get the customer into the store in the first place? He may put on display a product that possesses features which, quite independently of their intrinsic consumption value, are good at attracting attention.[1] Alternatively, the manager may try to *persuade* the customer that there is a good reason for him to enter the store - specifically, that the products sold at the store are superior to the customer's outside option according to some criterion (terms of payment, extended warranties, an aspect of quality). In both cases, it may well be that after giving serious consideration to the products sold at the store, the customer will conclude that he prefers his outside option after all. However, in order to bring the customer to the point where he actually exercises his preferences, the store manager had to employ some "door opening" marketing technique.

We construct a simple model of consumer choice that incorporates the idea that an important preliminary phase in consumers' decision process is the formation of consideration sets, and that consequently, employing marketing techniques to influence the consideration set is a major aspect of firms' competitive behavior. In our model,

[1] A vivid example of this marketing technique involves a soda company that issues a "limited holiday edition" including absurd flavors such as Christmas ham or latke - see http://www.jonessoda.com/files/holiday_2007.php.

a choice problem is a pair (M, s), where M is a menu of products and s is a product that serves as a "reference point". The consumer's choice procedure is based on two primitives: a standard preference relation $\succsim$ and another binary relation R, called the "*consideration relation*". The relation yRx means that the consumer is willing to consider y at x. The consumers' choice procedure consists of two stages. In the first stage, he constructs a consideration set, which is $M \cup \{s\}$ if xRs for some $x \in M$, and $\{s\}$ otherwise. (Figuratively put, when the consumer's reference point is x, he agrees to enter a store whenever y belongs to the set of products it offers.) In the second stage, the consumer chooses one of the most preferred products in the consideration set (with a tie-breaking rule that favors the reference point s).

For most of the paper, we assume that R is complete and transitive. This restriction captures the above-mentioned idea that consumers demand a *reason* that would justify considering products in M as potential substitutes to s. A preference criterion is a natural example of such a reason. The criterion underlying the consideration relation need not coincide with actual preferences. The former reflects a superficial impression, while the latter is the outcome of careful deliberation or consumption experience. One may wonder why the consumer applies R as a screening criterion if it fails to coincide with his actual preferences. First, the screening criterion can be instinctive and "hard-wired" rather than a result of deliberation (think of a rule such as "bigger is better"). Second, the criterion may be a reasonable predictor of preferences across a large set of markets. Using the criterion as a screening device saves cognitive costs because it can be applied to many market situations. However, it may diverge from the consumer's preferences in any particular instance.

The heart of the paper is a pair of market applications, in which two identical firms compete over a homogeneous population of consumers who follow the consideration-sets procedure. These applications have the property that if consumers were rational, the equilibrium outcome would be manifestly competitive and no marketing device would ever be employed. We first analyze a model in which firms simultaneously choose menus of products and aim to maximize market share minus the fixed cost of adding products to the menu. For half the population of consumers, M coincides with firm 1's menu and s is the most preferred product on firm 2's menu, and vice versa for the other half. We assume that the consumers' most preferred product also has the highest menu cost, and that the cost of offering the grand set is lower than the value of a 50% market share. If consumers were rational, firms would offer the consumers' most preferred product as a singleton in Nash equilibrium.

In contrast, as long as the most preferred product is not a R-maximal product as well, the outcome of symmetric Nash equilibrium is non-competitive, in the sense that firms offer inferior products with positive probability. Moreover, firms necessarily offer non-singleton menus with positive probability in symmetric equilibrium. A non-singleton menu necessarily contains "irrelevant alternatives" - namely, products which consumers never choose. The function of these products is to attract consumers' serious attention to other, better products on the same menu. Thus, the costly use of irrelevant alternatives as a marketing device is a necessary counterpart of a non-competitive equilibrium outcome. The expected cost of irrelevant alternatives may be viewed as a "deadweight loss" resulting from consumer bounded rationality.

Our main result is that although the equilibrium outcome may be non-competitive, firms necessarily earn "competitive payoffs" in symmetric Nash equilibrium- that is, industry profits are as if consumers were rational. From this perspective, the function of irrelevant alternatives is to restore an aspect of market competitiveness which is initially eroded by consumer bounded rationality. The competitive-payoff result has a subtle corollary concerning the effectiveness of irrelevant alternatives as a marketing device. Recall that the function of irrelevant alternatives is to "get the customer into the store". It turns out that in symmetric equilibrium, whenever the consumer considers a firm because of an irrelevant alternative in its menu, he necessarily ends up buying from that firm (unless his outside option is the best possible product). We refer to this result as the "*effective marketing property*".

To further demonstrate the scope of our framework, we also study a model of competitive advertising. The consideration-sets procedure allows two products to be equivalent in terms of the preference relation yet distinct in terms of the consideration relation. One possible interpretation is that the two products are in fact the same product as far as the consumption experience is concerned, but they are *framed* differently and therefore differ in the set of products from which they attract serious attention.

In the competitive-advertising model, each firm sells a single product which is characterized both by its set of actual features P and by the subset of features $A \subseteq P$ which the firm chooses to advertise. We assume that in terms of preferences, consumers care only about the number of actual product features. The consideration relation R, in contrast, is defined over pairs (P, A). Different specifications of R capture different assumptions regarding the way advertising manipulates consumers' consideration sets. Our main result in this section is that whenever R is complete and transitive, firms earn competitive payoffs in symmetric Nash equilibrium. We also characterize symmetric equilibria under an example of such a consideration relation. The characterization consists of a description of the products that firms offer as well as the way they choose to advertise them.

Our paper contributes to a growing theoretical literature on market interaction between profit-maximizing firms and boundedly rational consumers. Piccione and Rubinstein (2003) study intertemporal pricing when consumers have diverse ability to perceive temporal patterns. Spiegler (2006a,b) analyzes markets in which profit-maximizing firms compete over consumers who use naive sampling to evaluate firms. Shapiro (2006) studies a model in which firms use advertising to manipulate the beliefs of consumers with bounded memory. DellaVigna and Malmendier (2004), Eliaz and Spiegler (2005,2006), and Gabaix and Laibson (2006) study interaction with consumers having limited ability to predict their future tastes. Mullainathan, Shleifer and Schwartzstein (2007) study the role of uninformative advertising when consumers apply "coarse reasoning". Kamenica (2006) examines different forms of information asymmetries between firms and consumers that may give rise to observed consumer behavior that violates independence of irrelevant alternatives. For a field experiment that quantifies the effects of various marketing devices in terms of their price-reduction equivalent, See Bertrand, Karlan, Mullainathan, Shafir and Zinman (2006).

The plan of the paper is as follows. We present the consideration-sets procedure in Section 2, where we also discuss some related choice-theoretic literature. We analyze the competition-in-menus model in Section 3. Section 4 is devoted to the competitive advertising model. We conclude in Section 5 with a discussion of variations and extensions. Proofs not given in the main text are relegated to the Appendix.

2 Consideration Sets

Let X be a finite set of *products*. A *menu* is a non-empty subset of X. A *choice problem* is a pair (M, s), where M is a menu and $s \in X$ is a "*reference point*", which may or may not belong to M. The interpretation we will favor in this paper is that M is a set of products offered by a new store the consumer is yet to be familiar with, while s is a product the consumer regularly buys or used to buy from another supplier. A *choice correspondence* assigns to each choice problem a non-empty subset of $M \cup \{s\}$.

The consumer chooses according to a procedure based on two primitives: a standard preference relation $\succsim$ over X and another binary relation R over X, which we call the "*consideration relation*". For any menu M, let $b(M)$ denote the set of $\succsim$-maximal products in M. The relation yRx means that the consumer is willing to consider y at x. As explained in the Introduction, in the sequel we will often assume that R is complete and transitive. The procedure consists of two stages. In the first stage, the consumer constructs a consideration set according to the following rule. If xRs for some $x \in M$, the consideration set coincides with the feasible set $M \cup \{s\}$; otherwise, the consideration set is $\{s\}$. In the second stage, the consumer chooses a $\succsim$-maximal product in the consideration set he constructed in the first stage. The reference point is chosen whenever it is a $\succsim$-maximal element in the consideration set. The procedure thus induces the following choice correspondence: $c(M,s) = b(M \cup \{s\})$ if xRs for some $x \in M$ and $b(M) \succ s$, and $c(M,s) = s$ otherwise.

The consideration relation R enriches our description of the consumer's psychology. In addition to his preferences, the consumer is characterized by his willingness to consider new products, and consequently by the lengths to which a marketer has to go in order to draw his serious attention to new products. Personality psychologists often regard "*openness to experience*" as one of the basic traits that define an individual's personality (see Goldberg (1993)). The consideration relation may be viewed as a representation of this trait: a more "open" personality corresponds to a consideration relation that induces a larger set $\{(x,y) \in X \times X \mid yRx\}$.

The choice correspondence induced by the consideration-sets procedure violates the axiom of Independence of Irrelevant Alternatives. For instance, let $X = \{x,y,z\}$, $x \succ y \succ z$, zRy, $x\not{R}y$. Then, $c(\{x,z\},y) = \{x\}$ whereas $c(\{x\},y) = y$. Thus, despite the fact that the primitives of the choice procedure are two preference relations, the induced choice behavior may violate rationality. The following example illustrates another non-standard effect of our model: adding s to the menu M may cause the consumer to switch from the reference point to another product in M. For instance, let $X = \{x,y\}$, $x \succ y$, yRy, $x\not{R}y$. Then, $c(\{x\},y) = \{y\}$ whereas $c(\{x,y\},y) = \{x\}$. Note that the consideration-sets procedure contains rational choice as a special case, when yRx for every two products x,y.[2]

Related choice-theoretic literature

Complete choice-theoretic analysis of the consideration-sets procedure is outside the scope of the present study, which is more concerned with the procedure's implications for market behavior. Nevertheless, it is insightful to draw some comparisons with other decision models in the choice-theoretic literature.

The *Rational Shortlist Method* due to Mariotti and Manzini (2007) is a procedure which, like our model, applies two binary relations sequentially. The first binary relation is used to shrink the menu into a "shortlist" (Mariotti and Manzini rely on the

[2]In light of the anomalous examples, the reader may wonder whether our model can accommodate any arbitrary choice behavior. This is not the case. In particular, expanding M can never cause the consumer to switch from a product $x \neq s$ to s.

standard definition of a choice problem as a menu), while the second binary relation selects a unique product from the shortlist. To see the difference between the two procedures, note that under the Rational Shortlist Method, if both binary relations are complete and transitive, the procedure is reduced to standard lexicographic preferences. In contrast, our model accommodates non-rational choice even when both binary relations are preference relations.

Our notion of a choice problem with a reference point follows a number of recent choice models, e.g., Rubinstein and Zhou (1999) and Masatlioglu and Ok (2005). The latter paper interprets the reference point as a status quo and provides an axiomatization of a decision rule based on an incomplete preference relation over X. The decision maker chooses an element $x \neq s$ from M if and only if there exists such an element which is superior to s according to the incomplete preference relation. The important behavioral difference between the Masatlioglu-Ok model and the consideration-sets procedure is that the former model satisfies Independence of Irrelevant Alternatives. More broadly, the two models can be viewed as capturing different types of a status quo bias. The Masatlioglu-Ok model addresses this bias as a feature of the decision maker's preferences, whereas in our model, the conservative bias in favor of the reference point occurs at an earlier stage of the decision process, in which the decision maker constructs the set to which he will ultimately apply his preferences.

The most closely related choice-theoretic paper to ours is Masatlioglu and Nakajima (2007), who characterize a class of choice correspondences that accommodates endogenously determined reference points. In particular, a decision maker who falls into this class behaves as if he applies a preference ranking to an endogenously determined "comparison set". This class of decision rules is quite general. First, it allows the comparison set to depend on the alternative under consideration. As a result, eventual choices are "fixed points": the decision maker chooses an alternative which is optimal given the comparison set assigned to this very alternative. Second, Masatlioglu and Nakajima are agnostic about the determinants of the comparison set, and therefore do not specify how an outside agent could manipulate it. At the same time, the generality of their framework allows them to subsume a large variety of reference-dependent decision models, including our consideration-sets procedure, as special cases. Indeed, from the point of view of the present paper, the contribution of Masatlioglu and Nakajima (2007) is to characterize the weakest possible decision model with consideration sets. In this sense (as well as in their focus on choice-theoretic analysis rather than market applications), their study complements ours.

3 Competitive Marketing

In this section we present a market model that incorporates the choice procedure presented in Section 2. Two identical firms compete for a continuum of measure one of identical consumers. Each firm i simultaneously chooses a menu M_i, which can be any non-empty subset of a finite set of products $X = \{1, 2, ..., n\}$, $n > 1$. Let c_x be a strictly positive fixed cost associated with adding the product x to the menu, where $c_1 > c_x$ for all $x \neq 1$. The cost of a menu M is thus $c(M) = \sum_{x \in M} c_x$. Each firm maximizes the fraction of consumers who choose a product from its menu, minus the costs associated with that menu. (We abstract from price setting - see a discussion in Section 5.2.) We assume that $c(X) < \frac{1}{2}$. That is, if a firm offers the grand set and gets a market share of 50%, it earns a strictly positive profit.

Consumers choose a product according to the consideration-sets procedure. Consumer preferences are strict: $1 \succ 2 \succ \cdots \succ n$. Slightly abusing the notation introduced in Section 2, $b(M)$ denotes the $\succ$-maximal product in M. Given a strategy profile (M_1, M_2), the choice problem is $(M_1, b(M_2))$ for one half of the population of consumers and $(M_2, b(M_1))$ for the other half. The interpretation is that for some consumers, firm 1 is the incumbent. Therefore, as the market interaction begins, their outside option is effectively the best product in firm 1's menu. The market interaction between the firms then determines whether these consumers will consider the unfamiliar firm 2. For the other consumers, firm 2 is the incumbent and firm 1 is the challenger. We impose a "*minimal richness*" condition on the consideration relation R: for every $x \in X$, there exists $y \in X$ (possibly x itself) such that yRx.[3]

The consumers' choice rule induces a strategic game between the firms. The minimal richness condition implies that the max-min strategy is the pure strategy $\{1\}$ and the max-min payoff is $\frac{1}{2} - c_1$. When the consumer is rational, these are also the Nash equilibrium strategy and Nash equilibrium payoffs, respectively. We refer to the Nash equilibrium outcome in the rational-consumer case as the "*competitive outcome*" and to $\frac{1}{2} - c_1$ as the "*competitive payoff*".

We will make use of the following terminology and notation. A menu M *beats* another menu M' if $xRb(M')$ for some $x \in M$ and $b(M) \succ b(M')$. A mixed strategy σ for a firm is a probability distribution over menus. The support of σ is denoted $S(\sigma)$. Given a mixed strategy σ, let $\beta_\sigma(x)$ and $\alpha_\sigma(x)$ denote the probabilities that x is offered as a $\succ$-maximal and as a $\succ$-inferior, "irrelevant" product. That is, $\beta_\sigma(x) =$

[3] Since R is assumed to be complete and transitive, minimal richness holds automatically for any product which is not R-maximal. The condition thus means that the consumer is willing to consider any R-maximal product at any R-maximal product.

$\sum_{M \in S(\sigma), x=b(M)} \sigma(M)$ and $\alpha_\sigma(x) = \sum_{M \in S(\sigma), x \in M \setminus \{b(M)\}} \sigma(M)$.

3.1 Equilibrium Analysis

We begin with a pair of examples that are based on two extreme cases of our model. First, suppose that for all $x, y \in X$, yRx if and only if $y \succ x$ or $y = x$. That is, the criteria underlying the consideration and preference relations perfectly coincide (except that $\succ$ is a linear ordering whereas R is reflexive). In this case, it is easy to check that the Nash equilibrium outcome is competitive.

For purely illustrative purposes, let us now examine the opposite case in which the criteria underlying the consideration and preference relations are diametrically opposed (except that as in the previous example, $\succ$ is a linear ordering whereas R is reflexive).

Remark 1 *Assume that for all $x, y \in X$, yRx if and only if $x \succ y$ or $x = y$. Assume further that $c_n < c_x$ for all $x \neq n$. Then, there is a unique symmetric Nash equilibrium strategy. The equilibrium strategy is mixed and given as follows:*

$$\begin{aligned} \sigma\{n\} &= 2c_n \\ \sigma\{1, n\} &= 2c_1 - 2c_n \\ \sigma\{1\} &= 1 - 2c_1 \end{aligned}$$

This equilibrium has several noteworthy properties. First, the equilibrium strategy is mixed and consumers end up choosing an inferior product with positive probability. Second, firms offer an "irrelevant alternative" with positive probability in equilibrium. This alternative is never chosen by the consumer, but it serves to attract his serious attention to a better product on the menu. Third, while the equilibrium outcome is not competitive, the firms' equilibrium payoff is at the competitive level. To see why, observe that the menu $\{1\}$ belongs to the support of the equilibrium strategy. Yet, the consumer is not willing to consider 1 at n, and therefore the market share that the menu $\{1\}$ generates is exactly $\frac{1}{2}$. Our task in this sub-section is to investigate the generality of these observations.

Non-competitive equilibrium outcomes

As we observed earlier, firms choose $\{1\}$ and earn a payoff of $\frac{1}{2} - c_1$ in Nash equilibrium when the consumer is rational, as well as when $\succ$ and R coincide. The following result states a necessary and sufficient condition for a competitive equilibrium outcome.

Proposition 1 *Firms play $\{1\}$ with probability one in Nash equilibrium if and only if $1Rx$ for every $x \neq 1$.*

Proof. Assume $1Rx$ for every $x \neq 1$. Clearly, if both firms play $\{1\}$, no firm has an incentive to deviate to another menu because $\{1\}$ beats any menu M with $b(M) \neq 1$, while any other menu M with $b(M) = 1$ attains the same market share and costs more. Now suppose that there exists a non-competitive (potentially asymmetric) Nash equilibrium (σ_1, σ_2). If both firms assign positive probability only to menus M with $b(M) = 1$, then $\{1\}$ is the unique best-reply for each firm, a contradiction. Therefore, at least one firm assigns positive probability to menus M with $b(M) \neq 1$. Let $M \in S(\sigma_1) \cup S(\sigma_2)$ such that $b(M') \succeq b(M)$ for all $M' \in S(\sigma_1) \cup S(\sigma_2)$. Without loss of generality, let $M \in S(\sigma_1)$. Note that $b(M) \neq 1$ and that M does not beat any menu in $S(\sigma_2)$. Suppose that firm 1 deviates from M to X. Then, the firm increases its market share by at least $\frac{1}{2}[\beta_{\sigma_2}(1) + (1 - \beta_{\sigma_2}(1))] = \frac{1}{2}$, because the deviation prevents being beaten by menus $M' \in S(\sigma)$ with $b(M') = 1$, and it allows beating any menu $M' \in S(\sigma)$ with $b(M') \neq 1$. The cost of this deviation is below $\frac{1}{2}$, hence the deviation is profitable.

Now assume $1 \not\!R x$ for some $x \neq 1$. If both firms play $\{1\}$, then it is profitable for any firm to deviate from $\{1\}$ to $\{x\}$, increasing its payoff from $\frac{1}{2} - c_1$ to $\frac{1}{2} - c_x$. ■

Thus, a non-competitive equilibrium outcome emerges when the preference and consideration relations induce a different ranking between the most preferred product and some other product in X.

Equilibrium use of irrelevant alternatives

As we saw in Section 2, the consumer's choice procedure violates Independence of Irrelevant Alternatives. The question arises, whether firms respond to this feature in equilibrium by offering irrelevant alternatives - i.e., products which affect the consumer's choice without themselves being chosen. Our next result answers in the affirmative. This result is preceded by a lemma stating that in any symmetric equilibrium, firms offer the most preferred product with positive probability.

Lemma 1 *Let σ be a symmetric Nash equilibrium strategy. Then, $\beta_\sigma(1) > 0$.*

Proof. Assume the contrary. Given symmetry of equilibrium, a firm's equilibrium payoff is below $\frac{1}{2}$. If the firm deviates to the pure strategy X, it ensures that all consumers choose it, yielding a payoff of $1 - c(X) > \frac{1}{2}$. ■

Proposition 2 *If $1 \not R x$ for some $x \neq 1$, then in any symmetric Nash equilibrium σ, $S(\sigma)$ contains a non-singleton menu.*

Proof. Assume the contrary that $S(\sigma)$ consists of singletons only. By Lemma 1, $\{1\} \in S(\sigma)$. At the same time, by Proposition 1, there exists $x \neq 1$ such that $\sigma(\{x\}) > 0$. Suppose that a firm deviate from $\{1\}$ to X. The cost of this deviation is $c(X) - c_1$. The benefit from this deviation is $\frac{1}{2}\sum_{x \neq 1, 1 \not R x} \sigma(\{x\})$. The firms' decision not to carry out this deviation implies that $\frac{1}{2}\sum_{x \neq 1, 1 \not R x} \sigma(\{x\}) \leq c(X) - c_1$.

Suppose that $\sigma(\{x\}) > 0$ for some $x \neq 1$, $1Rx$. Let x^* denote the $\succ$-minimal such product. Suppose that a firm deviates from $\{x^*\}$ to $\{1, x^*\}$. The cost of this deviation is c_1. The benefit from this deviation is at least $\frac{1}{2}\sigma\{1\} + \frac{1}{2}\sum_{x \neq 1, 1Rx} \sigma(\{x\})$, because the deviation prevents being beaten by $\{1\}$ and allows beating any $\{x\}$ with $x \neq 1$, $1Rx$. From the firms' decision not to carry out this deviation, we conclude that $\frac{1}{2}\sigma\{1\} + \frac{1}{2}\sum_{x \neq 1, 1Rx} \sigma(\{x\}) \leq c_1$.

Combining the two inequalities, we obtain $\frac{1}{2} \leq c(X)$, a contradiction. It follows that $\sigma(\{x\}) = 0$ for every $x \neq 1$, $1Rx$. Therefore, $\{1\}$ generates a payoff of $\frac{1}{2} - c_1$. Now consider the $\succ$-maximal product $y \neq 1$ for which $\sigma(\{y\}) > 0$. By our previous step, $1 \not R y$. Therefore, $\{y\}$ generates a payoff of at least $\frac{1}{2} - c_y > \frac{1}{2} - c_1$. a contradiction. ■

Since an irrelevant alternative is costly to offer and never chosen by the consumer, the equilibrium use of irrelevant alternatives is socially wasteful. However, it is an integral part of the firms' competitive strategy whenever the condition for a competitive equilibrium outcome is not met.

It is interesting to compare the role of irrelevant alternatives in our model to the role of "loss leaders" in papers such as Lal and Matutes (1994). In both cases, the direct revenues from the product do not cover its costs, yet firms offer this product because it enables them to earn a profit from another product. However, loss leaders in these models are not "irrelevant alternatives", as consumers purchase them with certainty.

Competitive equilibrium payoffs

Our results so far relied only on the minimal richness condition. The next result is the first to utilize the assumption that R is complete and transitive. Recall that $\frac{1}{2} - c_1$ is the max-min payoff under any R. Thus, firms can never earn less than the competitive payoff in symmetric Nash equilibrium. Does this mean that when the conditions for a competitive equilibrium outcome are not met, the outcome is *collusive* from the firms' point of view? The answer turns out to be negative.

Proposition 3 *Firms earn a payoff of $\frac{1}{2} - c_1$ in any symmetric Nash equilibrium.*

Proof. Assume the contrary. By Lemma 1, the set of menus M$= \{M \in S(\sigma) \mid b(M) = 1\}$ is non-empty. Every menu in $\mathcal{M}$ beats some other menu in $S(\sigma)$ - otherwise, the firms' equilibrium payoff could not exceed the competitive level. Define the menu $M^* \in S(\sigma)$ as follows. For every other menu M in $S(\sigma)$, $b(M)Rb(M^*)$, and either *(i)* $b(M^*)\not{R}b(M)$ or *(ii)* $b(M^*)Rb(M)$ and $b(M) \succ b(M^*)$. By the completeness and transitivity of R, M^* is well-defined. Moreover, M^* does not beat any other menu in $S(\sigma)$, and it is beaten by every menu in $\mathcal{M}$. Therefore, if a firm deviates from M^* to the menu X, it increases its market share by at least $\frac{1}{2}\beta_\sigma(1) + \frac{1}{2}(1 - \beta_\sigma(1)) = \frac{1}{2}$, which by assumption is strictly higher than the added menu cost $c(X) - c(M^*)$. Thus, the deviation is profitable, a contradiction. ■

Thus, firms earn competitive payoffs in symmetric Nash equilibrium. This result is important for several reasons. First, it carries an immediate welfare implication. On one hand, Proposition 1 implies that when $1\not{R}x$ for some $x \neq 1$, consumers select inferior products with positive probability, hence they are strictly worse off than in the full-rationality benchmark. On the other hand, Proposition 3 implies that firms earn competitive payoffs. It follows that whenever the conditions for a competitive equilibrium outcome are not met, the symmetric equilibrium outcome is necessarily *Pareto inferior* to the rational-consumer equilibrium outcome.

The competitive-payoff result also has implications for market entry. Although our model abstracts from entry concerns, one could construct more elaborate, multiple-industry models in which consumers in any individual market behave according to the consideration-sets procedure, where the consideration relation R could be market-specific. One might suspect a priori that consumers' departure from rational choice could lead to a distortion in the firms' market entry decisions. However, Proposition 3 implies that the firms' decision whether to enter a market is the same as when consumers are rational.

Effective marketing

Proposition 3 implies another, subtler corollary, which concerns the equilibrium effectiveness of irrelevant alternatives as marketing devices. Firms in our model use irrelevant alternatives as marketing devices whose objective is to "get the customer into the store". Clearly, the fact that a marketing device attracts the consumer's serious attention to a store does not automatically guarantee that he will *buy* at that store. When this event occurs nonetheless, we say that marketing is "effective".

Definition 1 (Effective Marketing) *A mixed strategy σ satisfies the effective marketing property if for every $M, M' \in S(\sigma)$ satisfying $b(M') \neq 1$, $b(M) \not R b(M')$ and $xRb(M')$ for some $x \in M$, M beats M'.*

The effective marketing property means that whenever a consumer considers a firm only because of an irrelevant alternative it offers, he ends up buying from that firm (unless his reference point is the best possible product). To see how this property could be violated, let $X = \{1, 2, 3, 4\}$, $1 \succ 2 \succ 3 \succ 4$, $4R2$ and $3\not R2$. Suppose that $\{2\}, \{3, 4\} \in S(\sigma)$. The choice problem for half of the consumers is $(\{3, 4\}, 2)$. These consumers add $\{3, 4\}$ to their consideration set only because this menu contains the irrelevant alternative 4, yet they end up choosing their reference point 2. This means that σ violates the effective marketing property. Our next result demonstrates that such a state of affairs is impossible in symmetric equilibrium.

Proposition 4 *Any symmetric Nash equilibrium strategy satisfies the effective marketing property.*

Proof. Assume the contrary - i.e., there exist menus $M, M' \in S(\sigma)$ such that $b(M) \succeq b(M')$, $b(M) \neq 1$, $b(M')\not R b(M)$, and $xRb(M)$ for some $x \in M'$. That is, consumers expand the consideration set from M to $M \cup M'$ thanks to some irrelevant alternative x in M', yet end up choosing from M. The marginal contribution of x to the market share generated by M' is at most $\frac{1}{2} \sum_{xRy,\ b(M') \succ y} \beta_\sigma(y)$. By the fact that firms choose to include x in M' as part of a best-reply to σ, we conclude that $\frac{1}{2} \sum_{xRy,\ b(M') \succ y} \beta_\sigma(y) \geq c_x$. Now suppose that a firm deviates to the menu $\{1, x\}$. By Proposition 3, firms earn competitive payoffs in equilibrium. Therefore, $\beta_\sigma(x) = 0$ for all $x \neq 1$, $1Rx$ - otherwise, the menu $\{1\}$ would yield a payoff above $\frac{1}{2} - c_1$. Therefore, the marginal contribution of x to the market share generated by the menu $\{1, x\}$ is at least $\frac{1}{2}\beta_\sigma(b(M)) + \frac{1}{2} \sum_{xRy,\ b(M) \succ y} \beta_\sigma(y)$, which strictly exceeds c_x. Therefore, the menu $\{1, x\}$ yields a payoff strictly above $\frac{1}{2} - c_1$, a contradiction. ■

The next result presents a different sense in which equilibrium marketing is "effective". If two products belong to the same menu in the support of a symmetric equilibrium strategy, the sets of products from which they attract serious attention in equilibrium are necessarily disjoint. In other words, firms exercise specialization in their equilibrium use of marketing devices.

Proposition 5 *Let σ be a symmetric Nash equilibrium strategy. Then, for any non-singleton menu $M \in S(\sigma)$ and any pair of products $x, x' \in M$, the sets $\{y \in X \mid \beta_\sigma(y) > 0$ and $yRx\}$ and $\{y \in X \mid \beta_\sigma(y) > 0$ and $yRx'\}$ are disjoint.*

Proof. Assume the contrary - i.e., that there exists a menu $M \in S(\sigma)$ containing two products x, x', $x' \succ x$, such that the sets $\{y \in X \mid \beta_\sigma(y) > 0$ and $yRx\}$ and $\{y \in X \mid \beta_\sigma(y) > 0$ and $yRx'\}$ have non-empty intersection. Therefore, the marginal contribution of x to the market share generated by M is strictly below $\sum_{xRy,\ b(M)\succ y} \beta_\sigma(y)$. >From the firms decision to include x in M as part of a best-reply to σ, we conclude that $\sum_{xRy,\ b(M)\succ y} \beta_\sigma(y) > c_x$. Using the same argument as in the proof of Proposition 4, a deviation to $\{1, x\}$ is strictly profitable. ■

It should be emphasized that Propositions 4 and 5 rely on the assumption that R is complete and transitive only indirectly. They are direct corollaries of Proposition 3, and would therefore hold under any consideration relation that induces competitive payoffs in symmetric equilibrium.

Summary

The results in this sub-section shed light on the nature of competitive behavior when consumers follow the consideration-sets procedure. As long as consumers are not willing to consider the most preferred product at all other products, the equilibrium outcome fails to be competitive. However, it does retain an aspect of competitiveness, in the sense that firms earn the same equilibrium payoff as in the rational-consumer benchmark. Irrelevant alternatives are employed with positive probability. It is the use of irrelevant alternatives as a marketing device that generates the competitive force that brings industry profits to the competitive level. Another aspect of competitive equilibrium behavior is that the deployment of marketing devices is "cost-effective". First, when an irrelevant alternative is responsible for "getting the consumer into the store", he ends up buying there. Second, different products on a firm's menu attract consumers away from different outside options.

3.2 Other Classes of Consideration Relations

So far we restricted attention to complete and transitive consideration relations, reflecting the idea that consumers' resistance to considering new alternatives finds expression in a demand for a *reason* that would justify looking into M. However, the formation of consideration sets may be based on alternative criteria. Chakravarti and Janiszewski (2003) present experimental evidence suggesting that when people are asked to choose from a large set of diverse alternatives, they tend to focus on a small subset of "easy-to-compare" options having alignable or overlapping attributes. A pair of alternatives is more likely to enter the consideration set as their alignability or number of overlapping features increase. One implication of this finding is that consumers may be reluctant to consider products that are hard to compare with their outside option.

In this sub-section we relax the assumption that R is complete and transitive. We retain the minimal richness condition - for every $x \in X$, there exists $y \in X$ (possibly x itself) such that yRx. As already mentioned, some of the results in the previous sub-section do not rely on the completeness or transitivity of R. Specifically, Propositions 1 and 2, as well as Lemma 1, only rely on minimal richness. Furthermore, our results on "effective marketing" (Propositions 4 and 5) are corollaries of the competitive-payoff result. If, for whatever reason, firms earn competitive payoffs in symmetric Nash equilibrium under some R, these two propositions continue to hold.

The remainder of this section thus deals with the following question: *Does the competitive-payoff result survive the extension to other families of consideration relations*?

Similarity-based consideration relations

The findings of Chakravarti and Janiszewski (2003) suggest that the formation of consideration sets may be based on similarity judgments. Since there is no single obvious definition of similarity, and different intuitions about similarity give rise to different definitions (see Tversky (1977) and Rubinstein (1988)), we propose two examples.

1. Equivalence relations. A consideration relation R is an equivalence relation if it is reflexive, symmetric and transitive. This case fits situations in which products are divided into mutually exclusive categories, such that two products are deemed similar if they belong to the same category.

2. Linear similarity. Let $\psi : X \to \mathbb{R}$ be a one-to-one function that assigns to each product a distinct location along the real line. Assume that for every $x \in X$, yRx if $\psi(y)$ belongs to some arbitrary neighborhood of $\psi(x)$. We refer to such R as a linear similarity relation. It fits situations in which products are represented by points on a efficiency frontier in $\mathbb{R}^2_{++}$ (e.g., soft drinks are characterized by tastiness and healthiness, such that the tastier the beverage, the less healthy it is), and two products are similar if they lie close to each other along the frontier. In general, a linear similarity relation need not be symmetric, transitive or complete. However, it is obviously reflexive, since any neighborhood of $\psi(x)$ necessarily contains itself.

Proposition 6 *Suppose that R is an equivalence relation or a linear similarity relation. Then, firms earn a payoff of $\frac{1}{2} - c_1$ in any symmetric Nash equilibrium.*

The identity relation (yRx if and only if $x = y$) is a special case of both equivalence and linear similarity relations. It turns out that symmetric Nash equilibria in this case have a particularly clean characterization.

Proposition 7 *If R is the identity relation, then in any symmetric Nash equilibrium σ, $\beta_\sigma(x) = 2c_x$ and $\alpha_\sigma(x) = 2c_1 - 2c_x$ for all $x \neq 1$.*

Thus, all products $x \neq 1$ are offered with the same probability $2c_1$ in symmetric equilibrium. The more costly the product, the higher (lower) the probability it is offered as a $\succ$-maximal product (an irrelevant alternative).

Given the equilibrium characterization of $\beta_\sigma(\cdot)$ and $\alpha_\sigma(\cdot)$, we can calculate the fraction of consumers who switch a supplier in symmetric equilibrium, denoted $\lambda(\sigma)$. In order for a consumer to switch from one firm to the other, the best product in the former's menu must be offered by the other firm as an irrelevant alternative. This leads to the following expression:

$$\lambda(\sigma) = \sum_{x \neq 1} \beta_\sigma(x)\alpha_\sigma(x) = \sum_{x \neq 1} 4c_x(c_1 - c_x)$$

Our assumptions on menu costs ensure that $\lambda(\sigma) \in (0, 1)$. Thus, consumers switch suppliers in equilibrium. By comparison, no switching occurs in the rational-consumer benchmark. Note that $\lambda(\sigma)$ behaves non-monotonically in menu costs, and approaches an upper bound of $(n-1) \cdot c_1^2$ as the costs of all products $x \neq 1$ cluster near $c_1/2$. The reason for this non-monotonicity is that as a product becomes more costly, it is offered less frequently as an irrelevant alternative and more frequently as a $\succ$-maximal product.

The switching fraction is exactly equal to the expected cost of irrelevant alternatives, because for each $x \neq 1$, the probability it is offered as an irrelevant alternative by each firm is by definition $\alpha_\sigma(x)$, while by Proposition 7, $\beta_\sigma(x)$ is equal to twice the cost of x. Thus, the general relation between the social cost of irrelevant alternatives and their role in attracting consumers' attention is especially transparent when R is the identity relation, because the "deadweight loss" associated with irrelevant alternatives is equal to consumers' switching frequency.

The identity relation is an example of a reflexive binary relation - i.e., xRx for all $x \in X$. We say that a consideration relation R' is *richer* than R if yRx implies $yR'x$ for all $x, y \in X$. Any reflexive relation is richer than the identity relation, and the consideration relation that fits a rational consumer is the richest of them all. In the two extreme cases of the identity relation and the rational-consumer relation, firms earn competitive payoffs in symmetric Nash equilibrium. Intuitively, as the consideration relation becomes richer, the market friction due to consumers' bounded rationality gets weaker, because the consideration set coincides with the feasible set for a larger family of strategy profiles. Therefore, one might expect that the competitive payoff result would hold for all reflexive consideration relations.

This intuition turns out to be false, as the following counter-example demonstrates. Let X be the set of all three-digit binary numbers, excluding 000. Assume that $111 \succ x$ for every $x \neq 111$. Also, assume that yRx if x and y have at least two identical digits. This R is richer than the identity relation. Assume that $c_{111} = \frac{1}{3}$, whereas $c_x = c < \frac{1}{30}$ for all $x \neq 111$. It can be shown that there exists a continuum of symmetric equilibria with the following properties: *(i)* the support of the equilibrium strategy consists of $\{111, 110\}$, $\{111, 101\}$, $\{111, 001\}$, $\{100\}$, $\{010\}$ and $\{001\}$; *(ii)* the equilibrium payoff is strictly above the competitive level of $\frac{1}{2} - c_{111}$. (There is also a symmetric equilibrium that induces competitive payoffs.)

Small menu costs

The previous counter-example means that in general, firms may earn above-competitive equilibrium payoffs under an arbitrary consideration relation. We conjecture that for generic cost structures, firms indeed earn competitive payoffs in symmetric Nash equilibrium under any minimally rich consideration relation. So far, we have been able to obtain a general competitive-payoff result for sufficiently low menu costs.

Proposition 8 *Suppose that R satisfies the minimal richness condition. Then, firms earn a payoff of $\frac{1}{2} - c_1$ in symmetric Nash equilibrium whenever $c_1 < 1/(2^n + 2n)$.*

This result is somewhat unsatisfactory, in the sense that when menu costs are sufficiently small, the outcome itself is close to the competitive benchmark, as our next result demonstrates.

Proposition 9 *As c tends to zero, $\beta_\sigma(1)$ converges to one under any symmetric Nash equilibrium strategy σ.*

Thus, a competitive-payoff result that holds only when the outcome is close to competitive anyway takes the sting out of the distinction between competitiveness of the market outcome and competitiveness of industry profits. The (open) problem is what kind of a general competitive-payoff result could emerge under our original, significantly weaker assumptions on menu costs.

4 Competitive Advertising

The model in Section 3 assumes strict consumer preferences. In particular, this rules out configurations such as $x \sim y$, xRz and $y\not{R}z$. One interpretation of such a configuration is that x are y are the same product as far as the consumption experience is concerned, yet they differ in the way they are *framed*. Thus, allowing weak preferences enables us to capture situations in which firms decide not only which product to sell but also how to frame it for consumers. We now use this basic insight to construct a model of competitive advertising.[4]

In a classical model of price competition with advertising, Butters (1977) assumes that firms attract consumers' attention by posting ads. Butters' "advertising technology" can be embedded in our framework: every product x comes in two variants, advertised and unadvertised, denoted x^a and x^u, such that: *(i)* $x^a \sim x^u$, and *(ii)* for any $x, y \in X$ and any $i \in \{a, u\}$, $x^a R y^i$ and $x^u \not{R} y^i$. However, our model goes beyond Butters in providing a language for formulating a richer family of "advertising technologies".

Let $F = \{1, ..., K\}$ be a finite set of product features. Let X be the set of all pairs (P, A), where $A \subseteq P \subseteq F$ and $P \neq \varnothing$. Firms simultaneously choose pairs (P, A) from X. A firm's cost of offering (P, A) is $c_{(P,A)} = c_p \cdot |P| + c_a \cdot |A|$, where $c_p, c_a > 0$. Assume that $K(c_p + c_a) < \frac{1}{2}$. Given a strategy profile (x_1, x_2), the consumer's choice problem is equally likely to be $(\{x_1\}, \{x_2\})$ or $(\{x_2\}, \{x_1\})$. Assume that $(P, A) \succ (P, A')$ if $|P| > |P'|$. When consumers are rational, firms play $(F, \varnothing)$ in Nash equilibrium and earn a payoff of $\frac{1}{2} - Kc_p$, which is henceforth referred to as the "competitive payoff".

As before, we restrict attention to complete and transitive consideration relations. In addition, we assume that $(F, F)R(P, A)$ for every (P, A). That is, if a firm produces the highest-quality product and advertises all its features, it secures the consumer's attention. The following are examples of consideration relations that satisfy these conditions: *(i)* $(P', A')R(P, A)$ if $|A'| \geq |P|$; *(ii)* $(P', A')R(P, A)$ if $|A'| \geq |A|$. In example

[4]For an alternative framework that extends the standard choice model to incorporate framing effects, see Rubinstein and Salant (2007).

(i), the decision whether to consider the new product depends on some relation between its advertised features and the actual features of the outside option. (This relation happens to be mirror the criterion underlying actual preferences, but this need not be the case - see below.) In example *(ii)*, the decision depends on an analogous relation between the advertised features of both products. The interpretation of the former case is that the consumer has experienced the product (P, A) and therefore he is familiar with its actual features. In order to persuade him to consider the new product, the ads associated with it must declare that it is better than the original product, according to some preference criterion. The interpretation of the latter case is that although the consumer initially consumes one product, he is not intimately familiar with its features at the time he starts considering alternatives. This fits telecommunication or banking services, where the consumer does not actually experience features such as the price per minute or banking fees, hence his initial impression is purely based on advertising.

The main result in this section establishes that firms earn competitive payoffs in symmetric Nash equilibrium.

Proposition 10 *Firms earn a payoff of $\frac{1}{2} - Kc_p$ in any symmetric Nash equilibrium.*

Proof. The proof follows the same logic as the proof of Proposition 3, which is the analogous result for the model of Section 3. Let σ be a symmetric Nash equilibrium (mixed) strategy, and let $S(\sigma)$ denote its support. Our first step is to show that there exists a subset $A \subseteq F$ such that $(F, A) \in S(\sigma)$. Assume the contrary. Then, it is profitable for a firm to deviate from one of the $\succsim$-minimal elements (P, A) in $S(\sigma)$ to (F, F) - the deviation will increase the firm's market share by $\frac{1}{2}$, whereas the added cost is by assumption below $\frac{1}{2}$.

It is easy to see that if $(F, \varnothing) \in S(\sigma)$, firms earn $\frac{1}{2} - Kc_p$ under σ. It follows that in order for firms to earn a payoff above the competitive level, for each element of the form (F, A) in $S(\sigma)$ there must exist some $(P', A') \in S(\sigma)$ such that $P' \subset F$ and $(F, A)R(P', A')$. By completeness and transitivity of R, we can identify a class of R-minimal elements in $S(\sigma)$. Among these elements, select some $\succsim$-minimal element (P, A). By our previous argument, $P \neq F$. By construction, (P, A) fails to beat any element in $S(\sigma)$, and it is beaten by all elements of the form (F, A) in $S(\sigma)$. Therefore, if a firm deviates from (P, A) to (F, F), it raises its market share by $\frac{1}{2}$ while increasing its cost by less than $\frac{1}{2}$, a contradiction. ■

For a simple illustration of this result, consider the following example: $(P', A')R(P, A)$ if and only if $\max(A') \geq \max(P)$. The interpretation is that when the consumer decides whether to consider a new product, he relies on an intuitive ranking of features.

According to this ranking, feature k is more important than feature j if $k > j$. The consumer's preference ranking, in contrast, treats all features symmetrically. The following mixed strategy is a symmetric equilibrium strategy:

$$\begin{aligned}\sigma(\{K\}, \varnothing) &= 2c_a \\ \sigma(F, \{K\}) &= 2(K-1)c_p \\ \sigma(F, \varnothing) &= 1 - 2c_a - 2(K-1)c_p\end{aligned}$$

In fact, it can be shown that this is essentially the unique symmetric equilibrium under this specification of R. Any other symmetric equilibrium is the same as given above, except that $(\{K\}, \varnothing)$ is replaced with any collection of $(\{k\}, \varnothing)$, such that $\sum_k \sigma(\{k\}, \varnothing) = 2c_a$. Thus, firms produce the most or least preferred products, and avoid offering intermediate-quality products. When a firm offers the most preferred product, it either advertises the most important feature according to the consideration relation, or it avoids advertising altogether.

Note that this equilibrium satisfies an effective marketing property: whenever the consumer's reference point is not the most preferred product, considering a new product only because of its advertising culminates in buying that product. However, the effective marketing property does not necessarily hold in the competitive advertising model. The reason is that in the competition-in-menus model, irrelevant alternatives affect the fraction of consumers who switch *away* from the rival firm, but they cannot affect the fraction of consumers who switch *to* the rival firm. In contrast, in the current model, firms' advertising decision can affect consumer switching in both directions.

5 Variations and Extensions

5.1 Alternative Models of Consideration-Set Formation

The choice procedure presented in Section 2 is arguably the simplest way to introduce the idea of consideration sets into a model of consumer choice. In this sub-section we discuss a pair of somewhat more complicated variants. First, suppose that we replace the reference point s with a *reference set* of products S. In this variant, the consideration set is $S \cup M$ whenever yRx for some $x \in S$ and $y \in M$, and S otherwise. The interpretation is that at the time the consumer determines whether to consider the new menu M, he has not yet reached a tentative choice from the set of products he is already familiar with. Since every product in the reference set is a potential choice, it makes sense to apply the consideration relation to all of them. This variant is particularly suited to the competition-in-menus model of Section 3, where it is natural to identify S with the menu offered by the consumer's "incumbent" firm. It turns out that under the restriction to complete and transitive R, all the results in Section 3 continue to hold. Proofs are available upon request.

Another variant addresses our assumption that once the consumer finds a product $x \in M$ such that xRs, he considers the *entire* menu M. This variant assumes that the consumer constructs the consideration set iteratively: $CS^0 = \{s\}$ and for any $k > 0$, $CS^k = CS^0 \cup \{y \in M \mid yRx \text{ for some } x \in CS^{k-1}\}$. In this case, the consideration set ends up being the largest subset of $M \cup \{s\}$ having the property that for every $x \in M$ there is $y \in M \cup \{s\}$ such that xRy. The interpretation is that even when "inside the store", the consumer continues to display his original reluctance to consider new alternatives. When R is complete and transitive, this variant curbs the incentive to use irrelevant alternatives. In fact, it can be shown that when R is transitive, firms would not offer products that the consumer never chooses in the model of Section 3.

5.2 Price Setting

Throughout this paper, we abstracted from price setting and assumed that firms try to maximize market share minus fixed costs. Our motivation was analytic simplicity. As in older marketing models - a prime example of which is the Hotelling strategic location model - it is easier to start by assuming that firms care only about market share and defer the incorporation of price setting. The reason price setting complicates the model is that when a firm decides whether to add an irrelevant alternative to its menu, it needs to know its payoff when consumers choose the $\succ$-maximal product on the menu. In the model of Section 3, this payoff is fixed, whereas in a model with price setting, it depends on the price associated with the $\succ$-maximal element.

Incorporating price setting is of particular interest because many natural specification of the consideration relation involve price comparisons. For instance, consider the following extended model. Let $X = Z \times [0, \infty)$, where Z is a finite set of product types and $[0, \infty)$ is the set of feasible prices. Consumer preferences satisfy the following condition: $(z, p) \succ (z, p')$ whenever $p < p'$. A natural example of a complete and transitive consideration relation in this context is $(z, p)R(z, p')$ if and only if $p \leq p'$. Another plausible definition is $(z', p')R(z, p)$ if and only if $z = z'$ and $p' < p$ (under this alternative definition, R is transitive but incomplete). One would like to know whether the main results of Section 3 persist under these assumptions. In addition, the extended model allows for both "product dispersion" (different products being offered with positive probability) and "price dispersion" (the same product being offered at two different prices). Characterizing the equilibrium patterns of product and price dispersion is an interesting challenge for future research.

5.3 Consumer Heterogeneity

Since consumers in our model are characterized by two primitives, $\succsim$ and R, heterogeneity may exist in both dimensions. Consider heterogeneity in R first. It can be shown that in the model of Section 3, if consumers differ only in their consideration relations and the distribution of consumer types is held fixed, firms earn the competitive payoff $\frac{1}{2} - c_1$ in any symmetric Nash equilibrium for sufficiently low menu costs. In contrast, it is not true that for a fixed cost structure, if the competitive-payoff result holds for all consideration relations in the collection $\mathcal{R} = \{R^1, .., R^K\}$, then it must also hold for a heterogeneous consumer population with some distribution over $\mathcal{R}$. To see why, let $X = \{1, 2\}$ and suppose that R^1 is the identity relation whereas yR^2x for all $x, y \in X$. It is easy to show that for a certain range of distributions over $\{R^1, R^2\}$, the symmetric Nash equilibrium strategy has a support $\{\{1\}, \{2\}, \{1, 2\}\}$. This means that firms earn an expected payoff above the competitive level, because the menu $\{1\}$ attracts R^2-consumers away from the menu $\{2\}$.

Now consider the case of heterogeneity in $\succsim$. The reason we avoided heterogeneous consumer preferences is that we wanted to have a clear rational-consumer benchmark, and the equilibrium outcome under consumer rationality is sensitive to the exact distribution over consumer preferences. Some of our results are extreme as a result of this modeling strategy. First, when the set of preferences in the population is sufficiently rich, every product will be found to be optimal by some consumers, and therefore no menu can contain totally "irrelevant" alternatives. Second, the effective marketing property is unlikely to hold under heterogeneous preferences. One of the challenges of extending our models in this direction is to formulate analogues of these two effects when consumer preferences are diverse.

5.4 Endogenizing the Consideration Relation: Internet Search

Our model is non-standard in the sense that it introduces a new primitive into a model of consumer choice. The reader may wonder to what extent our model can be "rationalized". In particular, can the consideration relation be *endogenized* as a best-reply to market environments such as those studied in Sections 3 and 4? If one views R as a description of a personality trait such as "openness to experience", then endogenizing R has the same urgency as endogenizing $\succsim$. In this section, however, we propose an alternative interpretation of R, in which case it may seem desirable to derive R from more conventional primitives.

The alternative interpretation is based on the notion of consumer search. Models of search are typically based on some concrete image of the underlying search technology. Sequential-search models evoke a picture of a consumer literally going from one store to another. Models of simultaneous search conjure up an image of a consumer sitting at home and calling up a sample of stores. In the case of internet search, a more relevant image is of a process that takes the form of a query. When the consumer browses for a substitute to his outside option, he submits a query through a search engine. For instance, he may enter keywords which represent features that the potential substitute might share with the outside option. The set $\{x \in M \mid xRs\}$ thus consists of the relevant matches that the consumer's query elicits. A higher number of keywords corresponds to a thinner R and implies a narrower, more focused search. This means that the search will be less costly, but it also means that the consumer is more likely to miss relevant matches.

Bearing in mind this search-based interpretation, consider the following two-stage variant on the competition-in-menus model of Section 3. In stage 1, firms choose menus just as in our model, and consumers simultaneously choose a search intensity level $l \in \{1, ..., L\}$. Each search intensity l corresponds to some consideration relation R^l, such that a higher l corresponds to a richer R^l, and L corresponds to the rational-consumer case (yR^Lx for all $x, y \in X$). Search cost increases with its intensity. In stage 2, consumers choose according to our procedure, given the R they effectively chose in stage 1.

It is easy to show that if search costs are sufficiently high, consumers choose the lowest possible search intensity $l = 1$ in equilibrium, and so the search-theoretic model is trivially reduced to our model with a consideration relation R^1. Conversely, as search costs tend to zero, consumers necessarily choose $l = L$ with positive probability in stage 1. However, this probability is strictly below one. To see why, assume the contrary. Then, consumers choose rationally as if there were no search costs. Firms' best-reply must be to play $\{1\}$ with probability one. But if this is the case, then it is optimal for consumers to choose $l = 1$ rather than $l = L$, a contradiction.

This result means that when search costs are small, the search-theoretic model is consistent with the extended competition-in-menus model with *heterogeneous consumers* discussed in the previous sub-section, where the distribution of R is derived from the consumers' first-stage mixed equilibrium strategy. The result also means that the firms' equilibrium payoff is necessarily *above the competitive level*. The reason is that in equilibrium, consumers choose the maximal search intensity with positive probability, and therefore if a firm plays $\{1\}$, it generates a market share above $\frac{1}{2}$. This result is somewhat surprising: lower search costs lead to higher industry profits.

It should be emphasized that although the methodology embodied in this search model is standard, the actual model is unconventional both in its search technology and in the effect that the consumer's sample is influenced by irrelevant alternatives. However, these aspects make the model suited for analyzing internet search. In particular, the use of irrelevant alternatives can be viewed as an instance of the branch of internet marketing known as "*search engine optimization*", which is concerned with firms' strategic reaction to the way consumers employ search engines. Developing an internet search model based on our consideration-sets formalism is left for future work.[5]

References

[1] Athey, S. and G. Ellison (2007): "Position Auctions with Consumer Search", mimeo, MIT.

[2] Bertrand, M., D. Karlan, S. Mullainathan, E. Shafir and j. Zinman (2006), "What's Psychology Worth? A Field Experiment in the Consumer Credit Market," mimeo, Yale University.

[3] Butters, G. (1977): "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies* **44**, 465-491.

[4] Chakravarti, A. and C. Janiszewski (2003): "The Influence of Macro-Level Motives on Consideration Set Composition in Novel Purchase Situations," *Journal of Consumer Research* **30**, 244-258.

[5] DellaVigna, S. and U. Malmendier (2004): "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics* **119**, 353-402.

[6] Eliaz, K. and R. Spiegler (2005): "Speculative Contracts," mimeo, New York University and University College London.

[5] See Athey and Ellison (2007) for a recent attempt to combine a model of internet search and a model of position auctions for search-engine keywords. Their model abstracts from search engine optimization.

[7] Eliaz, K. and R. Spiegler (2006): “Contracting with Diversely Naive Agents,” *Review of Economic Studies* **73**, 689-714.

[8] Gabaix, X. and D. Laibson (2006): “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets,” *Quarterly Journal of Economics* **121**, 505-540.

[9] Goldberg, L. (1993): “The structure of phenotypic personality traits,” *American Psychologist* **48**, 26-34.

[10] Kamenica, E. (2007): “Contextual Inference in Markets: On the Informational Content of Product Lines,” mimeo, University of Chicago.

[11] Lal, R. and C. Matutes (1994): “Retail Pricing and Advertising Strategies,” *Journal of Business* **67**, 345-370.

[12] Manzini, P. and M. Mariotti (2007): “Sequentially Rationalizable Choice,” *American Economic Review,* forthcoming.

[13] Masatlioglu, Y. and E. Ok (2005): “Rational Choice with Status Quo Bias,” *Journal of Economic Theory* **121**, 1-29.

[14] Masatlioglu, Y. and D. Nakajima (2007): “Choice by Constraint Elimination,” mimeo, University of Michigan.

[15] Mullainathan S., J. Schwartzstein and A. Shleifer (2007), “Coarse Thinking and Persuasion,” forthcoming in *Quarterly Journal of Economics.*

[16] Piccione, M. and A. Rubinstein (2003), “Modeling the Economic Interaction of Agents with Diverse Abilities to Recognize Equilibrium Patterns,” *Journal of European Economic Association* **1**, 212-223.

[17] Roberts, J. and J. Lattin (1997): “Consideration: Review of Research and Prospects for Future Insights,” *Journal of Marketing Research* **34**, 406-410.

[18] Rubinstein, A. (1988), “Similarity and Decision-Making Under Risk,” *Journal of Economic Theory* **46**, 145-153.

[19] Rubinstein, A. and Y. Salant (2007), “(A,f): Choices with Frames,” mimeo, NYU.

[20] Rubinstein, A. and L. Zhou (1999), “Choice Problems with a "Reference" Point,” *Mathematics of Social Sciences* **37**, 205-210.

[21] Shapiro, J. (2006): "A 'Memory-Jamming' Theory of Advertising," mimeo, University of Chicago.

[22] Spiegler, R. (2005), "The Market for Quacks," *Review of Economic Studies* **73**, 1113-1131.

[23] Spiegler R. (2006), "Competition over Agents with Boundedly Rational Expectations," *Theoretical Economics* **1**, 207-231

[24] Tversky A. (1977), "Features of Similarity," *Psychological Review* **84**, 327–352.

Appendix: Proofs

All the proofs in the Appendix pertain to results stated in Section 3. Note that they make use of Lemma 1, the proof of which was given in the main text. In what follows, σ denotes a symmetric Nash equilibrium strategy.

Proof of Remark 1

Let us begin with two observations. First, note that R violates the necessary and sufficient condition for a competitive equilibrium outcome given by proposition 1. Therefore, $\beta_\sigma(x) > 0$ for some $x \neq 1$. Second, any menu in $S(\sigma)$ is either a singleton or a menu of the form $\{x, n\}$, where $x \neq n$. Otherwise, it is profitable to substitute any irrelevant alternative with n, thus saving menu costs without damaging market share.

Our next step is to show that $\{1\} \in S(\sigma)$. By Lemma 1, if $\{1\} \notin S(\sigma)$, then $\{1, n\} \in S(\sigma)$. It follows that $\{1, n\}$ beats any other menu $M \in S(\sigma)$. Among these menus M, let M^* be a menu with a $\succ$-minimal $b(M)$. If a firm deviates from M^* to $\{1, n\}$, it increases its market share by $\frac{1}{2}\sigma\{1, n\} + \frac{1}{2}(1 - \sigma\{1, n\}) = \frac{1}{2} > c\{1, n\} - c(M^*\}$, hence the deviation is profitable. Therefore, $\{1\} \in S(\sigma)$. In particular, this means that firms earn a competitive payoff, because the menu $\{1\}$ generates a market share of $\frac{1}{2}$.

Next, let us show that $\{n\} \in S(\sigma)$. The menu M^* defined in the previous paragraph does not beat any other menu in $S(\sigma)$, because $b(M) \succeq b(M^*)$ for all $M \in S(\sigma)$. Note that M^* must be a singleton, because an irrelevant alternative would be costly to add yet it would generate no added market share. If $M^* \neq \{n\}$, then a firm can profitably deviate from M^* to $\{n\}$, thereby saving menu costs without changing its market share.

The key step in the proof is to show that $S(\sigma)$ contains exactly one menu of the form $\{x, n\}$, with $x = 1$. Assume that $\{x, n\} \in S(\sigma)$ for some $x \neq 1, n$. From the firms' decision to include n in this menu as part of a best-reply to σ, we conclude that $\frac{1}{2}\sum_{x \succ y} \beta_\sigma(y) - c_n \geq 0$. But since by assumption $\beta_\sigma(x) > 0$, this means that $\frac{1}{2}\sum_{1 \succ y} \beta_\sigma(y) - c_n > 0$. Thus, if a firm plays the menu $\{1, n\}$, it will generate a payoff strictly above $\frac{1}{2} - c_1$, contradicting a previous step. Now suppose that $\{1, n\} \notin S(\sigma)$. We have already established that $\{n\} \in S(\sigma)$. But if $S(\sigma)$ contains no menu of the form $\{x, n\}$ with $x \neq n$, the menu $\{n\}$ generates a payoff of $\frac{1}{2} - c_n > \frac{1}{2} - c_1$, a contradiction.

Thus, we have demonstrated that $S(\sigma) = \{\{1\}, \{1, n\}, \{n\}\}$. The play probabilities follow immediately from the requirement that all three menus generate a payoff of $\frac{1}{2} - c_1$. ■

The proof of Proposition 6 is preceded by a pair of lemmas concerning symmetric equilibria under reflexive consideration relations.

Lemma 2 *Suppose that R is reflexive. Then, $\beta_\sigma(x) \leq 2c_x$ for all $x \neq 1$.*

Proof. Assume the contrary. Let x be the $\succ$-minimal product for which $\frac{1}{2}\beta_\sigma(x) > c_x$. Suppose that there exists a menu $M \in S(\sigma)$ such that $b(M) \succ x$ and $y \not{R} x$ for all $y \in M$. Then, M does not beat any menu M' with $b(M') = x$. If a firm deviates from M to $M \cup \{x\}$, then since $b(M) \succ x$, the probability that some menu M'' with $b(M'') \succ b(M)$ beats M does not change. Therefore, by reflexivity of R, the deviation increases the firm's payoff by at least $\frac{1}{2}\beta_\sigma(x) - c_x > 0$, hence it is profitable. It follows that for every $M \in S(\sigma)$ for which $b(M) \succ x$, there exists some $y \in M$ such that yRx, so that M beats any M' with $b(M') = x$.

Now consider a menu $M \in S(\sigma)$ with $b(M) = x$ (there must be such a menu, since by assumption, $\frac{1}{2}\beta_\sigma(x) > c_x > 0$), and suppose that a firm deviates to $M \cup \{1\}$. The cost of this deviation is c_1, whereas the gained market share is at least $\frac{1}{2}\sum_{y \succeq x} \beta(y)$. The reason is that first, $M \cup \{1\}$ beats any menu M' with $b(M') = x$; and second, whereas prior to the deviation every menu $M' \in S(\sigma)$ with $b(M') \succ x$ had beaten M (as we showed in the previous paragraph), after the deviation no menu beats $M \cup \{1\}$. In order for this deviation to be unprofitable, we must have $\frac{1}{2}\sum_{y \succeq x} \beta(y) \leq c_1$. By the definition of x, $\frac{1}{2}\beta_\sigma(z) \leq c_z$ for all $z \prec x$. Adding up these inequalities, we obtain $\frac{1}{2}\sum_{y \in X} \beta(y) \leq c_1 + c(\{x+1, ..., n\}) < c(X)$ But since the L.H.S of this inequality is by definition $\frac{1}{2}$, we obtain $\frac{1}{2} - c(X) \leq 0$, contradicting condition (ii). ■

Lemma 3 *Suppose that R is reflexive. Then, for every $M \in S(\sigma)$ with $b(M) \neq 1$ there exists $M' \in S(\sigma)$ with $b(M') = 1$ such that M' does not beat M.*

Proof. Assume the contrary and let $M \in S(\sigma)$ be a menu which is beaten by all $M' \in S(\sigma)$ with $b(M') = 1$. If a firm deviates from M to $M \cup \{1\}$, it increases its market share by more than $\frac{1}{2}\beta_\sigma(1)$. In order for this deviation to be unprofitable, we must have $\beta_\sigma(1) \leq 2c_1$. Combined with Lemma 2, we obtain $\sum_x \beta_\sigma(x) \leq 2c(X)$. Since the L.H.S is equal to one, we obtain a contradiction. ■

Proof of Proposition 6
Let us first introduce three pieces of notation. first, recall the definition $\mathcal{M} = \{M \in S(\sigma) \mid b(M) = 1\}$. Second, for every $M \in \mathcal{M}$, define $B(M)$ as the set of products $z \neq 1$ for which $\beta_\sigma(z) > 0$ and yRz for some $y \in M \setminus \{1\}$. Second, define $\Delta(M) \equiv \sum_{z \in B(M)} \beta_\sigma(z) - 2\sum_{y \in M} c_y$.

Assume that firms earn a payoff above $\frac{1}{2} - c_1$ under σ. By Lemma 1, $\beta_\sigma(1) > 0$. By Lemma 3, $\beta_\sigma(x) = 0$ for all $1Rx$, $x \neq 1$. Therefore, in order for a menu $M \in \mathcal{M}$ to generate a payoff above $\frac{1}{2} - c_1$, it must be the case that $\Delta(M) > 0$. Suppose that $B(M) \cap B(M') = \varnothing$ for some pair of menus $M, M' \in \mathcal{M}$. If a firm deviates from M' to $M' \cup M$, it increases its payoff by $\Delta(M) > 0$, hence the deviation is profitable. It follows that $B(M) \cap B(M') \neq \varnothing$ for any pair of menus $M, M' \in \mathcal{M}$. When R is an equivalence relation, these pairwise intersections imply that $B(M)$ is the same for all $M \in \mathcal{M}$, contradicting Lemma 3. When R is a linear similarity relation, these pairwise intersections imply that $\cap_{M \in \mathcal{M}} B(M) \neq \varnothing$, again contradicting Lemma 3. ■

Proof of Proposition 7
Because identity is an equivalence relation, Proposition 6 implies that firms earn competitive payoffs in symmetric Nash equilibrium. Observe that under the identity consideration relation, M beats M' if and only if $b(M) \succ b(M')$ and $b(M') \in M$. Suppose that $\alpha_\sigma(x) = 0$ for some $x \neq 1$. Then, if a firm plays $\{x\}$, it earns $\frac{1}{2} - c_x$, which is above the competitive level, a contradiction. Therefore, $\alpha_\sigma(x) > 0$ for all $x \neq 1$. Let $M \in S(\sigma)$ be a menu that includes some $x \neq 1$ as a non-maximal product. Since the identity relation is reflexive, Lemma 2 implies $\beta_\sigma(x) \leq 2c_x$. If the inequality is strict, it is profitable for a firm to deviate from M to $M \backslash \{x\}$. It follows that $\beta_\sigma(x) = 2c_x$. But this means that any menu $M \in S(\sigma)$ with $b(M) = x$, $x \neq 1$, yields the same payoff against σ as the singleton $\{x\}$. Therefore, $\frac{1}{2}[1 - \alpha_\sigma(x)] - c_x = \frac{1}{2} - c_1$, which implies $\alpha_\sigma(x) = 2c_1 - 2c_x$. ■

The two results that characterize symmetric equilibria when menu costs are small rely on the following lemma.

Lemma 4 $\sum_{x \neq 1} \beta_\sigma(x) \leq 2c(X)$

Proof. By Lemma 1, $S(\sigma)$ contains at least one menu M with $b(M) = 1$. List all these menus as follows: $M^1, ..., M^L$. If one of these menus M^l does not beat any other menu in $S(\sigma)$, then it must be the case that $\sum_{x\neq 1} \beta_\sigma(x) \leq 2(c(X) - c(M^l))$ - otherwise, it is profitable to deviate from M^l to X - and the result follows immediately. Now suppose that each of these menus M^l beats some other menu in $S(\sigma)$. Then, for any $l \in \{1, ..., L\}$, it must be the case that $\sum_{x\neq 1, yRx \text{ for no } y \in M^l} \beta_\sigma(x) \leq 2(c(X) - c(M^l))$ - otherwise, it is profitable to deviate from M^l to X. In addition, it must be the case that $\sigma(M^l) + \sum_{x\neq 1, yRx \text{ for some } y\in M^l} \beta_\sigma(x) \leq 2c(M^l)$ - otherwise, it is profitable to deviate from one of the menus that M^l beats into M^l itself. Summing over these inequalities, we obtain $\sigma(M^l) + \sum_{x\neq 1} \beta_\sigma(x) \leq 2c(X)(c(X) - c(M^l))$, which immediately implies the result. ■

Proof of Proposition 8

Assume $c_1 < 1/(2^n + 2n)$. Suppose that $S(\sigma)$ contains a menu M such that $1Rb(M)$. Then, $\frac{1}{2}\beta_\sigma(1) \leq c_1$ - otherwise, it is profitable for a firm to deviate from M to $\{1\}$. Combined with Lemma 4, we obtain the inequality $\sum_{x=1,...,n} \beta_\sigma(x) \leq 2(c(X) + c_1) < 2(n + 1)c_1$. Note that the L.H.S of this inequality is by definition equal to one. Therefore, we obtain the inequality $c_1 > 1/2(n + 1)$, which contradicts our assumption on c_1 for any $n \geq 2$.

Now suppose that $\beta_\sigma(x) = 0$ for all $1Rx$, $x \neq 1$. Note that $S(\sigma)$ contains at most 2^{n-1} menus M with $b(M) = 1$. If $\{1\} \in S(\sigma)$, then this menu generates a payoff of $\frac{1}{2} - c_1$ against σ, which concludes the proof. Thus, assume $\{1\} \notin S(\sigma)$. Then, each $M \in S(\sigma)$ with $b(M) = 1$ must beat some other $M' \in S(\sigma)$. Moreover, it must be the case that $\sigma(M) \leq 2c_1$- otherwise, it is profitable to deviate from M' to $\{1\}$. Summing over all these menus M, we obtain $\beta_\sigma(1) \leq 2^n \cdot c_1$. Combined with Lemma 4, we obtain $1 \leq 2c(X) + 2^n \cdot c_1 < (2^n + 2n)c_1$, a contradiction. ■

Proof of Proposition 9

This is an immediate corollary of Lemma 4. ■