A "Sheening" Theory of Product Attribute Advertising

Simon P. Anderson,[*] Federico Ciliberto,[†] and Jura Liaukonyte.[‡]

PRELIMINARY WORK. PLEASE DO NOT CITE

Abstract

If advertising improves the allure of promoted product attributes, then, in a competitive setting, advertising of these attributes affects the structure of demand as a function of shared characteristics. We develop a theory model that predicts which brands will choose to advertise which attributes and to what extent. We show that optimally, firms will specialize in what characteristic they advertise the most and that advertising will exaggerate the perceptual differences in characteristics. We test our model with the OTC analgesics category data, which includes detailed expenditure information about advertised attributes.

Keywords: Product differentiation, advertising, attribute valuation..

JEL Classification: L13, M37, L15

[*]Department of Economics, University of Virginia, Charlottesville VA 22904, USA. sa9w@virginia.edu

[†]Department of Economics, University of Virginia, Charlottesville VA 22904, USA. ciliberto@virginia.edu

[‡]Dyson School of Applied Economics and Management, Cornell University, Ithaca NY 14850, USA. Jurate@cornell.edu

1 Introduction

We reconsider the economics of product differentiation by looking at how advertising of characteristics in a competitive setting affects the structure of demand as a function of shared characteristics of products. The theory determines which brands advertise which characteristics of their products, and to what extent.

The starting point is the idea that advertisements emphasize certain properties of products, and that brands choose which attributes to emphasize. Such advertising entails spillovers between brands. While this effect is often mentioned in casual discussion of advertising, it has not been formally modeled.

We suppose that products can be described by vectors of characteristics that they embody, and that advertising improves the allure of particular characteristics.[1] Then advertising has positive and negative spillovers. It has positive benefits to the extent that other brands with similar characteristics to those being advertised will see an increase in demand. An ad that makes consumers thirst for low-calorie beer may help all low calorie ales, and not just the one paying for the advertisement. But an ad will tend to steal business from products in the same broad class that are less similar, i.e., have fewer common characteristics. Advertising will also alter equilibrium prices: this must also be factored into the advertising choice.

To say more about the relations between brands and advertising, we need to put more structure on the profit function. To illustrate the analysis, we first describe the equilibrium configuration in terms of a logit approach. This has the advantage that it is readily related to much work on differentiated product demand systems in the structural empirical literature. We then extend to a more general demand system (still rooted in the discrete choice approach to product differentiation) where we emphasize the substitutability patterns between pairs of products through their diversion ratios.

Apart from some recent papers on disclosure games, described below, the economics literature has scarcely addressed the choice of ad content (Bagwell, 2007, delivers an excellent survey on the Economics of Advertising). The literature on informative advertising (see Shapiro 1980 for the monopoly case and Butters, 1977, for a competitive analysis) has been

[1] One benefit to emphasizing the characteristics embodied in goods is to revisit the characteristics models associated with Lancaster (1966) from a different angle.

mostly concerned with advertising "reach" (the number of consumers seeing the ad), and whether or not this is socially insufficient. Because it is typically assumed that products are homogeneous, the ad needs to communicate is the product price and where the consumer can buy it.[2]

Konishi and Sandfort (2002) describe the role of price assurance in ads when consumers face search costs in buying. Bar-Isaac et al. (2010) and Kuksov and Villas-Boas (2010) allow the consumer to buy in ignorance and save the cost of learning about attributes. The persuasion game literature (Milgrom, 1981 and Grossman, 1981) shows that all brand types reveal their true quality (quality unraveling) or else they are taken as the worst possible quality. Information is also conveyed by quality signaling. Signaling models of advertising allow consumer inference of high product quality from observing high enough advertising spending, although the ad need not convey any direct information about the product (see Nelson 1970, 1974, Kihlstrom and Riordan 1984, and Milgrom and Roberts 1986b).

The recent literature on disclosure games has mainly focussed on monopoly. Sun (2010) analyzes horizontal product information quality disclosure. Mayzlin and Shin (2011) analyze two quality attributes and a limited communication technology. Koessler and Renault (2011) analyze the general monopoly case. Guo and Zhao (2009) address duopolists' incentives to reveal quality information, assuming that each is ignorant of the other's quality; Board (2009) does similarly assuming that they know each other's quality.

Anderson and Renault (2006) look at a monopoly firm's problem of how much horizontal information to convey to consumers. Anderson and Renault (2012) extend the analysis to include quality information. By contrast, in the current paper, instead of being informative, the advertising improves the allure of the characteristic to the consumer.

2 An exploratory model: Logit approach

To fix ideas, assume that product choice is governed by a logit demand structure of the type analyzed in Anderson and de Palma (2001) and in many econometric studies of demand for

[2] See Grossman and Shapiro (1984), Meurer and Stahl (1994), and Christou and Vettas (2008) for analysis of differentiated products.

differentiated products. The demand for product i is given by

$$D_i = \frac{\exp\left[q_i - \alpha p_i\right]}{\sum_{j=1}^{n} \exp\left[q_j - \alpha p_j\right] + \exp V_0}, \qquad i = 1, ..., n,$$

where α represents the marginal utility of income,[3] q_j represents product j's quality, p_j its price, and V_0 represents the value of the outside option of not buying. A key property for the analysis below is that $\frac{\partial D_i}{\partial p_i} = \alpha D_i \left(D_i - 1\right) < 0$.

Now, to highlight the characteristics content of products, we suppose that

$$q_j = \sum_{k=1}^{K} \tilde{\beta}^k X_j^k \qquad j = 1, ..., n \tag{1}$$

where $\tilde{\beta}^k$ represents the consumer valuation per unit of characteristic k, and X_j^k is the amount of characteristic k embodied in good j, which is assumed to be fixed. Think of chemical molecules in analgesics, for example. Assume finally that the valuation of the characteristic is an increasing function of the total amount of advertising of the characteristic. Think of total minutes spent on television for extolling the characteristic, for example. Then we can write

$$\tilde{\beta}^k = \beta^k \Phi_k \left(A^k\right) \tag{2}$$

where $\Phi_k(.)$ is increasing and strictly concave in the amount A^k spent on promoting characteristic k. Set too $\Phi_k(0) = 1$ so that β^k represents the base level of characteristic k's valuation. Because only the total spending by all brands matters, then

$$A^k = \sum_{j=1}^{n} a_j^k, \quad k = 1, ..., K,$$

where a_j^k is the amount of advertising by brand j of characteristic k.

Consider a Nash equilibrium to the game in which brands choose their advertising levels and prices. Assume constant marginal production costs, c_i. Brand i's profit is

$$\pi_i (A, p) = (p_i - c_i) D_i (A, p) - \sum_{k=1}^{K} \gamma^k a_i^k, \quad i = 1, ..., n. \tag{3}$$

[3] Here α is effectively $\frac{1}{\mu}$ in the version of the logit oligopoly model developed by Anderson, de Palma, and Thisse (1992), where μ measures the degree of differentiation between products.

where γ^k denotes the advertising cost for characteristic k, p denotes the vector of prices, and A denotes the vector of total advertising levels of characteristics. Brand i's demand, $D_i(A, p)$, is constructed from the steps above. Brand i is to maximize (3) by choice of its price, p_i, and the vector of advertising levels, $\left(a_i^1, ..., a_i^K\right)$.

The sequel describes the equilibrium.

2.1 One brand per characteristic

The Nash equilibrium to the advertising level problem described above has an important property. Namely, each characteristic will be advertised by only one brand (at most). The intuition is as follows. Consider a simple public goods contribution game with quasi-linear preferences - as expounded in Mas-Collel, Whinston and Green (1995), or Varian (1992). Then, absent identical preferences, the equilibrium has the individual whose demand is greatest at marginal cost doing all the contributing. Any other individual then has a demand price below marginal cost and so would not wish to contribute additional funds, whereas the individual with the highest demand would always want to "top up" any lower contribution level to his/her preferred level. The same logic applies when there are multiple public goods to which individuals may contribute. Absent ties, only one individual will contribute to each public good.

The logit structure above is specific, but it was boiled down to the relation (4). This means that any structure that boils down in this manner has the same equilibrium property of one brand per characteristic. Indeed, the above idea extends to games in which advertising levels and then prices are determined in a two-stage game (rather than simultaneously). In the second stage, prices are determined as a Bertrand-Nash equilibrium given the advertising levels and the inherent characteristics of products.

Then, for any value of A, in the price sub-game each brand will choose its price to maximize $(p_i - c_i) D_i(A, p)$, $i = 1, ..., n$. Denote the sub-game equilibrium values with a star. Then $\pi_i^*(A) = (p_i^* - c_i) D_i(A, p^*)$ depends only on the vector of total advertising levels of characteristics, A, and the matrix of characteristics contents of products. Thus, since we hold the latter fixed, the first stage objective function involves each brand i choosing a vector of advertising levels, $\left(a_i^1, ..., a_i^K\right)$ in order to maximize

$$\pi_i^*(A) - \sum_{k=1}^{K} \gamma^k a_i^k, \qquad i = 1, ..., n, \tag{4}$$

where γ^k denotes the advertising cost for characteristic k. Again, any game that boils down to such a form (and logit is not instrumental here) has the one-brand-per-characteristic property.

2.2 Some Examples

We give some preliminaries and then some illustrations. The first-order condition for pricing yields

$$(p_i - c_i) D_i (D_i - 1) \alpha + D_i = 0, \qquad i = 1, ..., n. \tag{5}$$

Hence the equilibrium prices for the brands are implicitly (see Anderson and de Palma, 2001):

$$p_i - c_i = \frac{1}{\alpha (1 - D_i)} \qquad i = 1, ..., n.$$

These indicate that brands with higher mark-ups are associated in equilibrium with higher demands, which in turn means that such brands must have higher net qualities $(q_i - \alpha p_i)$ in order for the demands to be higher. Such brands also garner higher gross profits (gross of advertising costs).[4] As we shall see, this profitability property is no longer true once we allow for endogenous advertising. Indeed, if one brand is has a slight advantage over another for a characteristic, but they are otherwise symmetric, that "superior" brand will end up doing all the advertising of the characteristic, and the accompanying expense will make its net profits lower than those of its rival. The rival free rides on the advertising expense.

2.2.1 Raise-all-boats advertising

Suppose that all brands have the first characteristic in common, and are otherwise symmetric.[5] In this case advertising of that characteristic is of the "Raise-all-boats" form insofar

[4]This also means that these are the firms with the highest net social contributions $(q_i - \alpha c_i)$: see Anderson and de Palma (2001) for more details.

[5]For example, they might have all the other characteristics in common, or they might have each an idiosyncratic characteristic of common value, etc.

as it brings in consumers to all brands and away from the outside option. Such advertising entails a positive externality, and so is underprovided by the brands in a classic free-rider problem. Consider any brand, say Brand A. the first one.

Its first-order condition for choice of ad level for characteristic 1, assuming it is the marginal advertiser, is

$$(p_A - c)\,\beta\Phi'\left(a^1\right) D_A D_0 = \gamma_A$$

where c is the common marginal cost and $\Phi(.)$ is the common transformation function. The intuition here is that an increase in a^1 gains the brand the value of its mark-up on all the consumers diverted to it from the outside option, 0. However, because all other brands are equally sheened, no consumers are diverted from them to A. The demand product $D_A D_0$ in the advertising derivative is just the logit model expression for diverted demand (and all other brands gain the same amount from A's advertising): the IIA property of the logit means that lost customers are picked up in proportion to market share. Using the pricing condition (5) yields

$$\frac{\beta}{\alpha}\Phi'\left(a^1\right)\frac{D_A}{1-D_A}\left(1-nD_A\right) = \gamma_A$$

The interpretation of this is as follows. Suppose that A increased by one unit its advertising and hence the allure for characteristic 1. It could then raise its price by $\beta\Phi'\left(a^1\right)/\alpha$ and lose no consumers. It would gain the extra revenue on its consumer base D_A. However, it also raises the allure for all its rivals (except the outside good)...

[how do we see the free-rider effect; how does it attenuate with n?]

2.2.2 Zero-sum advertising

At the other extreme, the model readily admits as another special case the idea of zero-sum advertising. To see this, suppose that $V_0 \to -\infty$ so advertising simply reshuffles demand among brands, and that each brand provides an idiosyncratic characteristic. Then, applying symmetry, the first order condition for any brand (using again the first-order pricing equation) boils down to

$$\alpha\beta\Phi'(a)\,D = \gamma, \tag{6}$$

where a is the common advertising level and D the common demand level. The interpretation is that a one-unit rise in advertising the characteristic raises the perceived quality by $\beta\Phi'(a)$ and so the brand can raise its price by $\beta\Phi'(a)/\alpha$ and leave its demand unchanged. It therefore earns this "neutralizing" price increase on its customer base, D, and so the LHS of (6) measures the dollar benefit of an extra unit of advertising, while the RHS is the marginal cost.

With no outside good, then simply $D = 1/n$ where n is the number of brands. This implies that the TOTAL equilibrium advertising level is

$$a^* = \Phi'^{-1}\left(\frac{\gamma n \alpha}{\beta}\right)$$

and the mark-up is $n/\alpha(n-1)$. This leads to an equilibrium level of profit per brand of

$$\frac{1}{\alpha(n-1)} - \frac{\gamma}{n}a^*$$

(where we assumed that the advertising is shared equally among firms). Note that when advertising costs, γ, rise, then profits rise if and only if Φ'^{-1} is elastic. This is because higher costs cause a more than proportionate reduction in advertising per brand.

More competition (higher n) has two competing effects on profit. First, it reduces the gross profit, but then it also decreases both the advertrising share of each brand and the total advertising level (via the free-rider effect that the advertising effectivenenss is now split more ways). *[is there a clean condition to give?]*

2.2.3 Asymmetries

Suppose for illustration that there are three brands and an outside option. Brands A and B have a common characteristic and so are similar in that respect. Firm C has an idiosyncratic characteristic.

Suppose for clarity that both the advertising cost, γ^k, and the valuation β^k is the same for all characteristics, and A and B share the first characteristic, while C is the only brand offering characteristic 2. Assume that each product has one unit of the relevant characteristic. We can now address the choice of advertising levels. From the result above, we cannot determine which of A and B advertises the first characteristic, or indeed whether they share that advertising.[6] However, the level of advertising is tied down by the first order condition. Considering Brand A's choice of advertising of characteristic 1, a^1 is determined by

$$(p_A - c_A)\,\beta\Phi'\left(a^1\right) D_A\left(1 - D_A - D_B\right) = \gamma,$$

while Brand C's first-order condition for characteristic 2 implies

$$(p_C - c_C)\,\beta\Phi'\left(a^2\right) D_C\left(1 - D_C\right) = \gamma.$$

The difference in the two expressions reflects the idea that advertising characteristic 1 has a detrimental effect on brand A through improving brand B's attractiveness, and so is expected to reduce the incentive to advertise characteristic 1. We now confirm this intuition, though the effect is more involved than might be anticipated from the simple argument. Substituting in the first order pricing conditions simplifies the above equations to

$$\Phi'\left(a^1\right) \frac{D_A}{1 - D_A}\left(1 - D_A - D_B\right) = \frac{\alpha\gamma}{\beta},$$

and

$$\Phi'\left(a^2\right) D_C = \frac{\alpha\gamma}{\beta}.$$

In ratio form, we have

$$\frac{D_A}{1 - D_A}\frac{\left(1 - D_A - D_B\right)}{D_C} = \frac{\Phi'\left(a^2\right)}{\Phi'\left(a^1\right)}. \tag{7}$$

Suppose first that the outside good is never an option (i.e., $V_0 \to -\infty$). Then, $1 - D_A - D_B = D_C$ and the ratio form (7) is $\frac{D_A}{1-D_A} = \frac{\Phi'\left(a^2\right)}{\Phi'\left(a^1\right)}$. Since D_A cannot exceed one half in equilibrium (because in equilibrium $D_A = D_B$ and $D_C > 0$), and $\Phi'(.)$ is a decreasing function, then we necessarily have $a^2 > a^1$. Even if V_0 is finite, the same result holds because the solution is a continuous function of V_0, and $a^2 = a^1$ can never solve the ratio form. This means the solution always lies on the same side. Hence, regardless of the value of V_0, we always have the result that $a^2 > a^1$. Loosely, more "competitive" characteristics (in the sense of ones that are shared by more brands) will be advertised less.

[6] This is simply an artefact of the assumption here that both A and B have the same amount of characteristic 1. Generically, the amounts will differ in practice, so that this knife-edge case will not occur. As seen below, it will suffice that A (say) has marginally more of the characteristic, so that it alone will advertise 1.

3 Equilibrium advertising spending

The "public goods contribution game" result above gives us a corner solution for all but one brand for each characteristic. The equilibrium involves a public good problem, so that a sheening ad raises all similar "boats" and suppresses dissimilar ones. We now expand these themes. Assuming (as before) a simultaneous price-advertising game, the pricing first-order condition is $(p_i - c_i) D_i (D_i - 1) \alpha + D_i = 0, \quad i = 1, ..., n$ (see (5)). The advertising first-order condition for brand i for characteristic k, conditional on i choosing a positive ad level for k, is:

$$(p_i - c_i) D_i \left[(1 - D_i) X_i^k - \Sigma_{j \neq i} D_j X_j^k \right] \beta^k \Phi_k' \left(A^k \right) = \gamma^k, \qquad i = 1, ..., n; \ k = 1, ..., K. \tag{8}$$

Substituting the pricing first-order condition (5) into the advertising one (8), we can write the first-order condition as

$$W_i^k D_i \Phi_k' \left(A^k \right) = \alpha \gamma^k, \qquad i = 1, ..., n; \qquad k = 1, ..., K; \tag{9}$$

where we have defined

$$W_i^k = \left[X_i^k - \Sigma_{j \neq i} \frac{D_j}{(1 - D_i)} X_j^k \right] \beta^k \tag{10}$$

as the index of $i's$ *diversity advantage* in characteristic k. This is a key summary statistic. It is the value of i's characteristic advantage in characteristic k. First, it indicates amounts of advertising of different characteristics. For example, if the cost of advertising two characteristics are the same for a brand advertising both of them (as above) then there will be more advertising of the one for which W_i^k is greater.

Proposition 1 *Assume a brand advertises at least two characteristics, and the cost of advertising each is the same, as is the sheening function $\Phi(.)$. Then the ranking of ad intensities follows the ranking of the W_i^k. That is, $A_i^k > A_i^m$ as $W_i^k > W_i^m$.*

Proof. Assuming the same sheening function for both characteristics, we can write (**??**) as

$$W_i^k \Phi' \left(A^k \right) = W_i^m \Phi' \left(A^m \right).$$

Suppose that $W_i^k > W_i^m$. Then, because $\Phi'(.)$ is a decreasing function, we must have $A^k > A^m$. ∎

This says that the "better" characteristic (as defined by the ranking of the W_i's) will be advertised more, ceteris paribus. These are the ones with more inherent ex-ante value to consumers, i.e., bigger β's. Likewise, advertising is larger the bigger is own strength in the characteristic, and the smaller is rival strength.

The diversity advantage index also tells us which brand will advertise which characteristic.

Proposition 2 *Rank characteristics for brands according to $W_i^k D_i$. Then the brand i with the highest value of $W_i^k D_i$ will be the one that will do all the advertising of characteristic k.*

Proof. From (9), the marginal benefit is highest for the brand with the greatest $W_i^k D_i$. The result then follows from the public good argument. ■

Define now $\tilde{W}_i^k = W_i^k \Phi\left(A^k\right)$ as the *post-advertising sheen* of characteristic k. The next Proposition tells us how characteristic differences are exaggerated through equilibrium sheening.

Proposition 3 *(Magnification) Assume a brand advertises at least two characteristics (k and m), for which the cost of advertising each is the same, as is the sheening function $\Phi(.)$. Equilibrium sheening will exaggerate the perceptual differences in characteristics; namely, $\tilde{W}_i^k - \tilde{W}_i^m > W_i^k - W_i^m$ for $W_i^k > W_i^m$.*

Proof. Suppose, w.l.o.g., that $W_i^k > W_i^m$. The post-advertising sheen of characteristic k is $\tilde{W}_i^k = W_i^k \Phi\left(A^k\right)$, and k is advertised more so $\Phi\left(A^k\right) > \Phi\left(A^m\right)$, and the result follows directly. ■

This result says that perceived differences in valuations are magnified through equilibrium advertising (equivalently, $\frac{\tilde{\beta}^k}{\beta^k} > \frac{\tilde{\beta}^m}{\beta^m}$, so characteristic k gets more sheened.) *[variant of above for different brands: use (9), though see next Propo below]*

We now further interpret the diversity index, W_i^k. It is the product of two terms, the base characteristic value, β^k and the term $X_i^k - \Sigma_{j\neq i}\frac{D_j}{(1-D_i)}X_j^k$. The latter term is the difference between i's characteristic level and the demand weighted (standardized by the non-i demand) characteristics of others. To help interpret the term, note that if all other products have the same demand, call it $\bar{D} = D_j$ for all $j \neq i$, then the weights are $\frac{\bar{D}}{(n-1)\bar{D}+D_0}$: if the market were fully covered ($D_0 = 0$), then the weights are simply $\frac{1}{n-1}$. So then W_i^k is just β^k times the difference over the average value of the others (pre-sheening "hedonic" value of superiority over the average other value).

Finally, there is a clean condition for determining advertising levels across different brands.

Proposition 4 *(Cross-brand) Assume two characteristics (k and m) have the same cost of advertising and the same sheening function* $\Phi(.)$. *Then* $A^k > A^m$ *as* $W_i^k D_i > W_l^m D_l$.

Proof. The first-order conditions yield

$$W_i^k D_i \Phi'\left(A^k\right) = W_l^m D_l \Phi'\left(A^m\right)$$

Hence, because $\Phi'(.)$ is a decreasing function, $A^k > A^m$ as $W_i^k D_i > W_l^m D_l$. ■

[interpretiation and follow-on]

The simple examples above indicate how the basic logit form can be used to describe patterns in the structure of product differentiation. The third example above was quite straightforward insofar as A's ads hurt C and help B. More generally, we should like to determine the conditions under which brands' advertising helps or hurts which rival products. As seen in the next section, the outcome depends on the characteristics overlap and the diversion ratios between brands. The focus on exclusive and complementary characteristics underscores a novel approach to product differentiation where advertising plays a key role.

4 General formulation

Here we reconsider the above results using a general diversion ratio approach.

[this seems quite tractable; maybe we don't need logit after all. To be seen. And, there are more results to get!] [other properties that we can look at? suggestions pls.]

Suppose that the demand is generated by a discrete choice model, where consumer (conditional indirect) utility is

$$u_i = q_i - \alpha p_i + \varepsilon_i$$

where q_i is given by (1) as $q_j = \sum_{k=1}^{K} \tilde{\beta}^k X_j^k$, $j = 1, ..., n$, and $\tilde{\beta}^k$ is given by (2) as $\tilde{\beta}^k = \beta^k \Phi_k\left(A^k\right)$, $k = 1, ..., K$. The logit is the special case where the random match terms, ε_i, are distributed Type 1 Extreme Value.

Now, let the market size be M and the share of brand j be denoted s_j, $j = 1, ..., n$, with "outside option" 0 with share s_0 so that the shares sum to one. The key here is that the shares depend only on the vector of net qualities, $q_i - \alpha p_i$, so that the derivative with respect to a price is minus-α times the derivative with respect to the same brand's quality. Shares are increasing in own net quality and decreasing in each rival's net quality, so that brands are substitutes. Brand i is to maximize

$$\pi_i = (p_i - c_i)\, M s_i - \sum_{k=1}^{K} \gamma^k a_i^k \qquad i = 1, ..., n,$$

by choice of $\left(p_i, a_i^1, ..., a_i^K\right)$. The equilibrium pricing condition is

$$(p_i - c_i) = \frac{s_i}{\alpha ds_i/dq_i} \qquad i = 1, ..., n, \tag{11}$$

while if i is the only brand advertising characteristic k its profit derivative is

$$\begin{aligned} \frac{d\pi_i}{da_i^k} &= (p_i - c_i)\, M \sum_{j=1}^{n} \frac{ds_i}{dq_j} \frac{dq_j}{da_i^k} - \gamma^k \qquad (12) \\ &= M s_i \left(\frac{dq_i}{da_i^k} - \sum_{j \neq i}^{n} d_{ji} \frac{dq_j}{da_i^k} \right) - \gamma^k \\ &= M s_i \beta^k \Phi_k' \left(A^k\right) \left(X_i^k - \sum_{j \neq i}^{n} d_{ji} X_j^k \right) - \gamma^k \end{aligned}$$

where we have made two key substitutions on subsequent lines. The first is to deploy the diversion ratio, $d_{ji} = -\frac{ds_i}{dq_j} / \frac{ds_i}{dq_i} \in (0, 1)$. Since equivalently $d_{ji} = -\frac{ds_j}{dp_i} / \frac{ds_i}{dp_i}$, this tells us the fraction of i's customers who are lost from a price hike are picked up by j.[7] Hence d_{ji} is the diversion from i to j, and the sum of such diversion ratios is one if we include the outside option, since all lost consumers go somewhere (i.e., $\sum_{j=0,..n;\ j \neq i} d_{ji} = 1$). The second substitution is from how ad levels impact perceived quality. Define $W_i^k = \left(X_i^k - \sum_{j \neq i}^{n} d_{ji} X_j^k \right)$ as the

[7] A key property here is $\frac{ds_i}{dq_j} = \frac{ds_j}{dq_i}$: this is a property of discrete choice models (see Anderson, de Palma, and Thisse, 1992), as well as representative consumer models (where it corresponds to the symmetry of the Slutsky matrix).

Proposition 5 *In equilibrium, each characteristic is advertised by just one brand, which brand is the one for which $s_i W_i^k$ is greatest.*

Proof. From the "public good" argument, and that the profit benefit is greatest for the brand with the largest $s_i W_i^k$. QED. ■

We now give some intuition and implications of this result. First, W_i^k measures the characteristic advantage of i for sheening characteristic k, duly subtracting the advantage conferred on rivals via the diversion ratios that indicate how much such advantage elsewhere detracts from the sheener. This net characteristic advantage is weighted by the market share, s_i, of the sheener, because the benefit is applied on all units sold in terms of a compensating price that i can charge and leave its output unchanged.

[should we engage the compensating price intuition? do it more carefully and explicitly.]

Thus *bigger brands are more likely to advertise a given characteristic*, ceteris paribus. That is, we expect more sheening advertising from bigger brands, even though they might not necessarily have the greatest characteristic advantage in characteristic k. The "strongest" brand does not necessarily do the sheening. However, if shares and diversion ratios are similar for different brands, the strongest brand is indicated as the advertiser. This latter effect is tempered by the diversion ratios. Ceteris paribus, the advertiser is the one with fewer customers diverted elsewhere. Notice that the characteristics X_i^k are not constrained to be positive. It may be that they are measured by reference to a benchmark, so that negative values represent characteristic levels falling short of a mean level. For example, the characteristic could be speed of action of a drug, with slower than some threshold being measured as a negative attribute. Of course, if a brand has a negative value for a characteristic, then sheening of the characteristic by a rival will hurt it. This harm is inflicted even with a small positive value if rivals have stronger values because of the demand diverted. We now turn to the benefits conferred or damages inflicted by sheening.

Consider then the benefit (or damage) to a rival of an incremental sheening advert by j. This is simply the previous expression (12) without the cost of the sheening.

$$\begin{aligned} \frac{d\pi_i}{da^k} &= (p_i - c_i) M \sum_{j=1}^{n} \frac{ds_i}{dq_j} \frac{dq_j}{da_i^k} \\ &= M s_i \beta^k \Phi_k' \left(A^k\right) W_i^k. \end{aligned}$$

Then, if all the X_i^k are positive, there is clearly insufficient advertising through the classic free-rider effect. However, there is harm to others whose characteristics are measured as negative, and also to those for which W_i^k is negative via the diversion effect.

The total marginal value of a sheening ad is therefore the sum over all non-sheening agents, because the value to the sheening brand is zero at the margin:

$$M\beta^k \Phi_k' \left(A^k\right) \sum_{i \neq j} s_i W_i^k$$

and so this has a negative value if and only if $\sum_{i \neq j} s_i W_i^k < 0$.

[can we do something with this, breaking out the divrats?]

[notice that the amount of advertising of a characteristic is given from $M s_i \beta^k \Phi_k' \left(A^k\right) \left(X_i^k - \sum_{j \neq i}^{n} d_{ji} X_j^k\right)$ γ^k. This might be estimated: a simple version might have Φ as a logarithmic function, so that then we would have the predicted amount of advertising of characteristic k as $A^k = \frac{M s_i \beta^k W_i^k}{\gamma^k}$.]

5 Empirical Evidence

We evaluate the theoretical model in light of findings for the US over-the-counter (OTC) analgesics industry. The OTC analgesics market covers pain-relief medication. The dominant active chemical ingredients are aspirin, acetaminophen, ibuprofen and naproxen sodium. The nationally advertised brands for each of these segments are familiar brand names: Tylenol (acetaminophen), Advil and Motrin (ibuprofen), Aleve (naproxen sodium), Bayer (aspirin or combination), Excedrin (acetaminophen or combination), and Midol and Pamprin (various formulae for menstrual pain relief). Since different brand pain relief mechanism is based on different active ingredients, these major brands vary in potential medical risks that they entail and in the efficiency of pain relief. They therefore embody varying base characteristics levels.

Our theory indicates that positive advertising of a characteristic is likely to "raise all boats" for the brands that are strong in that characteristic and will thus most benefit the brand that is perceived to best embody that characteristic. Thus, we predict that the product that best embodies that specific characteristic will advertise it.

5.1 Overview of the OTC Analgesics Industry and Data

The OTC analgesics industry is highly concentrated - the top 6 brands account for approximately 71% of the dollar market share. Generic store brands account for another 26%. The main brands (with national advertising), their market shares (calculated from our data), their ingredients, and ownership are given in Table 1.

As reported in Table 1, advertising-to-sales ratios for OTC analgesics typically range from 20% to 30% of sales, making them one of the most heavily promoted manufactured goods. Additionally, the type of cues mentioned (e.g., "strong") are clearly identifiable, which enables us to avoid making any subjective judgments while coding the information cue. All of these reasons make this industry ideal to test our theory model.

Table 1: OTC Analgesic Brands and Market Shares

Brand	Ownership	Active Ingredient	Average Price*	Market share	Advert. share	TA/ Sales
Advil	Wyeth	IB	3.75	16.49%	20.88%	23.54%
Aleve	Bayer	NS	3.77	7.91%	13.88%	32.60%
Bayer	Bayer	ASP	2.75	5.45%	9.22%	31.43%
Excedrin	Novartis	ACT	3.63	8.39%	12.75%	28.24%
Midol	Bayer	ACT/NS/IB	5.6	1.67%	2.29%	25.55%
Motrin	McNeil	IB	3.65	5.58%	8.71%	28.99%
Pamprin	Chattem	ACT/NS/IB	4.29	0.56%	0.62%	20.38%
Tylenol	McNeil	ACT	3.29	27.88%	31.65%	21.10%
Generic					0%	0%
		ACT	2.03	8.40%		
		IB	2.74	9.02%	26.06%	
		NS	2.83	1.61%		
		ASP	1.32	6.44%		

*average price of a representative 24 pill bottle of regular strength tablets
ACT-Acetaminophen; IB-Ibuprofen; NS-Naproxen Sodium; ASP-Aspirin
TA - Total Advertising

5.2 Advertising Data

The advertising data come from TNS Media Intelligence and cover the entire U.S. OTC analgesics product category. The data set contains video files of all advertisements, as well as monthly advertising expenditures, for each product advertised in the OTC analgesics category from 2001 to 2005. The advertising numbers also include expenditures on other media, but almost all the advertising budgets (approximately 90%) were spent on broadcast television advertising, including network and cable television. In our analysis, we examine only the television advertising data.

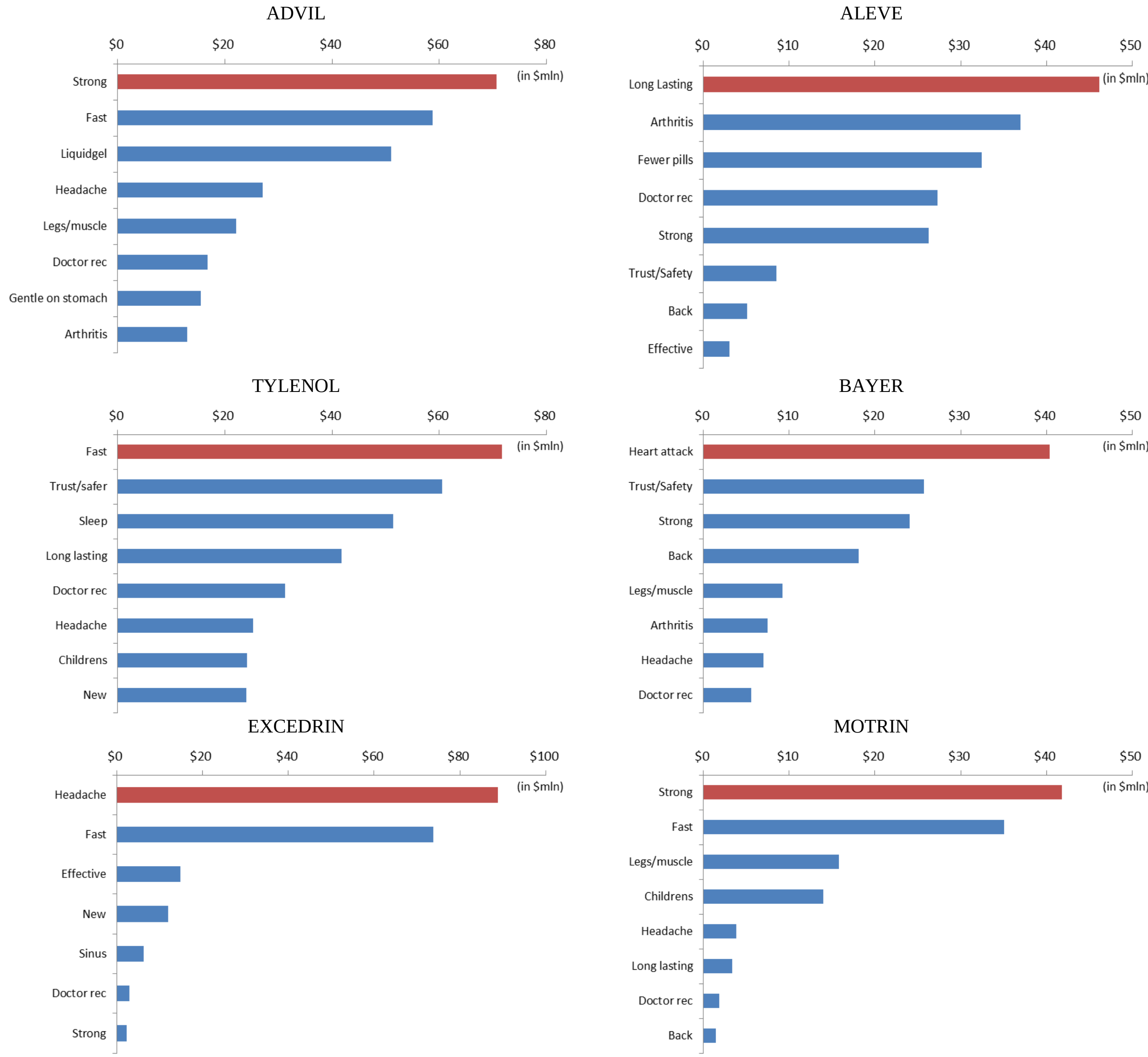
We watched 4503 individual commercials broadcast during the 2001–2005 period, 346 of which had missing video files. Each individual advertisement was usually shown multiple times. For each advertisement, we recorded all information cues mentioned, including the purpose of the drug (e.g., menstrual pain, arthritis, headache), drug efficiency (e.g., strength, speed), safety, and other characteristics. The type of information cues that were mentioned (e.g., "strong") are clearly identifiable, which enables us to avoid making any subjective judgments while coding the information content.

The breakdown of advertising expenditures by 6 major brands across advertised attributes is given in Figure 1 and Table 2 *[you can pick one - both illustrate the same thing].* Each top brand corresponds to a different active ingredient (Excedrin's active ingredient varies, but mostly it is acetaminophen combined with other active ingredients, and Motrin's active ingredient is Ibuprofen, identical to Advil's).

Table 2. Dominant Advertised Attribute by Brand

Attribute	**Advil**	**Attribute**	**Aleve**	**Attribute**	**Tylenol**
Strong	\$70.70	Long Lasting	\$46.18	Fast	\$71.73
Fast	\$58.81	Arthritis	\$36.93	Trust/safer	\$60.56
Liquidgel	\$50.96	Fewer pills	\$32.47	Sleep	\$51.39
Headache	\$27.11	Doctor rec	\$27.26	Long lasting	\$41.76
Legs/muscle	\$22.07	Strong	\$26.28	Doctor rec	\$31.18
Doctor rec	\$16.84	Trust/Safety	\$8.48	Headache	\$25.25
Gentle on stomach	\$15.55	Back	\$5.07	Childrens	\$24.15
Arthritis	\$12.94	Effective	\$3.02	New	\$24.04
Attribute	**Excedrin**	**Attribute**	**Motrin**	**Attribute**	**Bayer**
Headache	\$88.77	Strong	\$41.82	Heart attack	\$40.39
Fast	\$73.75	Fast	\$35.02	Trust/Safety	\$25.72
Effective	\$14.83	Legs/muscle	\$15.81	Strong	\$24.03
New	\$12.06	Childrens	\$13.97	Back	\$18.11
Sinus	\$6.22	Headache	\$3.86	Legs/muscle	\$9.19
Doctor rec	\$3.02	Long lasting	\$3.35	Arthritis	\$7.47
Strong	\$2.38	Doctor rec	\$1.89	Headache	\$7.02

Figure 1. Dominant Advertised Attribute by Brand



The main proposition from the contribution game analysis is that a specific characteristic will only be advertised by one brand. Each characteristic is unlikely to be advertised by more than one brand. However, it is possible (and consistent with the statement) that one brand could advertise more than one characteristic. The proposition is evaluated by looking at the matrix of information cues in advertising in Figure 2 (and Table 4).

It is clear that the strong form of the proposition *(#?)* is rejected (each brand advertises significantly more than one characteristic). However, the weak form of the "sheening" model holds: the highest expenditure characteristic is unique for most of the brands. Advil spends most money to emphasize that it is strong, Aleve, that it is long lasting, Excedrin that it is headache medicine. Tylenol promotes speed most, and Bayer that it is good also for heart attack prevention. Motrin, however, follows Advil, as is suggested by their common active ingredient.

5.3 Quantitative Measures of Attributes

Our analysis also incorporates data on strength of pain relief, relative efficiency, and safety for each brand. We collected this information from peer-reviewed medical journals. Clinically, all four main active ingredients have varying degrees of side effects. Since individuals react to each ingredient differently, clinical pain researchers hesitate to assign superiority to any single drug. Although each of the drugs generally treat pain, fevers and headaches (hence implying that they are close substitutes), there are some differences between analgesic types. While aspirin, naproxen sodium and ibuprofen are non-steroidal anti-inflammatory drugs (NSAIDs), acetaminophen is not. In general, ibuprofen and naproxen are more potent pain relievers, i.e., they reduce more pain than the same dose of acetaminophen or aspirin. On the other hand, acetaminophen is considered to be the safest pain reliever because it does not block prostaglandins, and therefore does not cause gastrointestinal (GI) bleeding. However, even though acetaminophen reduces pain and fever, it does nothing for inflammation. Additionally, high doses of acetaminophen may damage the liver. Aspirin is the only pain reliever shown to reduce the risk of heart attack. All active ingredients deliver pain relief, but their effectiveness and safety profiles differ greatly.

We quantify or rank all the "true" characteristics that were used in advertising associated with each active ingredient as follows:

First, we interpret "fast" as the time taken to achieve perceptible or meaningful pain relief (medical literature terms this "onset to perceptible pain relief"). Each active ingredient can be unambiguously ranked by the onset of pain relief using the results in published medical studies. According to these clinical studies, the fastest pain relief is achicved with ibuprofen-based pain relievers.

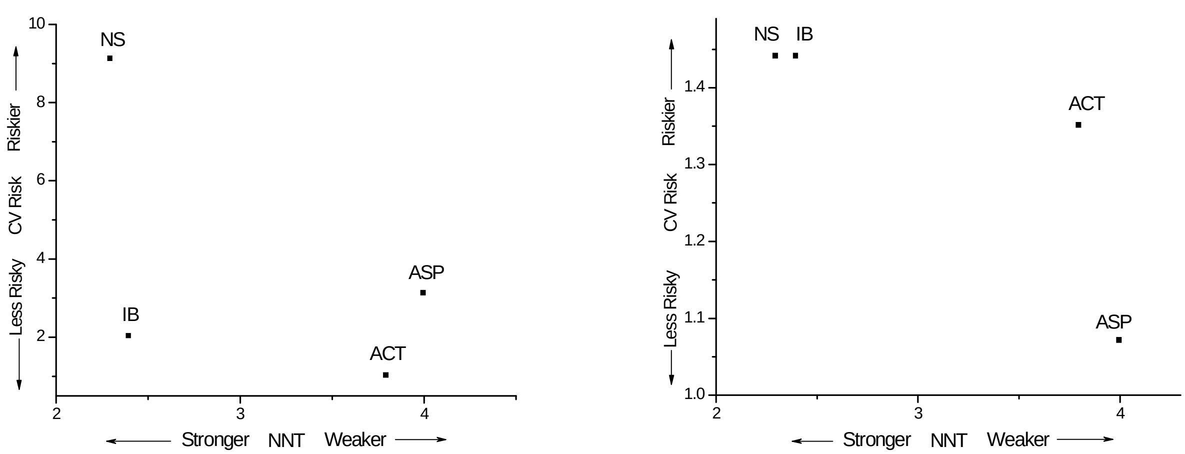
Second, we interpret claims such as "long lasting" as the duration of meaningful pain relief. The longevity measure is inversely related to the maximum number of regular strength pills allowed within a 24-hour period. Naproxen sodium tops the list in this regard, while acetaminophen has the lowest duration of pain relief.

Third, we interpret claims about strength (e.g., "strong," "stronger," "tougher on pain") as the maximum level of pain relief achieved; we use the number-needed-to-treat (NNT) measure to approximate analgesic efficiency claims. The NNT is a standard efficiency measure used in pain relief evaluation literature. Lower NNT number corresponds to higher pain relief potency.

Fourth, any claims related to gastrointestinal safety issues are interpreted as gastrointestinal side effects and in medical literature these risks are measured as GI Risk. Naproxen Sodium based pain relievers have profiles that are associated with the highest GI Risk.

Fifth, advertising claims mentioning cardiovascular side effects are interpreted as CV Risk. Naproxen Sodium and Ibuprofen have highest CV risks out of all OTC active ingredients.

Figure 2. Location of Active Ingredients in the Characteristics Space (GI/CV Risk and NNT)



These efficacy and safety measures for each brand are reported in Figure 2 and Table 3. Figure 2 plots GI Risk vs NNT (left panel) and CV Risk vs NNT (right panel). These panels in Figure 2 illustrate that there is a clear tradeoff between the potency of pain relief and the safety profiles of various pain relief molecules.

In Appendix A, we explain how NNT, cardiovascular risk, and gastrointestinal risk are calculated. In addition to using absolute risk and efficacy measures, we supplement the data with relative performance metric for speed and longevity of pain relief (see Table 3 for the relative rankings and Appendix A for explanation of all of these measures).

Table 3. Clinical Characteristics of OTC Analgesics

Active Ingredient	Dosage, mg. (maximum daily, mg)	Longevity Rating	Speed Rating	GI RR	CV RR
Aspirin	325-1000, every 4-6h (4000)	4	4	3.1	1.07
Acetaminophen	325-1000, every 4-6h (4000)	3	2	1	1.35
Ibuprofen	200-400, every 4-6h (1200)	2	1	2	1.44
Naproxen Sodium	220-440, every 8-12h (660)	1	3	9.1	1.44

6 Appendix A: Explanation of Clinical Measures

We reviewed 10 peer-reviewed medical journal articles to collect efficiency measures for:

1. Strength (Numbers Needed To Treat (NNT));
2. Speed (Onset to perceptible pain relief (Relative Speed));
3. Gastrointestinal side effects (Gastrointestinal risk (GI risk));
4. Cardiovascular side effects (Cardiovascular risk (CV risk).

These 4 efficacy measures for each brand are reported in Tables 2 and 3.

The peer-reviewed medical literature provides objective risk and efficiency measures for each product, based on its active ingredient (or combination of ingredients), strength, and recommended dosage. Each active ingredient has definitive maximum doses and durations of therapy. Differences exist across active ingredients in terms of the important safety issue of the potential for gastrointestinal toxicity and cardiovascular risk as well as relative strength and onset to perceptible pain relief. We collected the measurable characteristics for maximum OTC recommended dosage (single dose): Ibuprofen: 400 mg.; naproxen sodium: 440 mg.; aspirin: 1000 mg.; and acetaminophen: – 1000 mg.

Relative risk is the risk of an event (e.g., developing a disease) relative to exposure. Relative risk is the ratio of the probability of the event (E) occurring in the exposed group versus the control (nonexposed) group: $RR = (Pr\blacksquare(E|treatment))/(Pr\blacksquare(R|control))$

Relative risk is used frequently in clinical trial data to compare the risk of developing a disease in people not receiving the new medical treatment (or receiving a placebo) versus people receiving an established (standard of care) treatment. In the case of the gastrointestinal (GI Risk) and cardiovascular relative risk (CV Risk) numbers used herein, we use them to compare the risk of developing a side effect in people receiving a drug with people who do not receive the treatment (or receive a placebo). Thus, a cardiovascular relative risk of 1.44 means that cardiovascular problems arise with 44% higher likelihood using the drug (versus placebo).

The number needed to treat (NNT) is an epidemiological measure used in assessing the effectiveness of a health-care intervention, typically a treatment with medication. The NNT is the average number of patients who need to be treated to prevent one additional bad outcome (i.e. the number of patients that need to be treated for one to benefit compared with a control in a clinical trial). It is defined as the inverse of the absolute risk reduction. The ideal NNT is 1, where everyone improves with treatment and no one improves with control. The higher the NNT, the less effective is the treatment. More specifically, NNT is used with respect to two treatments, A and B, with A typically a drug and B a placebo. If the probabilities P_A and P_B under treatments A and B, respectively, are known, we can compute NNT as follows: $NNT = 1/(P_B - P_A)$

The NNT for a given therapy is simply the reciprocal of the absolute risk reduction($ARR = P_B - P_A$) for that treatment. For example, in a hypothetical migraine study, if risk decreased from $P_B = .30$ without treatment with drug M to $P_A = .05$ with treatment with drug M, for a relative risk of .17(.05/.3), a relative risk reduction of .83([.3 − .05]/.3), and an absolute risk reduction of .25(.3 − .05), the NNT would be 1/.25, or 4. In clinical terms, an NNT of 4 means that four patients need to be treated with drug M to prevent migraine from recurring in one patient. Typically, the lower the NNT number, the more potent and efficient the treatment is.

[some cites, or see - we can transfer this to appendix]

(1) Packman, B., E. Packman, G. Doyle, S. Cooper, E. Ashraf, K. Koronkiewicz, and S. Jayawardena (2000), "Solubilized Ibuprofen: Evaluation of Onset, Relief, and Safety of a Novel Formulation in the Treatment of Episodic Tension-Type Headache," Headache: The Journal of Head and Face Pain, 40 (7), 561-67.

(2) Hyllested, M., S. Jones, J.L. Pedersen, and H. Kehlet (2002), "Comparative Effect of Paracetamol, NSAIDs or Their Combination in Postoperative Pain Management: A Quali-

tative Review," British Journal of Anaesthesia, 88, 199-214.

(3) Forbes, J. A., C.K. Keller, J.W. Smith, J.R. Zeleznock, H. Sevelius, and W.T. Beaver (1986), "Analgesic Effect of Naproxen Sodium, Codeine, a Naproxen-Codeine Combination and Aspirin on the Postoperative Pain of Oral Surgery," Pharmacotherapy, 6 (5), 211-18.

(4) Milsom, I., M. Minic, Y. Dawood, M. Akin, J. Spann, N. Niland, and A. Squire (2002), "Comparison of the Efficacy and Safety of Nonprescription Doses of Naproxen and Naproxen Sodium with Ibuprofen, Acetaminophen, and Placebo in the Treatment of Primary Dysmenorrhea: A Pooled Analysis of Five Studies," Clinical Therapeutics, 24 (9), 1384-1400.

(5) Cooper, S.A., B.P. Schachtel, E. Goldman, S. Gelb, and P. Cohn (1989), "Ibuprofen and Acetaminophen in the Relief of Acute Pain: A Randomized, Double-Blind, Placebo-Controlled Study," Journal of Clinical Pharmacology, 29, 1026-1030.

(6) Olson, N.Z., A.M. Otero, I. Marrero, S. Tirado, S. Cooper, G. Doyle, S. Jayawardena, and A. Sunshine (2001), "Onset of Analgesia for Liquigel Ibuprofen 400 mg., Acetaminophen 1000 mg., Ketoprofen 25 mg., and Placebo in the Treatment of Postoperative Dental Pain," Journal of Clinical Pharmacology, 41 (11), 1238-47.

(7) Ong, C.K.S., P. Lirk, C.H. Tan, and R.A. Seymour (2006), "An Evidence-Based Update on Nonsteroidal Anti-Inflammatory Drugs," Clinical Medicine & Research, 5 (1), 19-34.

(8) Miller, D., C. Talbot, W. Simpson, and A. Korey (1987), "A Comparison of Naproxen Sodium, Acetaminophen and Placebo in the Treatment of Muscle Contraction Headache," Headache: The Journal of Head and Face Pain, 27 (7), 392-96.

(9) Lee, C., W.L. Straus, R. Balshaw, S. Barlas, S. Vogel, and T.J. Schnitzer (2004), "A Comparison of the Efficacy and Safety of Nonsteroidal Antiinflammatory Agents Versus Acetaminophen in the treatment of Osteoarthritis: A Meta-Analysis," Arthritis Rheum, 51, 746-54.

(10) Hersh, Elliot, Paul A. Moore, and Gilbert L. Ross (2000), "Over-the-Counter Analgesics and Antipyretics: A Critical Assessment," Clinical Therapeutics, 22 (5), 500-548.Conclusions

We have proposed here a theory of advertising sheening based on the desire to promote characteristics embodied in goods through persuasive advertising to render products more attractive through enhancing consumer appreciation of their embodied characteristics. The analysis has taken characteristics as exogenous. This fits well the OTC drugs context, and indeed drugs more generally, where chemical molecules are given (and there may be little scope from mixing molecules because of possible contra-indications). However, in other contexts the product characteristics themselves are likely endogenous. The analysis does suggest how characteristics choice might play out in equilibrium. In particular, when brands share common characteristics, they invoke a free-rider problem in their advertising of such characteristics. Choosing instead to differentiate enables brands to internalize better the sheening effects of advertising. We would thus expect the sheening motive to be a force towards product differentiation. This is borne out to some extent in the OTC drugs case in the sense that almost all main brands are based on different molecules.

7 Odd ends

here just holding unused parts for now

7.1 1

The "public goods contribution game" result above gives us a "corner" solution for all but one brand for each characteristic. The "MSI" paper tells us that we don't have support in the OTC data for the "full" hypothesis, but we do have a variant: each brand has a different ranking of characteristics advertised in terms of the dollars spent. Possible variants on the model that might lead this include

1) ad cost-side: the marginal cost of advertising a second characteristic may be lower – saying fast and strong may be cheaper than saying "fast" and "strong" separately. Federico has a paper on economies of scope (theory!)

2) Advertising "strong" may also sheen related words, like "fast" – so that other β's may be improved too. Or a word like "powerful" could sheen related words. Any mileage in a "word-use" variant on Gentzkow-Shapiro? (vague idea)

I don't think analysis is necessary here; we could just mention it. Or, a simple toy model version to illustrate (yet to be determined what – maybe something will come from what's below).

The "flaunt" idea: what you say depends on what you've got. Is it possible that you DON'T flaunt your dominant characteristic? Possible, to see.

Idea that some characteristics are exaggerated more than others, some differences over-played. Again, a bit vague for now.

Thinking of the structure of how much is spent on what characteristics, it seems that we have a structure reminiscent of Push-Pull, but with differences. Namely, if Advil advertises Strong, it has effects on ALL products' objective utilities (as opposed to just the target in PP). So the focs have some similar structure, but not clear yet that we can make similar substitutions – but we might know where to look (DIVRATS, e.g.). We might also introduce brand-enhancing ads into the model (like PP) to enable the pricing condition foc substitution.

7.2 Test-abilities

Notice first that we have the key variable W_i^k from knowing demands, and characteristics. (Suggesting we check if Medical Journals have new studies on molecules, better info on more charas would be great.) Jura to say something about looking for simpler fundamental charas – we have charas info for only 4 charas; do we associate charas in ads with one, and ignore other charas, can we collapse charas via multi-dimensional scaling, or such, can we engage hedonic regressions better ...

So, for structural empirics, the pattern of $W_i^k D_i$ predicts who does what (whoever i is biggest on k). How much it advertises is given by

$$W_i^k D_i \Phi' \left(A^k\right) = \alpha \gamma^k.$$

So we could set $\Phi(A) = \ln(k + A)$ to estimate

$$A_i^k = const. + const. W_i^k D_i.$$

Well, the model so far has one brand per chara; so we'll need to engage a version with multiple brands per chara! If this doesn't destroy the clean simplicity of the current theory results. Next challenge, what to do about this!

References

[1] Abernethy A.M., D. D. Butler. 1992. Advertising information: services versus products. *Journal of Retailing* 68(4) 398-419.

[2] Abernethy, A.M., G. R. Franke. 1996. The information content of advertising: a meta-analysis. *Journal of Advertising* 25(2) 1-17.

[3] Abernethy A.M., G. R. Franke. 1998. FTC regulatory activity and the information content of advertising. *Journal of Public Policy and Marketing* 17(2) 239-256.

[4] Amaldoss, W., C. He. 2009. Direct-to-Consumer Advertising of Prescription Drugs: A Strategic Analysis. *Marketing Science* 28(3) 472-487.

[5] Anand, B., R. Shachar. 2011. Advertising, the Matchmaker. *RAND Journal of Economics* 42(2) 205-245.

[6] Anderson, S. P., R. Renault. 2006. Advertising Content. *American Economic Review* 96(1) 93-113.

[7] Anderson, S. P., R. Renault. 2012. The Advertising Mix for a Search Good. *Management Science* forthcoming.

[8] Bagwell, K. 2007. The Economic Analysis of Advertising. *Handbook of Industrial Organization, Volume 3.* Eds. M. Armstrong and R. Porter, Elsevier, North Holland.

[9] Bar-Isaac, H., G. Caruana, V. Cuñat. 2010. Information Gathering and Marketing. *Journal of Economics and Management Strategy* 19(2) 375-401.

[10] Board, O. 2009. Competition and Disclosure. *Journal of Industrial Economics* 57(1) 197-213.

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

[12] Christou, C., N. Vettas. 2008. On Informative Advertising and Product Differentiation. *International Journal of Industrial Organization* 26(1) 611-629.

[13] Diamond, P. A. 1971. A Model of Price Adjustment. *Journal of Economic Theory* 3(2) 156-168.

[14] Dranove, D., G. Z. Jin. 2010. Quality Disclosure and Certification: Theory and Practice. *Journal of Economic Literature* 48(4) 935–63.

[15] Erdem, T., M. Keane, B. Sun. 2008. A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality. *Marketing Science* 27(6) 1111-1125.

[16] Farrell, J. 1986. Voluntary disclosure: robustness of the unraveling result and comments on its importance. In R. Grieson, ed.: *Proceedings of the U. C. Santa Cruz Conference on regulation and antitrust*. Lexington, MA., Heath-Lexington.

[17] Grossman, S. 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24(3) 461-483.

[18] Grossman, G., C. Shapiro. 1984. Informative Advertising and Differentiated Products. *Review of Economic Studies* 51(1) 63-81.

[19] Guo, L., Y. Zhao. 2009. Voluntary Quality Disclosure and Market Interaction. *Marketing Science* 28(3) 488-501.

[20] Iyer, G., D. Soberman, J. M. Villas-Boas. 2005. The Targeting of Advertising. *Marketing Science* 24(3) 461-476.

[21] Kalra, A., S. Li. 2008. Signaling Quality Through Specialization. *Marketing Science* 27(2) 168-184.

[22] Kamenica, E., M. Gentzkow. 2011. Bayesian Persuasion. *American Economic Review* 101(6) 2590-2615.

[23] Kihlstrom, R., M. Riordan. 1984. Advertising as a signal. *Journal of Political Economy* 92(3) 427-450.

[24] Koessler, F., R. Renault. 2011. When does a firm disclose product information? Mimeo, Universite de Cergy Pontoise, France.

[25] Konishi, H., M. T. Sandfort. 2002. Expanding Demand through Price Advertisement. *International Journal of Industrial Organization* 20(7) 965-994.

[26] Kuksov, D., J. M. Villas-Boas. 2010. When More Alternatives Lead to Less Choice. *Marketing Science* 29(3) 507-524.

[27] Madden, C. S., M. Caballero, S. Matsukubo. 1986. Analysis of information content in US and Japanese magazine advertising. *Journal of Advertising* 15(3) 38-45.

[28] Matthews, S., A. Postlewaite. 1985. Quality testing and disclosure. *RAND Journal of Economics* 16(3) 328-340.

[29] Mayzlin, Dina, Jiwoong Shin. 2011. Uninformative Advertising as an Invitation to Search. *Marketing Science* 30(4) 666-685.

[30] Mehta, Nitin, Xinlei (Jack) Chen, Om Narasimhan. 2008. Informing, Transforming, and Persuading: Disentangling the Multiple Effects of Advertising on Brand Choice Decisions. *Marketing Science* 27(3) 334-355.

[31] Mehta, N., S. Rajiv, K. Srinivasan. 2003 Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation. *Marketing Science* 22(1) 58-84.

[32] Meurer, M., D. O. Stahl. 1994. Informative Advertising and Product Match. *International Journal of Industrial Organization* 12(1) 1-19.

[33] Milgrom, P. 1981. Good news, bad news: representation theorems and applications. *Bell Journal of Economics* 12(2) 380-391.

[34] Milgrom, P. 2008. What the Seller Won't Tell You: Persuasion and Disclosure in Markets. *Journal of Economic Perspectives* 22(2) 115-131.

[35] Milgrom, P., J. Roberts. 1986a. Relying on the information of interested parties. *RAND Journal of Economics* 17(1) 18-32.

[36] Milgrom, P., J. Roberts. 1986b. Prices and Advertising Signals of Product Quality. *Journal of Political Economy* 94(4) 796-821.

[37] Nelson, P. 1970. Information and Consumer Behavior. *Journal of Political Economy* 78(2) 311-329.

[38] Nelson, P. 1974. Advertising as information. *Journal of Political Economy* 82(4) 729-754.

[39] Resnik, A., B. L. Stern. 1977. An analysis of information content in television advertising. *Journal of Marketing* 41(1) 50-53.

[40] Rob, R. 1985. Equilibrium price distribution. *Review of Economic Studies* 52(3) 487-504.

[41] Saak, A. E. 2006. The Optimal Private Information in Single Unit Monopoly. *Economics Letters* 91(2) 276-282.

[42] Salop, S. C. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10(1) 141-156.

[43] Shapiro, C. 1980. Advertising and Welfare: Comment. *Bell Journal of Economics* 11(2) 749-752.

[44] Stahl, Dale O.,II (1989): Oligopolistic Pricing and Sequential Consumer Search. *American Economic Review* 79, 700-712.

[45] Stahl, Dale O., II (1996): Oligopolistic Pricing with Heterogeneous Consumer Search. *International Journal of Industrial Organization* 14, 243-268.

[46] Stern, B. L., A. Resnik. 1991. Magazine Advertising: An analysis of its information content. *Journal of Advertising Research* 21 39-44.

[47] Sun, M. 2010. Disclosing Multiple Product Attributes. *Journal of Economics and Management Strategy* 20(1) 195-224.

[48] Weyl, E. G., M. Fabinger. 2009. Pass-Through as an Economic Tool, mimeo, Harvard University.

[49] Yee, M., E. Dahan, J. R. Hauser, J. Orlin. 2007. Greedoid-Based Noncompensatory Inference. *Marketing Science*26(4) 532-549.

[50] Zhang, J., L. Krishnamurthi. 2004. Customizing Promotions in Online Stores. *Marketing Science* 23(4) 561-578.

[51] Zhao, H. 2000. Raising Awareness and Signaling Quality to Uninformed Consumers: A Price-Advertising Model. *Marketing Science* 1994) 390-396.