Push-Me Pull-You: Comparative Advertising in the OTC Analgesics Industry

Anderson, Simon and Ciliberto, Federico and Liaukonyte, Jura and Renault, Regis

University of Virginia

14 April 2012

Push-Me Pull-You: Comparative Advertising in the OTC Analgesics Industry*

Simon P. Anderson[†], Federico Ciliberto[‡], Jura Liaukonyte[§], Regis Renault[¶]

Jan 2008; this Version: November 2012

Abstract

We model comparative advertising as brands pushing up own brand perception and pulling down the brand image of targeted rivals. We watched all TV advertisements for OTC analgesics 2001-2005 to construct matrices of rival targeting and estimate the structural model. These attack matrices identify diversion ratios and hence comparative advertising damage measures. We find that outgoing comparative advertising attacks are half as powerful as self-promotion in raising own perceived quality and cause more damage to the targeted rival than benefit to the advertiser. Comparative advertising causes most damage through the pull-down effect and has substantial benefits to other rivals.

Keywords: Comparative advertising, advertising targets, diversion ratios, attack matrix, push and pull effects, analgesics.

JEL Classification: L13, M37, L65.

---

*We thank Ross Rizley and gratefully acknowledge funding of the Marketing Science Institute under MSI Research Grant #4-1364. The first author thanks the NSF for support under grants SES 0452864 ("Marketing Characteristics") and GA10704-129937 ("Advertising Themes"); the fourth author thanks the Institut Universitaire de France. Anderson, Ciliberto and Liaukonyte thank the Bankard Fund for Political Economy at the University of Virginia. Catherine de Fontenay, JP Dubé, Fiona Scott-Morton, Massimo Motta, Joshua Gans, Phil Haile, Pauline Ippolito, Ginger Jin, Matt Shum, Andrew Sweeting, and Joel Waldfogel provided useful comments, as did participants at numerous seminars and at the Summer Workshop in Industrial Organization (Auckland, 2009) the NBER Summer Institute Meetings (Boston 2009), and CEPR (Toulouse 2010). We thank the Melbourne Business School and the Portuguese Competition Authority for their hospitality.

†University of Virginia and Center for Economic and Policy Research. Address: Department of Economics, University of Virginia, Charlottesville VA 22904, USA. sa9w@virginia.edu.

‡University of Virginia and Center for Economic and Policy Research. Address: Department of Economics, University of Virginia, Charlottesville VA 22904, USA. ciliberto@virginia.edu.

§Charles H. Dyson School of Applied Economics and Management, Cornell University, Ithaca NY 14850, USA. Jurate@cornell.edu.

¶Université de Cergy-Pontoise, ThEMA, 33 Bd du Port, F95000 Cergy-Pontoise, France. regis.renault@eco.u-cergy.fr.

1 Introduction

We investigate how brands strategically use comparative advertising. Comparative advertising both promotes positive perception of the advertiser's brand (the "push" effect) and detracts from the brand image of a targeted rival (the "pull" effect).[1] The distinctive characteristic of comparative advertising that sets it apart from purely self-promotional advertising is that specific rivals are targeted: advertising composition matters. Our fundamental objectives are i) to develop a novel theoretical model that explains who uses comparative advertising against whom and to what extent, ii) to apply this model to the US OTC analgesics industry where comparative advertisements are extensively used, and iii) to measure the damages inflicted.

We describe an equilibrium model, where brands simultaneously choose prices, self-promotion and comparative advertising expenditures, and derive the advertising first order conditions that predict oligopoly equilibrium relationships between advertising levels (for different types of advertising) and market shares. We use these first order conditions to rationalize the *attack matrix* of comparative advertising spending patterns against other brands. We show that the attack matrix identifies *diversion ratios* between brand pairs, where the diversion ratios measure the fraction of a target's lost consumers who are diverted to a rival brand following an attack. We employ the diversion ratios to estimate the damages inflicted by comparative attacks.

To estimate our model we use data constructed by Liaukonyte (2012), who watched over four thousand individual video files of all TV advertisements in the US OTC analgesics industry for 2001-2005 and recorded which brand(s) (or class of drugs) were compared (e.g. to Advil or Aleve; or to Ibuprofen-based drugs). The US OTC analgesics industry is particularly suitable for our analysis. First, comparative advertising is prevalent and represents a large fraction of industry sales. Second, data on advertising expenditures per ad is available for all brands for a reasonable time period. Video files are available and their content is readily

[1] The Pushmi-Pullyu is a fictitious two-headed llama befriended by Dr Doolittle. The heads are pointed in different directions. When one pushes forward, it pulls the other end back from its preferred direction.

coded to determine targets.

In the empirical analysis we deal with left-censoring of advertising (in some periods some brands do not engage in some types of advertising - there are corner solutions) and endogeneity of market shares and advertising expenditures. We control for left-censoring by running Tobit regressions. We control for endogeneity with brand fixed effects and two sources of exogenous variation: medical news shocks that hit the OTC analgesic market in the analyzed time period, and the prices of generic products, which we use as instrumental variables for the shares of the branded products.

Our empirical findings highlight how comparative advertising is inherently different from self-promotion. We find that outgoing attacks are about half as powerful as direct self-promotion ads in raising the brand's own perceived quality. But these attacks have a strong impact in terms of the damage that they cause to the target. This damage is heterogeneous across attacker-target pairs. For example, a marginal dollar of comparative advertising spent by Tylenol against Bayer reduces Bayer's profit by \$2, but a marginal dollar spent by Advil against Tylenol reduces Tylenol's profit by \$3. These losses are much larger than for pure self-promotion advertising. For instance if Tylenol increased its self-promotion expenditure by \$1, the decrease in its competitor's profit would range between 3 cents for Excedrin and 12 cents for Advil. Hence, much of the harm from comparative advertising comes from its negative impact on the target's perceived quality. We find that higher shares, ceteris paribus, are associated with higher self-promotion and comparative advertising advertising. Each extra consumer raises self-promotion advertising by 55 cents.

Comparative advertising also has substantial positive spillovers to rivals that are not being attacked. For example, a marginal dollar's comparative attack by Tylenol on Aleve increases Advil's profit by 20 cents. This means that the benefit the third party gets from denigration of the target's quality is larger than the loss from an improved perception of the attacker. These results indicate substantial "free-riding" in attacking any given target.

Despite the positive spillovers, the total damage to the industry (i.e., harm to target minus the benefits to other industry members) remains substantial. Our measures of the damage to the target are consistently and substantially above 1, which underscores the harm inflicted by comparative advertising: outgoing attacks cause much more damage to the target than benefit to the attacker. Spillovers are too small to make up for the difference. For example, the positive spillover to Advil of Aleve's marginal dollar attack on Tylenol is 39 cents, while the damage to Tylenol is over $3. These large numbers concur with the idea that comparative advertising can be very damaging overall, as suggested by the fact that they are used in few industries, and by commentators on the harmful effects of negative campaign ads in the political sphere.

Our paper is related to the growing empirical literature on the role of advertising content on market outcomes. In a complementary paper that we discuss at length below, Liaukonyte (2012) shows that the elasticity of demand is larger with respect to comparative advertising than self-promotion advertising. Ching et al. (2011) study the impact of content of media coverage on anti-cholesterol drugs on consumer demand and show that the impact varies by media type. Bertrand et al. (2010) develop a field experiment to show that advertising content significantly affects demand for loans, and conclude that advertising content persuades by appealing "peripherally" to intuition rather than reason. All these papers focus on the demand side. By contrast, we develop a model of firm strategic behavior (the "supply side") and use it to rationalize the comparative advertising attack patterns and to measure the magnitude of the damages inflicted with these attacks.

Our work is also related to papers that analyze equilibrium models of advertising. We contribute to this literature in two ways. First, due to our content data, we break down the ad expenditures into comparative and self-promotion expenditures, and the comparative expenditures are further broken down into attacker-target pairs. Second, we use exogenous shocks and brand-specific generic prices as sources of exogenous variation in the data. By contrast, Gasmi, Laffont, and Vuong (1992) use aggregate variables (e.g. the price of sugar) and Sovinsky Goeree (2008) uses entry and exit of products (Bresnahan, 1987), which is not feasible here because there is no entry of new brands.

Finally, our paper is related to the theoretical economics literature on comparative advertising. Anderson and Renault (2009) model it as directly informative revelation of horizontal match characteristics of products. Barigozzi, Garella, and Peitz (2009) and Emons and Fluet (2011) apply the signaling model of advertising (which goes back to insights in Nelson, 1974, and was formalized in Kihlstrom and Riordan, 1984, and Milgrom and Roberts, 1986). Our theory engages the complementary view of advertising (Stigler and Becker, 1977, and Becker and Murphy, 1993) with the added element of pulling down the rival. Thus, our approach is broadly consistent with advertising as a demand shifter (as in Dixit and Norman, 1978, and Johnson and Myatt, 2006).

2 The Model

2.1 Core Concepts

Using our coded advertising data we construct *attack matrices* of how much is spent by each advertiser against each rival target every month. These attack matrices allow us to identify *diversion ratios* that measure the substitutability between products. These diversion ratios are then used to find *damage measures* to a brand's profit from comparative advertising directed at that brand by different rivals. We now provide the intuition behind the use of diversion ratios, and link them to damage measures.

Let $\delta_j = Q_j - p_j$ be Brand j's attractiveness when it has quality Q_j and sets price p_j, and assume that market shares depend on j's attractiveness relative to its competitors. The *diversion ratio* from good j to k is the fraction of the market share lost by Brand j (due to a decrease in j's attractiveness) that is captured by Brand k.[2] It is defined as

$$d_{jk} = -\frac{ds_k/d\delta_j}{ds_j/d\delta_j} \in (0,1)\,, \tag{1}$$

where s_j is the market share of Brand j. One way to think of d_{jk} is in terms of consumers' second preferences: some consumers switch to their next preferred option when the first choice gets less attractive. For substitute differentiated products, d_{jk} is positive, and $\sum_k d_{jk} < 1$ because some customers no longer purchase at all when j gets less attractive.

[2]The diversion ratio has been proposed as a useful statistic for analyzing the price effects of mergers (see for example Shapiro, 1996, and recent development by Jaffe and Weyl, 2011).

It is useful to interpret the diversion ratio as the *neutralizing price change* that keeps j's market share the same after a drop of \$1 in k's attractiveness (e.g., following a rise in k's price by \$1). Such a lower rival attractiveness causes a $(-ds_j/d\delta_k)$ increase in j's market share. Now, this is exactly the market share picked up by k if j's attractiveness went down \$1, because the switching consumers are those broadly indifferent between j and k as first choice. This symmetry property implies that the increase in j's market share is equivalently $(-ds_k/d\delta_j)$.[3] On the other hand, a rise in j's price of Δp_j will cause j's market share to drop by $\Delta p_j\,(ds_j/d\delta_j)$. Equating these expressions gives the neutralizing price change as[4]

$$\Delta p_j = \frac{-ds_k/d\delta_j}{ds_j/d\delta_j} = d_{jk}. \tag{2}$$

The importance of the neutralizing price change is that we can measure the change in j's profit from a decrease in k's attractiveness as simply the price change applied to j's market, or $\Delta\pi_j = \Delta p_j M s_j = M s_j d_{jk}$, where M is the market base of potential consumers. This underscores why it is the outbound diversion ratio, d_{jk}, that matters in determining the worth of inbound customers. It also suggests that the diversion ratio should enter the marginal benefit for Brand j of targeting Brand k through comparative advertising, which adversely impacts Q_k. Indeed, let \$1 spent by j on comparative advertising against target k reduce Q_k by ΔQ_k (which is a positive number because it is defined as a reduction): this negative impact on k's attractiveness we call the "pull effect". The neutralizing price change argument above gives the marginal benefit for Brand j from the pull effect as $M s_j d_{jk} \Delta Q_k$.

Because comparative advertising is also advertising for Brand j, there is also a "push" effect from an increase in Brand j's attractiveness. This is the amount of pure self-promotion spending that would result in the same change in j's attractiveness as a \$1 increase in comparative advertising, and is therefore the marginal rate of substitution between them. We assume it is constant at rate λ. Because the push effect of a comparative ad returns λ per dollar, optimal arrangement of the ad portfolio implies the pull effect must return $1-\lambda$

[3] See Anderson, de Palma, and Thisse (1992), Ch.3, p. 67.

[4] If a \$1 price rise by k allows j to pick up 10 of the customers shed by k, and a \$1 price rise by j loses it 50 consumers (10 of which would go to j, incidentally, by the symmetry property), then the neutralizing price hike for j is 20 cents. The diversion ratio from j to k is 1/5.

per dollar (whenever comparative advertising is used against a target). Hence the optimal comparative advertising strategy of Brand j is characterized by $Ms_j d_{jk} \Delta Q_k = 1 - \lambda$ for any rival k it chooses to target. Diversion ratios may then be identified from the condition that comparative advertising expenditures should equate the marginal benefit to the marginal advertising cost (which is \$1).

The above condition also indicates that once we know the diversion ratios, we can write the drop in Brand k's attractiveness induced by one more dollar of comparative advertising by j targeted at k as $\Delta Q_k = \frac{1-\lambda}{Ms_j d_{jk}}$. This is therefore also the amount by which k must reduce its price to neutralize the hit to Q_k. Similarly, using the neutralizing price change interpretation of d_{kj}, it is readily shown that $\frac{d_{kj}}{Ms_j}$ is the drop in price that Brand k must incur in order to maintain its market share if Brand j were to raise Q_j by increasing its self-promotion by \$1 from its equilibrium level: a \$1 comparative ad only raises Q_j by a fraction λ of what \$1 self-promotion does. Pulling all this together, the harm to k's equilibrium profit of one more dollar of comparative advertising by j is:[5]

$$Ms_k \left(\frac{1-\lambda}{Ms_j d_{jk}} + \lambda \frac{d_{kj}}{Ms_j} \right), \tag{3}$$

where the first term in parentheses is the price drop that neutralizes the pull-down to Q_k and the second one is the price drop that neutralizes the push-up to Q_j.

2.2 Demand

Suppose that Brand $j = 1, ...n$ charges price p_j and has perceived quality $Q_j(.)$, $j = 1, ...n$. We retain the subscript j on $Q_j(.)$ because when we get to the estimation, exogenous variables such as medical news shocks and random variables summarizing the unobserved determinants of perceived quality will enter the errors in the equations to be estimated.

Brands can increase own perceived quality through both types of advertising, and degrade competitors' quality through comparative advertising. Comparative advertising, by its very nature of comparing, both raises own perceived quality and reduces the perceived quality of rival brands. The corresponding arguments of $Q_j(.)$ are advertising expenditure by Brand j

[5] Our analysis below derives this using the envelope theorem.

which directly promotes its own product, denoted by A_{jj}; "outgoing" advertising by Brand j targeted against Brand k, A_{jk}, $k \neq j$, which has a direct positive effect; and "incoming" comparative advertising by Brand k targeting Brand j, A_{kj}, $k \neq j$, which has a negative (detraction) effect on Brand j's perceived quality. Thus, we write j's perceived quality as $Q_j(A_{jj}, \{A_{jk}\}_{k \neq j}, \{A_{kj}\}_{k \neq j})$, $j = 1, ..., n$, which is increasing in the first argument, increasing in each component of the second (outgoing) group, and decreasing in each component of the third (incoming) group, with $\frac{\partial^2 Q_j}{\partial A_{jj}^2} < 0$ and $\frac{\partial^2 Q_j}{\partial A_{kj}^2} > 0$ for $k \neq j$.[6]

The demand side is generated by a discrete choice model of individual behavior where each consumer buys one unit of her most preferred good. We will not estimate this demand model from (aggregate) choice data; we simply use it to frame the structure of the demand system. Preferences are described by a (conditional indirect) utility function:

$$U_j = \delta_j + \varepsilon_j, \qquad j = 0, 1, ..., n, \tag{4}$$

in standard fashion, where ε_j is a brand-idiosyncratic match value and

$$\delta_j = Q_j\left(.\right) - p_j \tag{5}$$

is the "objective" utility, and where we let the "outside option" (of not buying an OTC pain remedy) be associated with an objective utility δ_0.

The distribution of the random terms determines the form of the market shares, s_j, $j = 0, ..., n$, and each s_j is increasing in its own objective utility, and decreasing in rivals' objective utilities.[7] Assume that there are M consumers in the market, so that the total demand for brand j is Ms_j, $j = 0, ..., n$.

[6] Throughout, we assume sufficient concavity that the relevant second order conditions hold.

[7] For example, $s_j = \frac{\exp[\delta_j/\mu]}{\sum_{k=0}^{n} \exp[\delta_k/\mu]}$, $j = 0, ..., n$ in the standard multinomial logit model.

2.3 Equilibrium Relations

Assume that product j is produced by Brand j at constant marginal cost, c_j. Brand $j's$ profit-maximizing problem is:

$$\underset{\{p_j, A_j\}}{Max}\, \pi_j = M(p_j - c_j)s_j - A_{jj} - \sum_{k \neq j} A_{jk} \qquad j = 1, ...n. \tag{6}$$

where the advertising quantities (the A's) are dollar expenditures.

Prices and advertising levels are determined simultaneously in a Nash equilibrium.

The price condition is determined in the standard manner by:

$$\frac{d\pi_j}{dp_j} = Ms_j - M(p_j - c_j)\frac{ds_j}{d\delta_j} = 0, \qquad j = 1, ...n, \tag{7}$$

which yields a solution $p_j > c_j$: brands always select strictly positive mark-ups.

Self-promotion advertising expenditures are determined (for $j = 1, ..., n$) by:

$$\frac{d\pi_j}{dA_{jj}} = \frac{d\pi_j}{d\delta_j}.\frac{\partial Q_j}{\partial A_{jj}} - 1 = M(p_j - c_j)\frac{ds_j}{d\delta_j}\frac{\partial Q_j}{\partial A_{jj}} - 1 \leq 0, \text{ with equality if } A_{jj} > 0 \tag{8}$$

where the partial derivative function $\frac{\partial Q_j}{\partial A_{jj}}$ may depend on any or all of the arguments of Q_j. Substituting the pricing first-order condition (7) into the advertising one (8) gives

$$Ms_j\frac{\partial Q_j}{\partial A_{jj}} \leq 1, \quad \text{with equality if } A_{jj} > 0, \qquad j = 1, ..., n.^8 \tag{9}$$

The interpretation is that raising A_{jj} by \$1 and raising price by $\$\frac{\partial Q_j}{\partial A_{jj}}$ too leaves δ_j unchanged. This change, therefore, increases revenue by $\$\frac{\partial Q_j}{\partial A_{jj}}$ on the existing consumer base (i.e., Ms_j consumers). This extra revenue is equated to the \$1 cost of the change, the RHS of (9). The relation in (9) implicitly determines self-promotion as a function of whatever advertising variables are in Q_j (these all involve brand j as either sender or target), along with j's share.

Recalling our assumption that $\frac{\partial^2 Q_j}{\partial A_{jj}^2} < 0$, from (9), brands with larger s_j will advertise more (choose a higher value of A_{jj}) than those with smaller market shares, ceteris paribus. The intuition is that the advertising cost per customer is lower for larger brands. The other

[8] Below we (implicitly) invoke sufficient concavity of Q_j for interior solutions to (9): if $\frac{\partial Q_j}{\partial A_{jj}}$ were constant (if ads entered perceived quality linearly), then this is unlikely.

relations in the following proposition follow similarly from the implicit function theorem through the dependence of perceived quality on self-promotion, and incoming and outgoing attacks. Through the next series of Propositions, we emphasize the various second derivatives of the Q function because these correspond to the parameters we estimate.

Proposition 1 *(**Self-promotion Advertising levels**) Brand j's choice of self-promotion advertising level is determined by $Ms_j \frac{\partial Q_j}{\partial A_{jj}} \leq 1$, with equality if $A_{jj} > 0$. For $A_{jj} > 0$, A_{jj} is an increasing function of s_j; A_{jj} is a decreasing function of A_{jk} iff $\frac{\partial^2 Q_j}{\partial A_{jj} \partial A_{jk}} < 0$; A_{jj} is an increasing function of A_{kj} iff $\frac{\partial Q_j}{\partial A_{jj} \partial A_{kj}} > 0$.*

The advertising relationships in the Proposition 1 hold for brands with large enough market shares.[9] They will be estimated below using a simple Q_j specification for which A_{jj} is written as a linear function of s_j and the other relevant advertising quantities.

We now turn to comparative advertising levels. An attack raises own perceived quality and decreases that of the targeted rival. We can determine the advertising spending against rivals by differentiating (6) to get (for $k \neq j$):

$$\begin{aligned}
\frac{d\pi_j}{dA_{jk}} &= \frac{d\pi_j}{d\delta_j} \cdot \frac{\partial Q_j}{\partial A_{jk}} + \frac{d\pi_j}{d\delta_k} \cdot \frac{\partial Q_k}{\partial A_{jk}} - 1 \\
&= \underbrace{M(p_j - c_j) \frac{ds_j}{d\delta_j} \frac{\partial Q_j}{\partial A_{jk}}}_{\text{own Q enhancement}} + \underbrace{M(p_j - c_j)(\frac{ds_j}{d\delta_k}) \frac{\partial Q_k}{\partial A_{jk}}}_{\text{competitor's Q denigration}} - 1 \leq 0,
\end{aligned}$$

with equality if $A_{jk} > 0$. We proceed by substituting the attacker pricing condition and its self-promotion condition to rewrite this comparative advertising condition in a form to be estimated. First, inserting the price first-order conditions (7) gives (for $k \neq j$):

$$\frac{d\pi_j}{dA_{jk}} = Ms_j \frac{\partial Q_j}{\partial A_{jk}} - Ms_j d_{jk} \frac{\partial Q_k}{\partial A_{jk}} \leq 1, \quad \text{with equality if } A_{jk} > 0, \tag{10}$$

where $d_{jk} > 0$ is the diversion ratio discussed in sub-section 2.1 above. Loosely, the diversion ratio measures how much custom is picked up from a rival per customer it sheds. The restriction on the diversion ratios ($d_{jk} \in [0, 1]$) motivates restrictions below in the estimation.

[9] Otherwise, from (7) the term $(p_j - c_j) \frac{ds_j}{d\delta_j}$ is small enough that the derivative $\frac{d\pi_j}{d\delta_j}$ in (8) is negative when $\frac{\partial Q_j}{\partial A_{jj}}$ is evaluated at $A_{jj} = 0$.

The comparative advertising derivative, (10), provides a bound on the size of the marginal rate of substitution between *outgoing* comparative advertising and self-promotion $(\frac{\partial Q_j}{\partial A_{jk}}/\frac{\partial Q_j}{\partial A_{jj}})$. Assume for the present argument that the solution for self-promotion spending (see (9)) is interior. Then, substituting the self-promotion condition $(Ms_j\frac{\partial Q_j}{\partial A_{jj}}=1)$ into (10) implies

$$\frac{\partial Q_j}{\partial A_{jk}}/\frac{\partial Q_j}{\partial A_{jj}} \leq 1 + Ms_jd_{jk}\frac{\partial Q_k}{\partial A_{jk}} \tag{11}$$

where the LHS is less than one because $\frac{\partial Q_k}{\partial A_{jk}} < 0$ on the RHS. In summary:

Proposition 2 ***(Self-promotion and outgoing comparative advertising)*** *If Brand j uses a strictly positive amount of self-promotion, then the marginal rate of substitution between outgoing comparative advertising against Brand k and self-promotion $(\frac{\partial Q_j}{\partial A_{jk}}/\frac{\partial Q_j}{\partial A_{jj}})$ is strictly below 1.*

If this were not the case, then comparative advertising would drive out self-promotion since it would give a direct own-quality benefit per dollar greater than self-promotion, while additionally helping the attacker by denigrating a rival. We will assume in the estimation that the marginal rate of substitution between outgoing comparative advertising and self-promotion in (11) is constant, at rate λ, so that the testable implication of Proposition 2 is that $\lambda < 1$. Then we can write from (11):

$$(0 <) \quad -Ms_jd_{jk}\frac{\partial Q_k}{\partial A_{jk}} \leq 1-\lambda, \qquad \text{with equality if } A_{jk} > 0. \tag{12}$$

The intuition is as follows for $A_{jk} > 0$. The term $1-\lambda$ on the RHS of (12) is the marginal cost of the pull effect once we subtract the value of the push component of the comparative attack. Hence the LHS should be the marginal benefit of the pull effect. To see that this is so, first note that the pull effect of raising A_{jk} by \$1 is equivalent to brand k raising its price by $\$\frac{-\partial Q_k}{\partial A_{jk}}$ (since the same δ_k is attained). The neutralizing price change for j that just keeps s_j intact per dollar increment in p_k is given by (2) as d_{jk}, and this benefit is reaped on j's market base of Ms_j. The LHS of (12) then follows directly.

To determine predictions for how A_{jk} depends on the other relevant advertising levels, we apply the implicit function theorem to (12) and recall that $\frac{\partial^2 Q_k}{\partial A_{jk}^2} > 0$.

Proposition 3 *(**Comparative Advertising levels**) The choice of comparative advertising level by Brand j against Brand k is determined by $-Ms_j d_{jk} \frac{\partial Q_k}{\partial A_{jk}} \leq 1 - \lambda$, with equality if $A_{jk} > 0$. For $A_{jk} > 0$, A_{jk} is: (i) an increasing function of d_{jk} and s_j; (ii) a decreasing function of A_{lk} iff $\frac{\partial^2 Q_k}{\partial A_{jk} \partial A_{lk}} > 0$; (iii) an increasing function of A_{kk} iff $\frac{\partial^2 Q_k}{\partial A_{kk} \partial A_{jk}} < 0$; (iv) an increasing function of A_{kl} iff $\frac{\partial^2 Q_k}{\partial A_{kl} \partial A_{jk}} < 0$; .*

From Proposition 3(i), there are more attacks for given diversion ratio d_{jk} the higher the attacker market share. This is roughly borne out in the raw data insofar as Advil and Aleve are the largest attackers of Tylenol. Likewise, for a given attacker share, attacks are larger for a bigger diversion ratio.[10] We shall proceed for the estimation by estimating d_{jk} for each pair. Thus we are implicitly constraining the diversion ratios to be constant over time.

From Proposition 3(ii), attacks by j against k increase with attacks on k by others if and only if $\frac{\partial^2 Q_k}{\partial A_{jk} \partial A_{lk}} < 0$. This cross partial sign implies that more harm is inflicted with a marginal attack by j when others' attacks render k more susceptible.

The third property in Proposition 3 depends on the sign of the cross partial $\frac{\partial^2 Q_k}{\partial A_{kk} \partial A_{jk}}$; we now argue that the last one does too. Indeed, the cross-partial $\frac{\partial^2 Q_k}{\partial A_{kl} \partial A_{jk}}$ (used in the fourth property) has the same sign as does $\frac{\partial^2 Q_k}{\partial A_{kk} \partial A_{jk}}$ because we know $\frac{\partial Q_k}{\partial A_{kl}} = \lambda \frac{\partial Q_k}{\partial A_{kk}}$ with both derivatives positive by assumption, and λ therefore positive, so the assumption of λ constant implies the two cross partials have the same sign.

Hence, the last two properties in Proposition 3 are both determined by the sign of the cross partial $\frac{\partial^2 Q_k}{\partial A_{kk} \partial A_{jk}}$, which is estimated in the self-promotion equation. Hence, applying Proposition 1 to 3(iii) and 3(iv) yields the next result.

Corollary. If self-promotion is increasing with incoming comparative advertising then comparative advertising decreases with target self-promotion and with target outgoing comparative advertising.

These are implications of the model, and not imposed by functional form. The intuition is that a brand is attacked less when it advertises more if having more outgoing ads reduces

[10] Alternatively, we can write $s_j d_{jk} = s_k D_{jk}$ where $D_{jk} = \frac{s_j}{s_k} d_{jk}$ is the ratio of cross elasticity of demand to own elasticity. In this case, for a given value of D_{jk}, a bigger target is attacked more. This roughly concurs with the data that the largest firm, Tylenol, is attacked most.

the negative impact of attacks (i.e., $\frac{\partial^2 Q_k}{\partial A_{kk}\partial A_{jk}} > 0$), and this is also the condition for a brand to want to engage in more self-promotion when attacked more (its marginal benefit rises with incoming attacks).

We now show how the damage to a rival from j's self-promotion depends on the diversion ratio. The effect on k's profits, $\pi_k^* = M(p_k^* - c_k)\, s_k^* - A_{kk}^* - \sum_{l \neq k} A_{kl}^*$ (where the stars denote equilibrium values) holding constant all other brands' actions (except the best-reply of k) is determined by the envelope theorem as

$$\begin{aligned} \frac{d\pi_k^*}{dA_{jj}} &= M(p_k^* - c_k)\frac{ds_k}{d\delta_j}\frac{\partial Q_j}{\partial A_{jj}} \\ &= -\frac{s_k}{s_j} d_{kj} \end{aligned} \tag{13}$$

where at the second step we have substituted in k's pricing condition (7) and the equality version of (9).

Similarly, the measure of the damage of an extra dollar of comparative advertising from Brand j against target k is a weighted average of push and pull effects, both of which can be written in terms of diversion ratios. Using the envelope theorem, the full effect of a marginal dollar of comparative advertising from j on $k's$ profits, with all other brands' actions fixed is

$$\frac{d\pi_k^*}{dA_{jk}} = M(p_k^* - c_k)\left(\frac{ds_k}{d\delta_k}\frac{\partial Q_k}{\partial A_{jk}} + \frac{ds_k}{d\delta_j}\frac{\partial Q_j}{\partial A_{jk}}\right).$$

Substituting in k's pricing condition (see (7)) implies

$$\begin{aligned} \frac{d\pi_k^*}{dA_{jk}} &= Ms_k\left(\frac{\partial Q_k}{\partial A_{jk}} - d_{kj}\frac{\partial Q_j}{\partial A_{jk}}\right) \\ &= -\frac{s_k}{s_j}\left(\frac{1-\lambda}{d_{jk}} + \lambda d_{kj}\right) \end{aligned} \tag{14}$$

where we have substituted in the equality versions of conditions (12) and (9) at the second step.[11] The interpretation of (14) in terms of neutralizing prices was given in Section 2.1 (see (3)). Basically, the first term here is the amount of self-promotion required to restore Q_k and the second term is the harm inflicted by the rival's increased self-promotion component

[11] Equivalently, we can write this as $\frac{d\pi_k^*}{dA_{jk}} = (1-\lambda)\, Pull_{jk} + \lambda Push_{jk} = \frac{(1-\lambda)}{D_{jk}} + \lambda D_{kj}$.

of the comparative advertising (hence the λ weight corresponding to the push effect). Note that the effect on profit here and below is measured in dollars: equivalently (by the target's optimality condition that the \$1 marginal cost of an extra dollar's advertising equals its marginal benefit), it is the amount of self-promotion advertising that would have to be spent to offset the harm. The empirical analysis will provide parameter estimates so the marginal harm can be estimated.

Proposition 4 ***(Damage Measure)*** *Assume that target k engages in self-promotion, and assume that outgoing comparative ads are perfectly substitutable with self-promotion at rate $\lambda \in (0,1)$. Then the profit lost by target k from an additional dollar of comparative advertising attack by Brand j is the sum of a pull damage, $\frac{1-\lambda}{d_{jk}} \frac{s_k}{s_j}$, and a push damage, $\lambda d_{kj} \frac{s_k}{s_j}$.*

In like manner we can determine the spillover benefit (related to free riding in comparative advertising) to l of an attack by j on k as

$$\frac{d\pi_l^*}{dA_{jk}} = \frac{s_l}{s_j} \left(d_{lk} \frac{(1-\lambda)}{d_{jk}} - \lambda d_{lj} \right). \tag{15}$$

The first term here is the direct benefit to l from the harm inflicted on k (pull); the second is (as above) the damage incurred by l from j improving its quality through the comparative advertising channel (push). This expression can readily be interpreted in terms of neutralizing price changes.

3 Description of Industry and Data

The OTC analgesics market is worth approximately \$2 billion in retail sales per year (including generics) and covers pain-relief medications with four major active chemical ingredients. These are Aspirin (ASP), Acetaminophen (ACT), Ibuprofen (IB), and Naproxen Sodium (NS). The nationally advertised brands are such familiar brand names as Tylenol (ACT), Advil and Motrin (IB), Aleve (NS), Bayer (ASP or combination), and Excedrin (ACT or combination). Table 1 summarizes market shares, ownership, prices and advertising levels in this industry.

TABLE 1. Market Shares and Advertising Levels of OTC Analgesics Brands

Brand	Active Ing.	Price / serving	Inside Market Share	Max Pills	TA/ Revenue	CA/ Revenue	CA/ TA	Owner-ship
Tylenol	ACT	$2.15	30.51%	7.2	17.4%	3.3%	19.3%	McNeil
Advil	IB	$1.60	24.21%	5.9	20.0%	13.3%	66.4%	Wyeth
Aleve	NS	$0.83	22.40%	3.0	26.0%	20.0%	75.7%	Bayer
Excedrin	ACT	$2.40	8.28%	9.2	26.4%	3.4%	13.2%	Novartis
Bayer	ASP	$1.85	6.98%	10.1	28.8%	6.4%	22.4%	Bayer
Motrin	IB	$1.71	7.68%	5.9	20.4%	8.1%	39.6%	McNeil
Generic	ACT	$1.17						
Generic	IB	$0.66						
Generic	ASP	$0.82						
Generic	NS	$0.57						

3.1 Sales Data

The sales data, collected by AC Nielsen, consist of prices and dollar total revenues of all OTC oral analgesics products sold in the U.S. national market from March of 2001 through December of 2005 (58 monthly observations).

We construct a measure of a *serving* of pain medication, or a *pain episode*, so that we can aggregate across different package sizes and across different medication strengths.[12] We define the *market size,* M*,* for OTC analgesic products as the US population 18 years or older minus the number of people who buy pain medication at Wal-Mart, a store that does not provide information on the sales of products. We then express each product's sales as the number of people whose pain could be relieved by that product for a period of three days, which is the average number of pain days per month in the population.[13]. To this end we assigned to each analgesic product in the sales dataset the strength of its active ingredient in milligrams and derived the maximum number of pills that a consumer can take for OTC analgesics consumption over 72 hours as defined by the FDA and required to be listed on the labels (e.g. 9 in the case of Aleve, and from 18 to 36 for Tylenol, depending on the ACT formulation). This we refer to as an episode of pain.

[12] A detailed description of how we construct the dataset is provided in Appendix A.

[13] This information is from the Morbidity and Mortality Weekly Report, Centers for Disease Control and Prevention, Feb 27, 1998/47(07); 134-140.

Then, we compute each brand's market share as the fraction of total number of episodes of pain sold by the brand over market size. The average price of a pain episode is computed as the ratio of the total sales revenue of a brand in a month to the total number of episodes of pain sold in that month. We do the same calculation for the generic products, which differ from each other only by their active ingredient. The resulting output is the time series of average prices of episodes of pain relief for each of the four active ingredients for the generic products. We maintain that the generic products are provided by a competitive fringe and that the generic prices are set equal to their marginal cost.

3.2 Advertising Data

Our advertising dataset is from TNS-Media Intelligence. The data include video files of all TV advertisements for 2001-2005 for each brand advertised in the OTC analgesics category and monthly advertising expenditures on each ad. The unit of observation in the raw dataset is a single ad. There are $4,506$ unique ads (346 of which are missing videos).

We watched all the ads and coded their content. We recorded whether the product was explicitly compared to any other products. If a commercial was comparative, we recorded which brand (or class of drugs) it was compared to (e.g., to Advil or Aleve). If an ad had multiple targets, the ad was assigned equally among them.

If an ad had no comparative claims, it was classified as a self-promotion ad. In the data we observe situations when brands made indirect attacks on their competitors. An indirect attack occurs when one brand makes a claim against "all other regular" brands. We code such indirect attacks as self-promotion. We discuss other coding scheme alternatives in Appendix E.

Table 2 presents the complete picture of cross targeting and advertising expenditures on each of the rival brands targeted. This table shows that *every* nationally advertised brand used comparative advertising during the sample period. However, only four (of the six) brands were targeted: Tylenol, Advil, Aleve, and Excedrin.[14] These data provide some in-

[14]Motrin does not attack Tylenol because the parent company is the same; likewise, Bayer does not attack Aleve for the same reason.

formal support that larger brands both used more comparative advertising and were targeted more. Entries on the diagonal are self-promotion expenditures.

TABLE 2. Advertising and Comparative Advertising Target Pairs

Adver-	TARGET:							
tiser ⇓	Advil	Aleve	Bayer	Excedrin	Motrin	Tylenol	Total CA	Total
Advil	92.1 [50]	17.8 [27]	-	4.3 [20]	-	160.2 [58]	182.2	274.3
Aleve	-	42.5 [45]	0.0 [3]	0.5 [7]	-	131.7 [58]	132.1	174.7
Bayer	13.8 [25]	-	104.9 [58]	-	-	15.7 [37]	29.5	131.8
Excedrin	-	1.9 [7]	2.2 [7]	158.4 [47]	-	19.9 [15]	24.1	182.5
Motrin	18.9 [27]	18.8 [27]	-	-	57.3 [54]	-	37.6	94.9
Tylenol	9.6 [16]	31.7 [31]	36.6 [27]	-	-	359.0 [58]	77.8	404.0
Total	42.6 [68]	70.2 [92]	38.7 [34]	4.7 [27]	-	327.5	483.4	

Notes: Row j indicates the advertiser brand and Column k indicates the target. The left part of cell jk is comparative ad expenditure in \$m.; the right part denotes how many time periods [out of 58] the attack pair jk happened. The diagonal entries are expenditures on self-promotional advertising.

3.3 News Shocks

The OTC analgesics market endured several major medical news shocks over the analyzed time period. Following the approach presented by Chintagunta, Jiang, and Jin (2009) we utilized Lexis-Nexis to search over all articles published between 2001 and 2005 on relevant topics. We recorded the article name, source, and date to construct a dataset of news shocks. Multiple articles reporting the same event were assigned to a unique shock ID. Additionally, we checked whether a news shock was associated with any new medical findings that were published in major scientific journals. Finally, we focused only on the events that were reported in a major national newspaper (USA Today, Washington Post, Wall Street Journal, New York Times). After this data cleaning, our news shock dataset includes 8 major news shocks between March of 2001 and December of 2005. Table 3 reports the news shocks by their title, date, and the original scientific publication.

After some experimentation, we determined that the effects of the news shocks fade out after three months. Still we consider two possibilities for the duration of each news shock in consumer memory. We construct a dummy variable for a short-term shock variant that takes value 1 at time t when the shock occurred, and for the next three months (i.e., t through $t+3$). Then, to check the robustness of our analysis, we construct another variable, which captures the possibility that consumers have a long-term memory. The dummy variable for the long-term shock takes value 1 at time t till the end of the sample period.

TABLE 3. Medical News Shocks

No	News Shock Description	Date	Source
1	Risk of Cardiovascular Events Associated With Selective COX-2 Inhibitors	8/21/2001	Journal of the American Medical Assoc (JAMA); 2001,286:954-959
2	Ibuprofen Interferes with Aspirin	12/20/2001	New England Journal of Medicine, 2001, 345:1809-1817
3	FDA Panel Calls for Stronger Warnings on Aspirin and Related Painkillers	9/21/2002	FDA Public Health Advisory
4	Aspirin Could Reduce Breast Cancer Risk/ NSAIDs Protect Against Alzheimer's	4/8/2003/ 4/2/2003	JAMA 2004; 291:2433-2440 American Academy Of Neurology
5	Anti-Inflammatory Pain Relievers Inhibit Cardioprotective Benefits of Aspirin	9/9/2003	Circulation, 9/9/2003
6	Vioxx Withdrawn From the Market	9/30/2004	
7	Long Term Naproxen (Aleve) Use may Increased Cardiovascular Risk	12/23/2004	FDA Public Health Advisory
8	Bextra Withdrawn	4/7/2005	

4 The Econometric Model

4.1 A Quality Function

After *extensive* experimentation, we implement the following perceived quality function:[15]

$$Q_j(.) = \alpha \ln\left(A_{jj} + \lambda \sum\nolimits_{k\neq j} A_{jk} - \beta \sum\nolimits_{k\neq j} A_{kj} + \bar{A}_{jj}\right) - \varphi \sum\nolimits_{k\neq j} \ln\left(\bar{A}_{kj} + A_{kj}\right) + \bar{Q}_j. \quad (16)$$

Variables other than advertising levels pertaining to j's perceived quality enter through $\bar{A}_{jj}$, $\bar{A}_{kj}$, and $\bar{Q}_j$. They include observed factors such as j's product characteristics or news shocks as well as unobserved factors that determine the realization of random shocks. They enter the equations to be estimated only if they interact with advertising levels, that is only if

[15]The use of logarithmic functions to ensure the concavity of the utility with respect to advertising is common in the literature. See, for example, Busse and Rysman (2005). See Bagwell (2007) for an extensive review of the literature on the economics of advertising.

they enter $\bar{A}_{jj}$ or $\bar{A}_{jk}$ for some k. Here, we interpret $\bar{Q}_j$ as the product differentiation from product characteristics and the remaining part of $Q_j(.)$ as the differentiation induced by advertising. This distinction is important when we discuss the identification strategy and we look into the nature of the structural unobservables because anything that enters into $\bar{Q}_j$ can be used as an instrumental variable in the advertising first order conditions.

The push effect is incorporated through the weighted sum of self-promotion and outgoing comparative ads ($A_{jj} + \lambda \sum_{k \neq j} A_{jk}$), where λ is the marginal rate of substitution between outgoing comparative and self-promotion ads, which is assumed to be constant. In order for self-promotion to favorably impact perceived quality, α should be positive. Recall from Proposition 2 that we should expect $\lambda < 1$. Whether there is a push effect for Brand j associated with its comparative advertising activity against rivals depends on whether λ is strictly positive or not.[16]

The pull effect from incoming comparative ads (A_{kj}) impacts the quality function in two ways. First, it enters the "net persuasion" term inside the first logarithm. The sign of β gives the sign of the cross effect between incoming attacks and outgoing ads. Second, incoming ads enter in a separable way with associated parameter ϕ. This additional term allows for disassociating the intensity of the overall pull effect from the intensity of the cross effect between incoming attacks and outgoing ads as measured by β.[17] Through this separable term, we also allow the A_{kj} to be imperfect substitutes with one another. Since attacks on target k constitute a public good for all the brands other than k, if expenditures attacking k were perfect substitutes, then there would be only one attacker in equilibrium in each period. The data show that this is not the case.

The comparative statics properties in Propositions 1 and 3 that link self-promotion and comparative advertising expenditures are determined by the signs of parameters α, β, and λ.[18]

[16] $\lambda < 0$ would mean that j's brand image is hurt by the use of comparative advertising, in line with conventional wisdom among marketers in continental Europe.

[17] With ϕ large enough, it also ensures that $\frac{\partial^2 Q_k}{\partial A_{jk}^2} > 0$ locally.

[18] This specification of Q imposes the sign of the cross effect between attacks by k on j and attacks by some other Brand l on j to have the sign of $-\alpha\beta^2$ so it is negative (provided that α is found to be positive). Then

4.2 The Equations To Be Estimated

The first order condition for self-promotion ads, corresponding to equation (9) above may be written as

$$\begin{cases} A^*_{jjt} = \alpha M s_{jt} - \lambda \sum_{k\neq j} A_{jkt} + \beta \sum_{k\neq j} A_{kjt} - \bar{A}_{jjt}, \\ \bar{A}_{jjt} \sim N\left(\mu_{jjt}, \sigma^2_{SP}\right), \ A_{jjt} = \max\left(A^*_{jjt}, 0\right), \qquad j = 1, ..., n. \end{cases} \tag{17}$$

A very attractive feature of our modeling strategy is that $\bar{A}_{jjt}$ incorporates the structural unobservable component of perceived quality that interacts with A_{jjt}. Subscripts j and t on the mean term reflect some possible brand fixed effect as well as the possible impact of some observable shocks such as news shocks. The equation above is a Tobit regression that is linear in the parameters.

The first order condition for comparative ads follows from first writing (12) for the specification of quality (16) above. This gives

$$-M s_{jt} d_{jkt} \left(\frac{-\alpha\beta}{A_{kkt} + \lambda \sum_{l\neq k} A_{klt} - \beta \sum_{l\neq k} A_{lkt} + \bar{A}_{kkt}} - \frac{\phi}{\left(\bar{A}_{jkt} + A_{jkt}\right)} \right) \leq 1 - \lambda,$$

with equality if $A_{jkt} > 0$. Second, using the target k's self-promotion equation (9) when $A_{kkt} > 0$ (namely $A_{kkt} + \lambda \sum_{l\neq k} A_{klt} - \beta \sum_{l\neq k} A_{lkt} + \bar{A}_{kkt} = \alpha M s_{kt}$), we obtain the following econometric specification:

$$\begin{cases} A^*_{jkt} = \varphi M s_{jt} \frac{s_{kt} d_{jk}}{(1-\lambda) s_{kt} - \beta s_{jt} d_{jk}} - \bar{A}_{jkt}, \\ \bar{A}_{jkt} \sim N\left(\mu_{jkt}, \sigma^2_{CA}\right), \ A_{jkt} = \max\left(A^*_{jkt}, 0\right), \qquad j = 1, ..., n. \end{cases} \tag{18}$$

as long as $A_{kkt} > 0$. Here again, the structural unobservable is in $\bar{A}_{jkt}$. In our estimation strategy, we assume that diversion ratios are constant over time, and given by $d_{jkt} = d_{jk}$. Equation (18) is a Tobit regression that is nonlinear in the parameters.

Some of the firms in our analysis are multi-brand firms. Motrin and Tylenol are owned by McNeil, and Aleve and Bayer are owned by Bayer. We treat each brand as making independent decisions. This is not a problem at all for the self-promotion equation, which is exactly the same if we allow firms to behave as multi-brand firms that maximize joint payoffs. However, the comparative ad equation would be modified to include cross-brand effects.

from Proposition 3, more attacks by other brands on j induce more comparative advertising by k against j.

This would require the estimation of a large number of additional diversion ratios with the same number of observations, which in exploratory work resulted in many diversion ratios being imprecisely estimated. We therefore treat brands as independent divisions maximizing brand profits, modulo the imposition that they do not attack sibling products as concurs with the data in this respect.

4.3 Identification Strategy

In both Tobit specifications above, the unobservables are correlated with the explanatory advertising and share variables because the brands take them into consideration when making their advertising and pricing decisions. The first, most straightforward, step to address the endogeneity of these variables is to exploit the panel structure of our data to account for time-constant differences across brands. Essentially, for the self-promotion equation, we set $\bar{A}_{jjt} = \bar{A}_{jj} + \Delta\bar{A}_{jjt}$, where $\bar{A}_{jj}$ is a brand fixed effect, while $\Delta\bar{A}_{jjt}$ are time-specific idiosyncratic shocks. We do not follow the same approach for the comparative ad equation since this would require estimating many pair specific dummy variables $\bar{A}_{jk}$, which cannot be achieved with much precision, given our limited number of observations. Hence the endogeneity of shares in the comparative ad equation (18) is only dealt with using instrumental variables, as described below. The dummy variables in the self-promotion equation (17) control for a brand's advertising base allure advantage, which picks up any persistent component of such an advantage. The remaining source of endogeneity in our regressions then comes from any potential correlation of temporary shocks, here picked up by $\Delta\bar{A}_{jjt}$ and $\bar{A}_{jkt}$, with advertising expenditures and shares.

The second step is to explore whether the data on news shocks can explain some of the correlation of $\Delta\bar{A}_{jjt}$ and $\bar{A}_{jkt}$ with advertising expenditures and shares. That is, brands observe the shocks, which affect their shares, and which affect their advertising and pricing decisions. Thus, if we include the news shocks as being part of $\Delta\bar{A}_{jjt}$, then we deal with some of the correlation between the temporary shocks on perceived quality and the advertising expenditures and shares. News shocks are clearly exogenous because they require new medical discoveries, which "surprise" both consumers and brands, and alter the perception of the health properties of the products.[19]

Finally and alternatively, we use an instrumental variable approach. Rather than assuming that news shocks contribute to the advertising base allure terms, $\bar{A}_{jjt}$ and $\bar{A}_{kjt}$, we suppose that their impact is separable from that of advertising expenditures, although they enter the brand's perceived quality. In other words they only enter the $\bar{Q}_j$ term in (16). They are therefore proper candidates for instrumental variables. In addition, generic prices and various functions of them can be used as instrumental variables as long as the marginal cost of production of a generic product does not depend on the quantity produced. In particular, all of the patents for the OTC analgesics have expired. After patents expire, generic counterparts are produced at prices that are substantially lower than the brand name product (Aleve). Over time, new generic entry brings the price of the generic counterpart down to marginal cost, as shown by Grabowski and Vernon (1992) and discussed by Ching (2004) and Scott Morton (2004). If the marginal cost is constant and the generic prices are set at the marginal cost, then the generic prices are independent of the prices set by the national brands and we can use them as instrumental variables.

To implement the estimation in our non-linear models, we use control functions (Heckman and Robb 1985, 1986). Our methodology follows Blundell and Smith (1986) and Rivers and Vuong (1988). Consider the self-promotion equation. Using control functions consists of rewriting the unobservable $\bar{A}_{jjt}$ as a linear function of v, the unobservable of the first stage reduced form regression, and of ϵ, a white noise term. For example, say that only shares are suspected to be endogenous. Then, v is the unobservable of a reduced form regression of the shares on all the exogenous variables, including the instrumental variables. We can then use the residuals from that reduced form regression, $\hat{v}$, and plug them in the regression (17) as follows: $A^*_{jjt} = \alpha M s_{jt} - \lambda \sum_{k \neq j} A_{jkt} + \beta \sum_{k \neq j} A_{kjt} + \theta \hat{v} + \epsilon$, where ϵ is now the unobservable that generates the Tobit model. The nice feature of this approach is that we

[19] This works better for consumers: firms are more likely to know when findings are in the offing. The idea of using a natural experiment to study the effect of advertising (on prices) is also in Milyo and Waldfogel (1999).

can test the exogeneity of the shares by testing whether $\theta = 0$. With three endogenous variables, we have three control functions, but the problem is conceptually identical. The only econometric difficulty in the application of this methodology is created by the fact that two of the explanatory variables in the self-promotion equation, $\sum_{k \neq j} A_{jkt}$ and $\sum_{k \neq j} A_{kjt}$, are left-censored, and thus the estimated residuals that are required to construct the control functions would be biased whenever the variables are zero. To address this econometric problem, we derive the generalized residuals, as proposed by Gourieroux et al. (1987). We describe the econometric approach in detail in Appendix B. Because of the nonlinear nature of all these problems we estimate the system of the two equations (17) and (18) separately rather than with the generalized method of moments (as in Sovinsky Goeree, 2008).

4.4 Methodological Remarks

We deliberately estimate only the supply side of the model and steer clear of distributional assumptions on the demand side.

First, we maintain that a supply-side analysis directly explains the supply-side question of who attacks who by how much, and the strategic use of different types of advertising to maximize profits; whereas a demand-side analysis is the best tool to investigate demand-side questions, such as the effects of different advertising content on demand. A contribution that we make is to show how to fully exploit the information on the supply side to estimate the parameters of the model without adding *unnecessary assumptions* on the demand side. In particular, we use the unique feature of comparative advertisement that it targets a specific rival. It enables us to find demand-side relations in the form of diversion ratios. This would not be possible in a market with purely self-promotional advertising or with standard equilibrium pricing relations that conflate all effects into a single variable (advertising or price).

Second, to estimate the supply side we need to make the equilibrium behavior assumptions explicit, which is not required in the demand estimation. However, any demand estimation that uses instrumental variables implicitly invokes an equilibrium model (although its explicit structure may be obscured) because the first stage instrumental variable regression is a reduced form equation of the right hand side endogenous variables (e.g. prices). Thus, any demand estimation implies an equilibrium behavior assumption, even though that equilibrium behavior assumption is not explicitly spelled out. In addition, in order to run counter-factuals on the supply side as in Bresnahan (1987), we would still need to make an equilibrium behavior assumption. Thus, we cannot address the question of this paper (what explains the attack matrix) without equilibrium behavior assumptions.

5 Empirical Analysis

5.1 Self-Promotion

Each column in Table 4 presents the results for the parameters α, β, and λ for various specifications of Equation (17). Across all specifications, α, β, and λ are positive and statistically significant. The results in Proposition 1 that larger shares are associated with more self-promotion advertising is reflected in the positive sign of α. Outgoing attacks have a push-up self-promotion impact measured by $\lambda > 0$. However, because $\lambda < 1$, comparative advertising does not drive out self-promotion, as per Proposition 2. The direct own-quality benefit per dollar is smaller than the benefit from self-promotion. Finally, $\beta > 0$ means that self-promotion increases with incoming advertising. This reflects a positive cross effect, which, by Proposition 3, implies that comparative advertising decreases with target self-promotion. None of these empirical results reject the theoretical model. Next, we investigate the economic significance of the results in Table 4.

Column 1 of Table 4 shows the results from a straightforward Tobit regression, where self-promotion ad expenditures are regressed on sales, outgoing attacks and incoming attacks. We estimate $\alpha = 0.123$, which means that a brand would spend 12 cents a month more in self-promotion per additional customer. The marginal rate of substitution between outgoing attacks and self-promotion ads is $\lambda = 0.768$, meaning that the self-promotion value of \$1 of outgoing comparative ads is the same as 77 cents of pure self-promotion. The value $\beta = 0.429$ provides a lower bound to how much additional self-promotion expenditures will offset one more dollar of attacks on the brand (43 cents).[20] We now investigate how the results change when we address the endogeneity of the explanatory variables.

In Column 2 we run the Tobit regression including a dummy variable that is equal to 1 if the observation is for one of the top brands (Advil, Aleve, Tylenol), and zero otherwise.[21] Thus, we have $\mu_{jjt} = \mu^{TB}$ for a top brand and $\mu_{jjt} = \mu^{OB}$ otherwise. Using this specification, the coefficient estimate of λ drops from 0.768 to 0.660 and the coefficient estimate of β drops from 0.429 to 0.297. In contrast, the coefficient estimate of α increases from 0.123 to 0.432. The contrasting direction of the bias between the advertising explanatory variables and the shares reflects the relationship between the unobserved component of perceived quality and the explanatory variables. In particular, it is reasonable to think that products with a higher unobserved component of perceived quality will have a larger market share, *ceteris paribus*. Then, the downwards bias on α when the fixed effect is omitted means that brands with a stronger unobserved component of perceived quality do less self-promotion advertising, *ceteris paribus*. Similarly, the upwards bias on the estimates for λ and β means that brands with a higher perceived quality are attacked less and attack rivals less than brands with a lower perceived quality. These predictions are consistent with our specification of perceived quality, which assumes a negative cross partial between $\bar{A}_{jj}$ and outgoing ads and a positive cross partial between $\bar{A}_{jj}$ and incoming attacks. This discussion is mirrored by the result on the coefficient estimate of the Top Brand dummy. The Top Brand fixed effect, $\bar{A}_{jj}^{TB}$ is equal to -0.353. It has a negative sign, which means that the larger brands, Aleve, Tylenol, and Advil have inherently higher advertising base allure than the other brands.

In Column 3 we add the variables that measure the occurrence of a news shock using the short term memory definition. With the exception of the estimate of α that increases from 0.432 to 0.513, the results in Column 3 are remarkably similar to those in Column 2, suggesting that adding the short-term memory news shocks as control variables does not change the way the model fits the data. This is consistent with the low values of the F

[20] It is not the full extent of the negative impact of attacks on the brand's perceived quality. This requires knowing ϕ, which is identified from estimating the comparative advertising equations (18).

[21] More discussion on the use of a top brand dummy variable is available in Appendix C.

TABLE 4. Self-Promotion Equation and Net Persuasion

Version	Baseline	Brand Dummy	News Shocks Short term	News Shocks Long Term	IV (Generics)	IV (Generics & Short Term Shocks	IV (Generics & Long Term Shocks)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Alpha	0.123	0.432	0.513	0.515	0.551	0.552	0.570
	(0.027)	(0.076)	(0.078)	(0.074)	(0.045)	(0.046)	(0.046)
Lambda	0.768	0.660	0.643	0.631	0.616	0.616	0.629
	(0.072)	(0.074)	(0.073)	(0.071)	(0.087)	(0.062)	(0.075)
Beta	0.429	0.297	0.251	0.258	0.447	0.446	0.430
	(0.063)	(0.068)	(0.070)	(0.066)	(0.037)	(0.038)	(0.032)
Control: Out. Ads					−0.018	−0.025	−0.014
					(0.071)	(0.053)	(0.059)
Control: Inc. Ads					−0.164	−0.165	−0.151
					(0.035)	(0.032)	(0.031)
Control: Shares					−0.309	−0.314	−0.332
					(0.043)	(0.039)	(0.043)
Brand dummy		−0.353	−0.440	−0.439	−0.525	−0.527	−0.541
		(0.081)	(0.085)	(0.079)	(0.054)	(0.051)	(0.054)
/sigma	0.195	0.189	0.181	0.175	0.185	0.185	0.185
	(0.008)	(0.008)	(0.007)	(0.007)	(0.004)	(0.003)	(0.003)
Log likelihood	47.955	57.082	75.089	83.085	63.680	63.807	64.039
F-tests Shocks In			$F(8,336) = 3.70$	$F(8,336) = 6.94$			
F-test: Outg. Ads					$F(6,341) = 6.27$	$F(14,333) = 3.92$	$F(14,333) = 4.08$
F-test: Inc. Ads					$F(6,341) = 22.58$	$F(14,333) = 11.09$	$F(14,333) = 11.05$
F-test: Shares					$F(6,340) = 40.09$	$F(14,332) = 18.46$	$F(14,332) = 17.68$
Obs	348	348	348	348	348	348	348

Note: Coefficient estimates of the constant (included in all the specifications) and of the news shocks are available from the authors. Bootstrapped standard errors are computed in columns 5-7.

statistic associated with the test that all the coefficients of the news shocks are equal to zero. The results in Column 4 show that adding the long-term memory news shocks as control variables does not change the way the model fits the data either.

To investigate whether we should still be concerned about any remaining endogeneity of s_j, $\sum_{k \neq j} A_{jkt}$, and $\sum_{k \neq j} A_{kjt}$, we run three instrumental variable regressions. In Column 5 the instrumental variables are the generic prices of the product that shares the same active ingredient and the sum of the generic prices over all the competing active ingredients. In Column 6 we add the short-term memory news shocks, which are then excluded from the second stage regression. In Column 7 we instead add long-term memory news shocks. We find that the instrumental variables do a fair job at explaining the first stage variation in outgoing comparative advertising and in incoming attacks. The first-stage F tests reject the null hypotheses that generic prices do not explain any of the first stage variation, and the F statistics are quite large, except for the one associated with the first stage regression for $\sum_{k \neq j} A_{jkt}$. Instrumental variables are less important to control for the endogeneity of shares, since the brand dummies predict most of the variation in shares.

Columns 5-7 show that $\alpha \approx 0.55$, which means that Brand j spends 55 cents per month in self-promotion advertising per additional consumer. We also find $\lambda \approx 0.6$ which means that each dollar spent in outgoing comparative advertising is worth approximately 60 cents in raising own perceived quality and the remaining 40 cents are gained from pulling down a competitor. $\beta \approx 0.44$ means that incoming attacks have at least a damage of 44 cents (and, as we calculate below, the full damage is much larger). The results in Column 5 shows that the variation in generic prices controls for the endogeneity of the variable $\sum_{k \neq j} A_{kjt}$ and of the variable s_j. Notice that the estimate of α is the same in Columns 3-5, suggesting that the instrumental variable approach controls for the endogeneity of s_j to the same extent as adding news shocks does. The control function for $\sum_{k \neq j} A_{jkt}$ is not statistically significant, suggesting (from Blundell and Smith, 1986) that the endogeneity of $\sum_{k \neq j} A_{jkt}$ is not empirically significant. Columns 6 and 7 show the coefficient estimates do not change when we add short or long term memory shocks to the generic prices, but in some cases the estimates become slightly more precise.

5.2 Comparative Advertising and Diversion Ratios

Table 5 presents the estimation results for the parameter φ and for the diversion ratios d_{jk}. The diversion ratios are treated as parameters to be estimated from the data and are restricted to be between 0 and 1. Treating diversion ratios as parameters avoids imposing a functional form on demand. Rather, we are implicitly using a linear approximation. This approximation strategy may be vindicated by the stability of market shares over the period. Berry (1994) shows that, under fairly lenient regularity conditions on the joint distribution of random terms in (4), there is an invertible relation between market shares and mean utilities, δ_j. Since diversion ratios are determined by the vector of mean utilities, they should be essentially unchanged if market shares do not vary much.[22] It is worth noting that with more observations our general methodology would allow the diversion ratios to be a function of market shares or of any variables that the researcher believes might determine the degree of substitutability between products, on top of a pair specific component.

Recall that we use a two-step approach. We first estimate (17). Then, we plug the estimates of β and λ into (18) to estimate φ and the diversion ratios. Thus, each Column in Table 5 corresponds to one specification of (18) in Table 4. In particular: Columns 1 and 2 use the estimates of β and λ that we obtain from Column 6 in Table 4; Column 3 uses the estimates of β and λ from Column 7 in Table 4; finally, Column 4 uses the estimates of β and λ from Column 5 of Table 4. All specifications use the same number of observations (601). Twelve diversion ratios are estimated. There are three reasons for a diversion ratio to be missing. First, there were too few or no attack months so the variable was omitted. For example, Aleve attacked Advil only three times (see Table 2). Second, there are no direct attacks on "siblings." For example, Bayer does not attack Aleve (both are owned by the same parent company). Third, we do not estimate (18) whenever the attacker or the target did no self-promotion (see also equation (11)).

[22] The exception is Aleve, which suffered a loss of market share in 2005, but recovered in a few months.

The coefficient estimates of the control functions for the shares of the attacker (s_{jt}) and of the attacked (s_{kt}) are statistically insignificant and of small magnitude in Columns 2-4, implying that the endogeneity of market shares is not empirically significant. This is not surprising in light of the fact that market shares are quite stable over time while advertising expenditures vary quite a bit (see Appendix A for more on this). Column 1, which presents the main results for this section, does not include control functions. Henceforth we discuss the economic implications of the coefficient estimates in Column 1.

Consider the entry d_{AlT}, the diversion ratio from Aleve to Tylenol. In the second column we estimate d_{AlT} equal to 0.153, meaning that if Aleve sheds 100 consumers through a price rise (say), then 15.3 of them go to Tylenol. Now consider the entry d_{AdT}, the diversion ratio from Advil to Tylenol. We estimate d_{AdT} to be virtually the same number. This is fairly large, suggesting that Tylenol is a fairly large gainer from both Aleve and Advil. The two brands attack Tylenol in very similar fashion. Looking back at Table 2, we observe that Advil and Aleve both attack Tylenol every month. More striking is the fact that their overall expenditures are very close, with Advil spending a total of \$160 million and Aleve spending \$132 million attacking Tylenol.

The figure for d_{ET} is surprisingly low (at 10.2%) since Excedrin and Tylenol share acetaminophen as active ingredient in many of its variants, but it might indicate that Excedrin serves specialty niches of consumers (Excedrin markets itself as a migraine medicine) interested in its combinations with caffeine and with aspirin (which Tylenol does not have). Motrin equally loses to Advil and Aleve an approximate 16%, despite sharing the same active ingredient with Advil.

Next, Bayer loses even more (20.3%) to Tylenol, which suggests that consumers perceive Tylenol as the closest substitute to Bayer. This concurs with the findings of a number of medical studies (e.g. Hyllested et al., 2002), according to which Tylenol is the second safest branded OTC pain reliever, after Bayer (based on cardiovascular and gastrointestinal risk profiles). Yet, Tylenol loses more to Aleve than to Bayer, suggesting that substitution patterns are not symmetric.[23] Indeed, a price rise loses Tylenol just 11.9% to its 3 main attackers, but it picks up at least that amount following a price increase by either of them.

The diversion ratios for each of the six brands sum to less than 1, as the theory hopes for (we imposed them each to be below one, but we did not restrict the sum). For example, we see that if a consumer leaves Tylenol, then that consumer will go with probability 2.6% to Advil, 5.0% to Aleve, and 4.3% to Bayer. With the remaining 88.1% probability a consumer will switch to the outside good or some other OTC analgesics, branded or generic.

There are three pairs for which we estimate the diversion ratios in both directions: $(Bayer, Tylenol,)$, $(Advil, Tylenol)$, $(Aleve, Tylenol)$. Comparing them indicates relative own demand derivatives. In particular, $\frac{d_{jk}}{d_{kj}} = \frac{ds_k/d\delta_k}{ds_j/d\delta_j}$. Take for example j =Tylenol and k =Bayer. Because we have $d_{TB} = 0.043$ and $d_{BT} = 0.203$ so $\frac{d_{TB}}{d_{BT}}$ is around $\frac{1}{5}$ (so $ds_T/d\delta_T \approx 5ds_B/d\delta_B$). This means the demand derivative is much more price sensitive for Tylenol. At first blush, this may seem to presage a poor prospect for the estimates, given that Tylenol has a much higher price than Bayer aspirin (suggesting a more inelastic demand). However, a rough calibration brings this into perspective. The price of a "serving" (here roughly 3 days of pain relief) of Tylenol is roughly \$2.15; taking the generic price of \$1.17 as representing marginal cost gives a mark-up of approximately \$1. A similar mark-up is found for Bayer, with a brand price of \$1.85 and a generic price of about \$0.8. The pricing equation (7) sets mark-up equal to demand over own demand derivative (in absolute value). Using the market shares of 0.3 for Tylenol and 0.07 for Bayer (these are rough inside market shares as a fraction of total market including generics, without outside good), the pricing formula predicts a demand derivative for Tylenol of 0.3 and for Bayer of 0.067, which gives us a 1-to-4.6 ratio that is very close to the one that we get from the ratio of diversion ratios.

[23] It is also possible, but we cannot check it given the data we have, that as far as Bayer is concerned, consumers leaving Tylenol switch to the generic version of aspirin. Because generics do not use comparative advertising, we cannot estimate those diversion ratios.

TABLE 5. Comparative Advertising Equation and Diversion Ratios

	No IV (Using β and λ from Column 2 of Table 4) (1)	IV: Generics and Short Term Shocks (Using β and λ from from Column 6 of Table 4) (2)	IV: Generics and Long Term Shocks (Using β and λ from from Column 7 of Table 4) (3)	IV: Generics Only (Using β and λ from from Column 5 of Table 4) (4)
ALEVE ON:				
Tylenol, d_{AlT}	0.153 (0.028)	0.119 (0.027)	0.138 (0.030)	0.201 (0.031)
ADVIL ON:				
Tylenol, d_{AdT}	0.153 (0.028)	0.120 (0.026)	0.139 (0.030)	0.199 (0.032)
Aleve, d_{AdAl}	0.045 (0.019)	0.026 (0.017)	0.037 (0.018)	0.045 (0.022)
Excedrin, d_{AdE}	0.014 (0.017)	0.001 (0.013)	0.011 (0.015)	0.000 (0.019)
TYLENOL ON:				
Advil, d_{TAd}	0.026 (0.015)	0.020 (0.009)	0.022 (0.013)	0.024 (0.021)
Aleve, d_{TAl}	0.050 (0.015)	0.030 (0.030)	0.041 (0.015)	0.056 (0.021)
Bayer, d_{TB}	0.043 (0.011)	0.029 (0.028)	0.038 (0.012)	0.049 (0.014)
BAYER ON:				
Advil, d_{BAd}	0.152 (0.067)	0.081 (0.055)	0.121 (0.060)	0.165 (0.078)
Tylenol, d_{BT}	0.203 (0.063)	0.136 (0.054)	0.184 (0.061)	0.251 (0.077)
MOTRIN ON:				
Advil, d_{MAd}	0.167 (0.060)	0.100 (0.054)	0.140 (0.052)	0.191 (0.084)
Aleve, d_{MAl}	0.162 (0.060)	0.090 (0.062)	0.128 (0.055)	0.167 (0.081)
EXCEDRIN ON:				
Tylenol, d_{ET}	0.102 (0.068)	0.058 (0.050)	0.092 (0.072)	0.104 (0.089)
Control Function for s_j		0.038 (0.083)	0.013 (0.077)	-0.000 (0.098)
Control Function for s_k		-0.021 (0.072)	-0.045 (0.062)	-0.058 (0.058)
ϕ	0.595 (0.135)	0.731 (0.144)	0.646 (0.142)	0.411 (0.148)
Constant Term	-0.159 (0.039)	-0.125 (0.041)	-0.150 (0.039)	-0.131 (0.065)
Variance Unobservable	0.140 (0.008)	0.140 (0.008)	0.138 (0.008)	0.140 (0.008)
Log-Likelihood Function	11.323	11.677	11.626	11.290
Number of Observations	601	601	601	601

Note: Bootstrapped standard errors are shown.

Whenever we have both diversion ratios, the diversion from small to large is greater than vice versa. This property would hold with a logit demand (recall for logit $d_{jk} = \frac{s_k}{1-s_j}$). For logit, d_{jk} is increasing in s_k (as customers are shed, they go to other brands in proportion to those brands' shares). This works well: the only, important, violation is from Tylenol to Advil and Aleve. However, other properties of the logit do not hold. For logit, d_{jk} is increasing in s_j but we see no clear relation in the table of diversion ratios on this count.

5.3 Damage and Spillover Measures

We now derive measures of the damage that comparative advertising delivers to the attacked brand and the spillovers to other brands. We use the coefficient estimates of λ from Column 6 of Table 4 and of the diversion ratios and φ from Column 1 of Table 5.

As discussed when deriving (14), the full damage can be decomposed into a push and a pull effect. Table 6 shows the damage measures that we can estimate given the pattern of attacks observed in the data. Targets are column entries, and attackers are on the rows. The entries are written as dollar damages to targets from a \$1 marginal increment in comparative advertising by the attacker. These are all positive numbers, so are all costs inflicted.[24]

The first entry is the impact on k of j's self-promotion push-up. From (13), the damage to k is given as $\frac{s_k}{s_j} d_{kj}$, and so this term is reported whenever d_{kj} is reported. If we multiply this by λ we get the impact of the push effect of outgoing comparative advertising by j, and hence the second term in (14). The second entry in the Table is the direct pull effect of an attack by j on k, which is given by the first term in (14) as $\frac{s_k}{s_j} \frac{(1-\lambda)}{d_{jk}}$, and so this is reported whenever d_{jk} is reported. When both effects are reported (i.e., when we have the diversion ratios in both directions), we can sum the pull effect with λ times the pure push effect to generate the third entry, which is the total damage on the target of a marginal dollar of comparative advertising. We report the bootstrapped 90% confidence intervals in square brackets underneath the point estimates. Several remarks follow from Table 6.

First, imprecision in the results of Table 5 feeds through to imprecise results in Table 6.

[24] The damage numbers can be interpreted as the amount of self-promotional advertising needed to compensate for the marginal attack dollar.

This can lead to very large numbers via small diversion ratios that appear in the denominator of the damage expressions. Still, some of the damage estimates (e.g. Advil on Tylenol) are very precisely estimated, and show that the damage is between 3 and 4 dollars for a marginal dollar of comparative advertising.

TABLE 6: Measures of Damage

	Target:					
Attacker:	Advil	Aleve	Bayer	Excedrin	Motrin	Tylenol
Advil		N/A 7.874 [3.752,17.015] N/A	0.044 [0.017,0.044] N/A N/A	N/A 9.694 [2.284,2.7e3] N/A	0.053 [0.024,0.080] N/A N/A	0.033 [0.004,0.059] 3.197 [2.094,4.835] 3.217 [2.112,4.853]
Aleve	0.049 [0.023,0.074] N/A N/A				0.056 [0.023,0.073] N/A N/A	0.068 [0.038,0.089] 3.457 [2.387,5.273] 3.499 [2.448,5.306]
Bayer	N/A 8.810 [4.265,23.786] N/A					0.191 [0.111,0.278] 8.358 [4.686,14.884] 8.475 [4.839,14.814]
Excedrin	0.040 [0.000,0.131] N/A N/A					N/A 14.011 [5.003,207.996] N/A
Motrin	N/A 7.284 [3.917,14.884] N/A	N/A 6.987 [3.708,15.787] N/A				
Tylenol	0.120 [0.080,0.150] 11.685 [4.839,71.037] 11.759 [4.925,71.093]	0.112 [0.075,0.140] 5.678 [3.238,9.906] 5.747 [3.347,9.962]	0.046 [0.027,0.067] 2.026 [1.113,3.552] 2.054 [1.170,3.569]	0.028 [0.002,0.061] N/A N/A		

Notes: A row-column entry denotes attacker-target $ damage from a marginal $1 comparative ad attack, split up from top down as push-up effect damage from attacker's self-promotion component of comparison; pull-down effect damage; and total damage as sum of these two. Bootstrapped 90% confidence intervals appear in square brackets underneath the point estimates.[25]

Second, the pull effect is much larger than the push effect (of self-promotion). For exam-

[25] Confidence intervals are based on 100 draws on the asymptotic distribution of the estimates from Column 6 of Table 4 and from Column 1 of Table 5.

ple, when Advil attacks Tylenol, Tylenol suffers a \$3.22 loss, but (marginal) self-promotion by Advil only causes a 3 cent loss. The pull effect is large because the target must be pulled down a lot in order to induce a brand to use comparative advertising, since the fall-out is shared among all other rivals (the size of the spillover is investigated below). This effect is exacerbated by the fact that the push effect of the comparative ad is only around half of what it would be with self-promotion.

TABLE 7. Spillover Effects

Attacker	Attacked	Advil	Bayer	Motrin	Tylenol
Advil	Aleve			0.404 [0.188,0.686]	0.517 [0.266,0.863]
Advil	Tylenol		0.120 [0.054,0.155]		
Aleve	Tylenol	0.387 [0.230,0.485]			
Bayer	Advil				0.172 [−0.036,0.432]
Excedrin	Tylenol	1.648 [0.560,13.916]			
Tylenol	Advil		0.484 [0.214,1.325]		
Tylenol	Aleve	0.202 [0.052,0.304]			

A row-column entry gives the dollar effect on the column brand of a \$1 increment in comparative advertising on the row link. Bootstrapped 90% confidence intervals are in square brackets below the point estimates. Confidence intervals are based on 100 draws on the asymptotic distribution of the estimates from Column 6 of Table 4 and from Column 1 of Table 5.

Third, the asymmetry between the Bayer-Tylenol numbers is striking. Tylenol needs \$8.48 to negate a marginal Bayer attack, but Bayer needs only \$2.05 to offset a marginal Tylenol attack. The difference between Aleve-Tylenol and Advil-Tylenol is striking for being in the opposite direction. For example, Aleve takes \$5.75 to negate a marginal Tylenol attack on it, whereas Tylenol needs \$3.50 to negate a marginal Aleve attack. These differences are explained by the fact that the main component of damage is the pull effect, given by (14) as $\frac{s_k}{s_j}\frac{(1-\lambda)}{d_{jk}}$. Differences in diversion ratios and market shares across pairs then explain the results. Because the diversion ratios from Tylenol to other brands are systematically smaller than in the opposite direction, the marginal effect of Tylenol attacks is larger on Advil and Aleve, whose shares are quite similar. However, the much smaller Bayer share reverses this.

Other brands are affected when brand j attacks brand k. First, the push-up effect on brand j hurts all other brands $l \neq j$, and the pull-down effect on brand k benefits all other brands $l \neq k$. The net effect (see (15) above) can a priori be positive or negative. For all our specifications, we find non-negligible, positive and statistically significant spillovers for all but one case (Bayer's attacks on Advil). These range from 12 to 52 cents for each dollar spent on a marginal attack, except for the outlier case of Excedrin on Tylenol, where Excedrin does very little comparative advertising and the estimates are unreliable due to the small number of observations of this target pair. Notice that the imprecise estimates in Table 5 feed through into imprecision in Table 7 (for example, Bayer vs. Advil). However, except for the outlier case of Excedrin against Tylenol, the intervals are smaller than those in Table 6. This is because the expression for spillover damage, (15), is written in terms of ratios of diversion ratios, whereas the damage to the target, (14), encompasses the reciprocal of a diversion ratio: small estimates of diversion ratios therefore give large damages and large confidence intervals.

Even though the results of Table 6 indicate much stronger pull-down effects on the target than push-up effects, the pull-down effect only benefits rivals to the extent that demand shed by the target is diverted to them.[26] But rivals are also harm by the attacker's push-up component of comparative advertising. Nonetheless, our results in Table 7 indicate that the net effect on other brands is positive. The positive spillovers on other brands are quite substantial. For example, a marginal comparative advertising dollar spent by Advil against Aleve benefits Motrin by 40 cents and Tylenol by 52 cents, while benefiting Advil by $1, and hurting Aleve by $7.87 (from Table 6). We are unable to estimate the spillovers on the other brands because we are unable to estimate the diversion ratios from those other brands to both target and attacker (Excedrin attacks neither, while Bayer does not attack Aleve). Indeed, estimating the spillover on l when j targets k requires estimates of the diversion ratios d_{jk}, d_{lk}, and d_{lj}. In turn, this requires there to be active attacks from j to k and l, and from l to k. Hence we cannot estimate any spillovers from Motrin attacks because no

[26] This dilution of pull-down is already reflected in the attacker's calculus: it only gets a fraction of the demand lost by its target.

brand attacks Motrin in return.

The estimates of spillovers indicate significant free-rider effects in comparative advertising, insofar as other brands are shown to benefit from comparative advertising (the harm from the push effect is dominated by the gains from the pull effect on the target). Lest this suggest that comparative advertising is insufficient (which it is if we exclude the target!), bear in mind that the costs to the target far outweigh the sum of benefits to attacker and other rivals. For example, a marginal dollar spent by Advil attacking Tylenol causes a \$3.22 loss to Tylenol (Table 6) and a 12 cent gain to Bayer (Table 7), and a \$1 benefit to Advil. The practice of comparative advertising causes far more loss in profit to the target (at the margin) than it recoups to the attacker and spills over to other rivals.[27] This, quite likely, explains why there are so few industries (in so few countries) where comparative advertising is used. Recognizing the mutual harm, companies refrain from attacks.

6 Conclusions

The paper models comparative advertising as having both a "push up" effect on own perceived quality, and a "pull-down" effect on a targeted rival's quality. The targeting of comparative advertising affords a unique opportunity for estimating diversion ratios between products solely from observed supply side comparative advertising expenditures. Diversion ratios are direct inputs into deriving estimates for the damage inflicted from comparative advertising and the spillover to other brands.

The empirical results for OTC analgesics indicate that push-up from a marginal comparative ad is about half the push-up from a marginal self-promotion ad. The benefit from pull-down is much smaller than the damage to the target, while conferring significant net benefits on other rivals. On net though, comparative advertising causes more harm to industry profit than benefit (and similar complaints are voiced about the destructive damage caused by negative political campaign ads). This is a likely reason why it remains quite

[27] Other ways of conceptualizing comparative advertising might soften this conclusion.

rare.[28]

The effects of advertising in the Push-Pull set-up are channeled through quality differences. This gives quite a negative view of comparative ads, in the sense that there is much wasteful battling to and fro between brands just to stay afloat. This feature is reminiscent of the critique of advertising that it serves solely to reshuffle demand and brands are better off if they could agree not to do it (they would save the expense). The critique is *a fortiori* true of comparative advertising.

Liaukonyte (2012) uses the same dataset as ours to estimate a random coefficient model of demand with the standard Berry, Levinsohn, and Pakes (1995) method. The analysis in her paper is carried at the product (e.g. Tylenol Cold) rather than at the brand level (e.g. Tylenol), which is what we consider here. We exploit different sources of exogenous variation. Liaukonyte (2012) uses variation in entry and exit of products, as in Bresnahan (1987). We use exogenous variation in news shocks and generic prices. While our approaches take quite different perspectives (demand vs. supply), it is striking that the marginal rate of substitution between self-promotion and comparative advertising is estimated to be between 0.5 and 1 in both papers.

Our conclusions from the push-pull set-up may also be consistent with alternative ways of thinking of how comparative advertising works. Indeed, when consumers have different tastes over different characteristics, then comparative advertising (done by different brands about different characteristics) can convey information to consumers. However, it should then be expected that comparative advertising contains information that the target would choose not to include in its own ads. It would then incur a loss in profit that could outweigh the benefits to other parties. Anderson and Renault (2009) show in such a setting that this theoretically possible. They find that comparative advertising, when it is used in equilibrium, may induce the target to decrease its price and the losses thus inflicted may be large enough to outweigh the benefits to the attacker and to consumers, so social surplus decreases. Even if such a detrimental welfare outcome does not arise, an informative advertising approach

[28] Comparative advertising is being used more recently (e.g., the “soup-wars” between Campbell’s and Progresso), coinciding with a recession, when quasi-collusion typically has more trouble surviving.

yields the same potential ambiguity as the push-pull set-up with regard to the desirability of comparative advertising for industry profit.

In future work, we hope to estimate a joint model of demand and supply. To do so will require taking a stand on the demand function to deploy. Such analysis will shed light on how much is added by a demand or supply side. It will also enable us to undertake an equilibrium analysis of the effects of comparative advertising, and, indeed, of advertising overall. That is, we could counterfactually ban comparative advertising and see how profits are affected. Currently, we only do this at the margin. Likewise, we could close down advertising completely to determine how constructive or destructive it is to profits. One fruitful avenue for future research is to estimate structural parameters for alternative formulations of how comparative advertising affects consumer choices.

We also to hope to further exploit the information in the dataset, which contains information on the topics of the attacks between brands and the characteristics of the products. A first step in this direction is taken by Anderson, Ciliberto, and Liaukonyte (2012), who employ a much simpler, non-strategic, theoretical framework than the one developed here to explain the number of information cues in ads.

7 References

Anderson, Simon P., André de Palma, and Jacques-François Thisse (1992), *Discrete Choice Theory of Product Differentiation.* Cambridge, MA: The MIT Press.

Anderson, Simon P. and Régis Renault (2009), Comparative Advertising: disclosing horizontal match information. *RAND Journal of Economics,* 40(3), 558-581.

Anderson, Simon. P., Federico Ciliberto, and Jura Liaukonyte (2012), Information Content of Advertising: Empirical Evidence from the OTC Analgesic Industry. *Working Paper*, University of Virginia.

Bagwell, Kyle (2007), The Economic Analysis of Advertising. In Mark Armstrong and Rob Porter (eds.) *Handbook of Industrial Organization*, 3, 1701-1844. Elsevier, Amsterdam, North Holland.

Barigozzi, Francesca, Paolo Garella, and Martin Peitz (2009), With a little help from my enemy: comparative vs. generic advertising. *Journal of Economics and Management Strategy*, 18(4), 1071–1094.

Becker, Gary S. and Kevin M. Murphy (1993), A Simple Theory of Advertising as a Good or Bad. *Quarterly Journal of Economics*, 108(4), 942–964.

Berry, Steven (1994), Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics,* 25(2), 242-262.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), Automobile Prices in Market Equilibrium. *Econometrica,* 63(4), 841-890.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman (2010), What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment. *Quarterly Journal of Economics,* 125(1), 263-305.

Blundell, Richard W. and Richard J. Smith (1986), An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply. *Econometrica*, 54(3), 679-685.

Bresnahan, Timothy F. (1987), Competition and Collusion in the American Automobile Market: The 1955 Price War. *Journal of Industrial Economics*, 35(4), 457-482.

Busse, Megan and Marc Rysman (2005), Competition and Price Discrimination in Yellow Pages Advertising. *RAND Journal of Economics*, 36(2), 378-390.

Ching, Andrew (2004), Some Observations in the U.S. Prescription Drug Market after Patent Expiration During the 80s. *Working paper,* Rotman School of Management, University of Toronto.

Ching, Andrew, C. Robert Clark, Ignatius J. Horstmann, and Hyunwoo Lim (2011), The Effects of Publicity on Demand: The Case of Anti-Cholesterol Drugs. *Working paper,* Rotman School of Management, University of Toronto.

Chintagunta, Pradeep K., Renna Jiang, and Ginger Zhe Jin (2009), Information, Learning, and Drug Diffusion: the Case of Cox-2 Inhibitors. *Quantitative Marketing and Economics*, 7(4), 399-443.

Dixit, Avinash K. and Victor D. Norman (1978), Advertising and Welfare. *Bell Journal of Economics*, 9(1), 1-17.

Emons, Winand and Claude Fluet (2011), Non-comparative versus comparative advertising as a quality signal. *CIRPEE Working Paper* No. 11-39.

Gasmi, Farid, Jean-Jacques Laffont, and Quang Vuong (1992), Econometric Analysis of Collusive Behavior in a Soft-Drink Market. *Journal of Economics & Management Strategy,* 1(2), 277–312.

Gourieroux, Christian, Alain Monfort, Eric Renault, and Alain Trognon (1987), Generalized Residuals. *Journal of Econometrics,* 34(1-2), 5–32.

Grabowski, Henry G. and John M. Vernon (1992), Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act. *Journal of Law and Economics*, 35(2), 331-350.

Heckman, James and Richard Robb (1985), Alternative methods for evaluating the impact of interventions: an overview. *Journal of Econometrics,* 30(1-2), 239–67.

Heckman, James and Richard Robb (1986), Alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes. In *Drawing Inferences from Self-Selected Samples*, ed. H. Wainer. New York: Springer. Repr. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.

Hyllested, M., S. Jones, J.L. Pedersen, and H. Kehlet (2002), Comparative effect of paracetamol, NSAIDs or their combination in postoperative pain management: a qualitative review. *British Journal of Anaesthesia*. 88(2), 199–214.

Jaffe, Sonia and Glen E. Weyl (2011), The First-Order Approach to Merger Analysis. *Working Paper,* Harvard University.

Johnson, Justin P. and David P. Myatt (2006), On the Simple Economics of Advertising, Marketing, and Product Design. *American Economic Review*, 96(3) 756-784.

Kihlstrom, Richard E. and Michael H. Riordan (1984), Advertising as a signal. *Journal of Political Economy*, 92(3), 427-450.

Liaukonyte, Jura (2012), Is comparative advertising an active ingredient in the market for pain relief? *Working Paper*, Cornell University.

Milgrom, Paul R., and John Roberts (1986), Prices and Advertising Signals of Product Quality. *Journal of Political Economy*, 94(4), 796-821.

Milyo, Jeffrey and Joel Waldfogel (1999), The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart. *American Economic Review*, 89(5),1081-1096.

Nelson, Phillip J. (1974), Advertising as information. *Journal of Political Economy*, 82(4), 729-754.

Rivers, Douglas and Quang H. Vuong (1988), Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. *Journal of Econometrics*, 39(3), 347–366.

Scott Morton, Fiona (2004), Horizontal Integration between Brand and Generic Firms in the Pharmaceutical Industry. *Journal of Economics and Management Strategy*, 11(1), 135-168.

Shapiro, Carl (1996), Mergers with Differentiated Products. *Antitrust*, 10(2), 23-30.

Sovinsky Goeree, Michelle (2008), Limited Information and Advertising in the US Personal Computer Industry. *Econometrica*, 76(5), 1017-1074.

Stigler, George J. and Gary Becker (1977), De Gustibus Non Est Disputandum. *American Economic Review*, 67(2), 76-90.