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

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

Incomplete and Preliminary Draft

JAN 2008: THIS VERSION: JUNE 2009

Abstract
We estimate the incentives to get ahead by hurting rivals in the context of comparative advertising. To do this, we watched all ads broadcast by the US OTC analgesics industry for a 5-year period and coded them according to which brands target which rival brands in comparisons. Data on how much was spent airing each ad then allows us to determine the dollar amounts spent in these attacks. We take these data to a structural model of targeting in which comparative advertising has a direct effect of pushing up own brand perception along with pulling down the brand images of targeted rivals. Brands’ optimal choices of advertising mix yield simple oligopoly equilibrium relations between advertising levels (for different types of advertising) and market shares. These we estimate by using as instruments the prices of equivalent generic drugs; and we use medical news shocks as further explanatory variables. We estimate that each dollar spent on comparative advertising has the same direct effect as 50 cents spent on non-comparative (purely direct) advertising: the remainder is attributable to pulling down rivals. There is strong evidence of damage to targets: each $1 spent against a target needs at least 40 cents to rectify.

Keywords: Comparative Advertising, persuasive advertising, targeted advertising, analgesics.

JEL Classification

1 Introduction

The economic analysis of comparative advertising offers a unique window into firms’ incentives to push themselves up and to pull their rivals down.1 Comparative advertising can do this by promoting one’s own product while benefiting from the fall-out from denigrating a rival product.2 Since the marketing mix can include purely direct advertising (that is, purely positive, self-promoting, advertising, which we henceforth refer to as non-comparative), we can untangle empirically the push and the pull effects. Moreover,

∗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. Catherine de Fontenay and Joshua Gans provided useful comments, as did participants at the Summer Workshop in Industrial Organization (Auckland, 2009). We thank Melbourne Business School and the Portuguese Competition Authority for their hospitality.
†Department of Economics, University of Virginia, Charlottesville VA 22904, USA. sa9w@virginia.edu
‡Department of Economics, University of Virginia, Charlottesville VA 22904, USA. ciliberto@virginia.edu
§Department of Economics, University of Virginia, Charlottesville VA 22904, USA; and Cornell AEM. jura@virginia.edu
¶Université de Cergy-Pontoise, ThÉMA, 33 Bd du Port, F95000 Cergy-Pontoise, France. regis.renault@eco.u-cergy.fr
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.
2 We discuss competing viewpoints on comparative advertising below.
comparative advertising can target particular rivals, and so we can determine whether large or small firms will be pulled down most by large or small rivals.

To do this requires first of all finding out how much is spent on comparative advertising. For coding reasons discussed below, a cross section study across industries is clearly infeasible, and so we need to analyze a particular industry. This is not a simple matter because advertising spending by firms, even when the data are available (which is already rare), is not broken down into comparative and non-comparative advertising. We must therefore look at each individual ad and determine whether or not it is comparative, and, if so, which is the target brand. This therefore requires a detailed coding of advertising content. Ideally, we should be able to analyze an industry for which comparative advertising is prevalent and represents a large fraction of industry sales, for which data on spending on ads is available for a full sample of firms and for a reasonably long period of time. Furthermore, video files (or audio files for radio ads or photographic files for newspaper/magazine ads) need to be available and their content readily coded for the desired information of comparison and targets. Fortunately, all these criteria are met with the Over-The-Counter (OTC) analgesics industry in the US.3

Non-comparative advertising involves only positive promotion. A comparative advertisement, by comparing one’s own product in favorable light relative to a rival, has both a positive promotion component (in common with non-comparative advertising) and an indirect effect through denigrating a rival. Denigration can be per se advantageous insofar as consumers who switch from the demeaned product are picked up by the denigrating firm. However, they may also be picked up by other rival firms. This logic indicates a possible free-rider situation in the provision of comparative advertising against any particular rival, but it also indicates an equilibrium at which each firm’s positive promotion (through both comparative and non-comparative channels) is devalued by others’ comparative advertising.

Our aim is to untangle these two effects in a structural model of firms’ allocations of advertising expenditures by determining just what extent of comparative advertising is pushing oneself up and how much is pulling down a rival. The push-pull model is based on a discrete choice approach to demand, in which firms’ perceived qualities are shifted by advertising. Promoting one’s own product increases demand directly, whether through non-comparative advertising or comparative advertising, while denigrating a rival helps a firm indirectly by decreasing perceived rival quality.4 By hurting the rival product directly, some consumers are diverted, and the comparative advertiser succeeds in attracting some portion of those consumers.

3 Indeed, while explicit comparative advertising has flourished in the United States over the past 20 years (with the blessing of the FTC), its prevalence varies widely across industries. The US OTC analgesics industry (basically, medicine for minor pain relief, involving as major brands Advil, Aleve, Bayer Aspirin, and Tylenol) exhibits high advertising levels in general, and extraordinary levels of explicit comparative claims on relative performance of drugs. Most of the advertising expenditures are for television ads.

4 A somewhat similar approach is expounded in Harrington and Hess (1996). These authors treat positive and negative advertising by 2 politicians with given locations in a policy space. Negative advertising shifts a rival candidate away from the median voter, while positive advertising shifts a candidate closer. This framework would indeed provide an interesting base to develop a product market model.
We develop a novel (and simple) model, which gives clean relations for estimation. To begin with, the way in which advertising enters the model is most simply thought of as persuasive advertising that shifts demand up.\(^5\) Then, we introduce comparative advertising into the equilibrium marketing mix. Next, we use a logit approach, wherein comparative advertising pulls down the perceived quality of targeted rivals’ products. We use our simple model to use the equilibrium pricing (first-order) conditions to eliminate prices from the relation between advertising and sales.\(^6\) We can then relate ad levels of the different ad types to other observable market variables, like market shares.\(^7\) Finally, we use functional form assumptions to be able to identify and estimate the structural parameters of the utility function that are related to non-comparative and comparative advertising.

To execute our empirical analysis we use data on national sales from AC Nielsen and advertising data on advertising expenditure (and movies) from TNS - Media Intelligence, respectively. To gather data on the precise extent of the practice of comparative advertising we actually watched more than four thousands ads and coded them by content. We then match the result to advertising expenditures data.

In our regression analysis we consider two sources of exogenous variation. First, we use data on the prices of the generic products to construct measures of the marginal costs that firms face to produce the corresponding branded product. Here, the generic price of a pill of Acetaminophen is used as an instrumental variable of the share of Tylenol, whose main active ingredient is Acetaminophen. Thus, the prices of the generic products are the variables that are excluded from the utility function and that we use as instrumental variables in the estimation. Second, we construct a dataset of news shock that hit the OTC analgesic markets in the time period of analysis.\(^8\) These shocks might interact with the advertising decisions, and thus we

\(^5\)This is, for example, consistent with “hype” in the Johnson and Myatt (2004) taxonomy of demand shifts. We can though also reconcile our formulation with other advertising types. Most simply, the formulation is consistent with complementary advertising of the type propounded by Stigler-Becker (1977) and Becker and Murphy (1993). Indeed, one can readily append advertising in the standard discrete choice approach underpinning to the logit demand, as we present below. Alternatively, it is easy to formulate a representative consumer utility function to underlie the demand model, along the lines of Anderson, de Palma, and Thisse (1988), and introduce advertising into it.

\(^6\)One advantage of this approach is that we bypass having to deal with price data, which involves multiple price points for multiple variants of the same brand, along with various other problems associated to price data. See the Appendix for more on this.

These variables are in turn determined simultaneously in a market equilibrium game between profit maximizing firms. Firms with a lot of advertising are also typically those with large market shares. They also tend to set high prices. This is of course not to say that high prices drive high market shares, nor, more subtly, that advertising creates high prices, nor indeed is it the high prices that create the desire to advertise. All of these variables are jointly determined, at a market equilibrium, and we show how they are determined within an industry from the firms’ equilibrium choices. What drives the results is the intrinsic brand “qualities” (fixed effects) and the marginal efficiency of advertising types across firms. See Anderson and de Palma (2001) for an analysis of how qualities correlate with market shares and prices, in a context without advertising. Here, with advertising in the choice set, and interacting with quality parameters, the results are more nuanced, though we still find some strong relations between market shares and advertising of various types.

\(^7\)As we discuss later on, we follow an approach similar to Chintagunta, Jiang and Jin (2007) when constructing our dataset of news shocks. In particular, between 2001 and 2005, the OTC analgesics market endured several major medical news related “shocks”. The most notable, but by no means the only ones, of these were the following. The withdrawals of the Prescription NSAIDs Vioxx (October, 2004) and Bextra (April, 2005) affected the OTC NSAIDs market (which excludes Tylenol). Naproxen sodium, the active ingredient in Aleve was linked to increased cardiovascular risk, which led to a significant sales decrease for Aleve (December, 2004). The main idea here is that these shocks act as many natural experiments. The idea of using a natural experiment to study the effect of advertising (on prices) is the crucial incisight in Milyo and Waldfogel [1999].
cannot use them straight up as instrumental variables. However, adding these news shocks improves our empirical analysis dramatically. In addition we can interact these shocks with the price of the generic products and increase the number of instrumental variables that we use.

The main results are the following. We estimate that each dollar spent on comparative advertising has the same direct effect as 50 cents spent on non-comparative (purely direct) advertising: the remainder is attributable to pulling down rivals. We also find that there is strong evidence of damage to targets: each $1 spent against a target needs at least 40 cents to rectify. We find evidence that comparative advertising is increasing in the shares of the attacked firm and in the shares of the attacking firm. This result has a nice and simple interpretation: the return to attacking a large firm is higher than the return to attacking a smaller firms, since by attacking a larger firm, the attacker can hope that a larger pool of consumers switch away from the attacked to the attacker. Similarly, a large firm has a stronger incentive than a smaller firm to attack because the probability that consumer switch to the larger firm is higher than the probability that consumers switch to the smaller firm.

The paper is organized as follows. In the next section we review the literature. Section 3 presents the theoretical model. Data and industry background are discussed in Section 4. We present the empirical specification and discuss identification of the model in Sections 5 and 6. Section 7 discusses results and Section 8 concludes.

2 Literature Review on Advertising

A lot of the economics literature on the economics of advertising has been concerned with the functions of advertising, and whether market provision is optimal. We here take more of a marketer’s stance that advertising clearly improves demand (otherwise firms would not do it), and we take a rather agnostic view of how it is the advertising actually works on individuals, and bundle it all into a single “persuasive” dimension. Since we do not cover here the normative economics of the advertising, this is excusable. The innovations we pursue are in advertising competition, and in the new strategic direction of comparative advertising.

2.1 Theoretical Literature

Much of the economic theory of advertising has been concerned with the mechanism by which advertising affects choice, and the welfare economics of the market outcome.\textsuperscript{9} Moreover, much work has considered very particular market structures, most often monopoly.\textsuperscript{10}

\textsuperscript{9}See Bagwell (2009) for a comprehensive survey.
\textsuperscript{10}Almost all the signaling literature considers monopoly, with the notable exception of Fluet and Garella (2002) who consider a duopoly. The classic Butters (1977) model of informative advertising considers monopolistic competition and a homogenous good with zero profits sent on each message. Grossman and Shapiro (1984) allow for oligopoly and product differentiation (around a circle), but they use symmetry assumptions liberally.
Persuasive Advertising. Much of the early work linked advertising to market power, and reached a fairly negative assessment that advertising is a wasteful form of competition. Kaldor (1950) and Galbraith (1958) saw the differentiation achieved by advertising as spurious and artificially created by persuasion. Such persuasive advertising was thought to decrease social welfare by deterring potential competition and creating barriers for new entrants. Dixit and Norman (1978), propose viewing persuasive advertising as shifting demand curves out, but they then take an agnostic view as to the welfare effects of the shift (i.e., whether the demand curve before or after the advertising is a better representation of the true consumer benefit from consuming the good).\(^1\) Regardless, they suggest that there is a tendency for too much advertising.

Informative Advertising. The persuasive view and the idea that advertising fosters monopoly was first challenged by Telser (1964) who argued that advertising can actually increase competition through improving consumer information about products (see also Demsetz (1979)).\(^2\) Butters (1977) later formalized a monopolistically competitive model of informative advertising about prices, in which the level of advertising reach is socially optimal. These results were tempered somewhat by Grossman and Shapiro (1984), who extended the advertising content to include (horizontal) product differentiation.\(^3\)

Another informative role, albeit indirect information, is at the heart of “money-burning” models of signaling product quality. Nelson (1970, 1974) claims that advertising serves as a signal of quality, especially in experience good markets, and reasons that consumers will rationally conclude that a firm doing a lot of advertising must be selling a product of high quality. These insights were later formalized and further developed, most frequently by using repeat purchases as the mechanism by which a high-quality firm recoups its advertising investment.\(^4\) Kihlstrom and Riordan (1984) show a role for dissipative advertising in a perfectly competitive model. Milgrom and Roberts (1986) break out different roles for signaling quality through (low) price and through advertising by a monopoly, again using a repeat purchase mechanism. Fluet and Garella (2002) show that under duopoly there must always be dissipative advertising by the high quality firm if qualities are similar enough.

Advertising as a Complementary Good. Another foundational role for advertising is proposed by Stigler and Becker (1977) and Becker and Murphy (1993), who argue that advertising can be viewed as part of consumers’ preferences in the same way as goods directly enter utility functions, and that there are complementarities between advertising levels and goods’ consumption. Hence, ceteris paribus, willingness to pay is higher the more a good is advertised. The complementary goods approach affords one clean

---

\(^1\)This analysis is not uncontroversial: see the subsequent issues of the RAND journal for comments, replies, and rejoinders. Dixit and Norman (1978) posited that advertising increases demand, and then perform the welfare analysis using consumer surplus measures from that starting point, according to which demand curve embodies “true” tastes.

\(^2\)Indeed, informative advertising can reduce consumers’ search costs to learn about the existence of products, their prices, qualities, and specifications.


\(^4\)Another mechanism is to suppose some consumers are informed already, so a low-quality firm has to distort its price so high to mimic the high-quality one that it does not wish to do so.
way for advertising to affect directly consumer well-being, and so gives a way of thinking about persuasive advertising.

The specification we use in our model is most directly interpreted in this vein of complementary goods, insofar as we can interpret that advertising expenditures as boosting demand. However, since we will not be doing a welfare analysis with the model, we are not constrained to this interpretation, but instead our approach is broadly consistent with advertising as a demand shifter (as in Dixit and Norman (1978)).

2.2 Theoretical Economics Comparative Advertising Literature: Modeling Comparative Advertising

The theoretical economics literature on comparative advertising is quite scarce. Modeling comparative advertising presents several alternative potential approaches. In common with much of the economics of advertising, these are perhaps complementary rather than substitute approaches, and elements of each are likely present (in different strengths) in different applications. Each though has drawbacks, and sometimes the predictions (e.g., comparative static properties) differ in direction.

One early contribution is Shy (1995), who argues that comparative advertising of differentiated products informs consumers about the difference between the brand they have purchased in the past and their ideal brand. The model explains only brand switching behavior, because according to that setting comparative advertising is meaningless for the inexperienced consumer as she would not be able to comprehend an ad involving a comparison of the brands’ attributes that she never consumed. Aluf and Shy (2001) model comparative advertising using a Hotelling-type model of product differentiation as shifting the transport cost to the rival’s product.

**Horizontal Match.** Anderson and Renault (2009) model advertising as purely and directly informative revelation of horizontal match characteristics of products.\(^{15}\) Revelation of such information increases product differentiation, although this does not always increase firm profits. Comparative advertising in this context is modeled as revelation of characteristics (match information) of the rival product along with own characteristics. One key finding is that (under duopoly) comparative advertising is carried out by the smaller firm against its larger rival, and arises if firms are different enough.\(^{16}\)

It is not immediately evident how these results extend to more firms, except insofar as an industry of

---

\(^{15}\)That paper builds on Anderson and Renault (2006), who show that a monopoly firm might limit information about its product attributes even if advertising has no cost. This result identifies situations where a firm is hurt by information disclosure about its own product, so there might be incentives for competitors to provide that information through comparative ads.

\(^{16}\)To understand the incentives to advertise requires understanding the benefits of more information on each firm’s profits. With no information at all, firms are homogenous apart from the quality advantage, and the large firm can price out its advantage and still serve the whole market. It has no incentive to advertise because, while such advertising will raise the willingness to pay of consumers who discover they appreciate its product, it will also decrease the valuations of those who discover they like the product less than average, and so the firm will lose customers to its rival as well as having to price lower to staunch the loss of consumer base. This means that the large firm does not want to advertise, while the smaller rival does. These incentives extend to comparative advertising, which further enhances differentiation and further erodes the customer base (and price) of the larger firm to the advantage of the smaller one.
roughly similar size firms would be expected to not deploy comparative advertising since individual incentives to broadcast own information should suffice. Otherwise, with firms of different sizes, there is a free-rider aspect to comparative advertising, that others (apart from the target) might benefit from it. A medium size firm might benefit from advertising relative to a large rival, but might lose relative to smaller ones. Small ones might have little to gain if indeed their small size stems from inherent disadvantages. However, it is not easy to introduce multiple firms in this context of asymmetric information divulging and hence asymmetric product differentiation.

The present model also relates the pattern of ads to market shares, but it treats the role of advertising differently. We do not model the informational content of the advertisement. Empirically we are unable to separate whether advertising was persuasive or informative, so we remain agnostic about the advertising effects and focus just on separation of comparative and non-comparative ads.

It is also important to note that the role of advertising in the Anderson-Renault (2009) model is only to divulge horizontal match information, which is two-edged sword – what characteristics one consumer likes, another dislikes. The analysis is phrased in terms of informing all consumers: it does not allow for advertising reach that tells only some. The same critique can be leveled at other models in the field, as well as (perhaps to a lesser degree) the model we actually propose here; and we return to this criticism in the conclusions.

**Signaling.** Another approach to modeling comparative advertising takes as staging point the signaling model of advertising, which goes back to insights in Nelson and was formalized in Milgrom and Roberts (1986). The original theory views advertising as "money-burning" expenditure which separates out low-quality from high quality producers. Equilibrium advertising spending, in this adverse-selection context, smokes out the low type because a low-type would never recuperate in repeat purchases the high level of spending indicated in equilibrium. The comparative advertising version of this theory expounded in Barigozzi, Garella, and Peitz (2006) relies on the possibility of a law-suit to punish an untrue claim. Recently, Emons and Fluet (2008) also took a signaling approach to comparative advertising, although their analysis relies on advertising being more costly the more extreme are the claims it makes, instead of a law-suit.

**Persuasion Games.** In parallel work, we are developing another approach along the lines of the Persuasion Game of Milgrom (1981) and Grossman (1981). In this work the firms must (truthfully) announce levels of product characteristics their products embody. Comparative advertising, through this lens, involves announcing characteristics levels of rivals that those rivals would prefer to keep silent. However, the actual ads are quite vague for the most part in specifics of actual claims (e.g., a product may act "faster" than another, but it is not usually specified how much faster, or indeed what the response time in minutes is for the two products or the statistical significance of the difference across different individuals, etc.)
2.3 Empirical Literature

In this Section we discuss the papers that are most closely related to ours and discuss the original contributions of our paper.\(^{17}\) To do this, we identify four modeling choices that have to be made when empirically studying advertising: how to measure advertising; whether to use a static or a dynamic model of advertising; whether to have a partial or a full equilibrium model, where both consumer and firm sides of the market are explicitly modeled; and whether to model advertising as having only a persuasive or informative effect, or both. Next, we discuss how the literature has dealt with these choices.

Advertising Content. Ours is the first paper to code the content of advertising into non-comparative and comparative ads and use the information in a structural model designed to address the incentives to use the different types of advertising.\(^{18}\) Previous papers have used total ad expenditures as the sole advertising explanatory variable (notable examples are Nevo [2000,2001] and Goeree [2008]). Here, because we have data on content, we break down the ad expenditures into comparative and non-comparative expenditures, and the comparative expenditures are further broken down into attacker-target pairs. We then look at the first order conditions of the advertising decisions, and so estimate the choice of advertising of the different types from the supply side. In related work with the same data, Liaukonyte (2009) estimates a model of demand where non-comparative and comparative advertising are found to have different quantitative effects on consumer choices.

Dynamic vs. Static and Partial vs. Full Equilibrium Models. We estimate a static model of firm behavior, where firms jointly choose product prices and advertising levels. We consider a full equilibrium static model of the advertising and product markets, where advertising is determined endogenously within the model. We use the first order conditions and demand equations for the product (analgesics) to solve the prices out of the first order conditions for advertising. This procedure yields simple relations between ad levels and market shares, which we term "quasi-reaction functions" (they are not the full reaction functions because they still include market shares, which in turn depend on all prices and all advertising). We estimate the structural parameters of the model from these advertising first order conditions.

Because advertising is likely to have long-run effects on demand, the decision to use a static model to study advertising needs to be carefully justified. This modeling decision is tightly linked to another one: whether or not to have a full equilibrium model of the advertising and product markets. In short, estimating a fully dynamic equilibrium model even of just the product market is beyond what is feasible at this stage.

\(^{17}\) For more detail on the broader findings of the literature, see Bagwell's [2007] superb review of the empirical literature on advertising.

\(^{18}\) Contemporaneous and independent work by Crawford and Molnar (2009) looks at advertising content of TV ads for Hungarian mobile telephony. They estimate a demand model, in the same fashion as Liaukonyte [2009]. Anderson and Renault (2008) study newspaper ads for airlines, and they code their content. In the former case, only 5% of ads are comparative, and even fewer in the latter case. For Hungarian telephones, much of the advertising concerns prices; in analgesics, virtually none. For airlines, mainly the low-cost carriers emphasize prices.
of the literature. Previous work in advertising has either estimated a dynamic model of demand (Hendel and Nevo [2006] and Gowrisankaran and Rysman [2009]) or has looked at a static model of demand and a dynamic model of supply (Roberts and Samuelson [1988], Dube, Hitsch, Manchanda [2005]).

Thus, a practical choice must be made. Either one models only one side of the market in a dynamic setting and must relinquish analyzing a full equilibrium model. Or else one can analyze a full equilibrium static model. In this paper we follow the second option. Clearly, these two approaches are complementary and provide different insights into the role of advertising. Most importantly, a static model simplifies the treatment of advertising as an endogenous variable. To our knowledge, all papers that study advertising in a dynamic context treat it as an exogenous variable (notable examples are Erdem and Keane [1996], Ackerberg [2001,2003] and Dube, Hitsch, Manchanda [2005]).

**Persuasive vs. Informative Advertising.** The last modeling choice is about the way that advertising affects consumer choice. Ideally, one would like advertising to have both an informative and persuasive effect. The informative effect has been modeled using a Bayesian learning model (Erdem and Keane [1996], Ackerberg [2003]), a limited consumer information model based on information sets (Goeree [2008]), or horizontal match information models (Anderson and Renault [2008] and Anand and Shachar [2004]). The persuasive effect is easier to model, as advertising is simply introduced into the utility function (e.g. Nevo [2001], Shum [2004]). There are only two papers that allow for both effects to be present, both by Ackerberg [2001,2003].

In order to identify the persuasive from the informative role, Ackerberg [2001,2003] analyzes consumer reactions to the advertising of a new product (the yogurt Yoplait 150). Essentially, advertising is only informative for first buyers, while it is both informative and persuasive for repeat buyers. This is a clever identification device, but we cannot use it here because we have aggregate and not individual data (that is, we cannot identify first buyers).

Our Push-Pull perspective on advertising is coherent with the persuasive view. In addition to positive persuasion on own quality, comparative advertising also gives negative persuasion on rivals.

---

19 The problem is both computational complexity and multiplicity of solutions. One would have to solve for rational and consistent expectations that consumers and producers have on the future values of the state variables, which means solving for a fixed point. There might be multiple future values of the state values for which such consistency requirements hold (that is, there might be multiple equilibria).

20 Although the latter paper presents a dynamic theoretical model of advertising, the econometric study estimates only the demand side parameters. These estimates are then used to calibrate the theoretical dynamic model.

21 Gasmi, Laffont, and Vuong [1992], Kadiyali [1996], and Slade [1995] postulate a set of residual demand functions, which include advertising. Thus, the interpretation of the role of advertising as persuasive or informative is not transparent.

22 Ackerberg (2001, 2003) argues that the observed facts that “experienced” consumers (those who have previously bought Yoplait 150) are much less sensitive to advertising than inexperienced ones is strong evidence in favor of advertising fulfilling an informative role rather than a “prestige” one. However, he does not control for the content of the particular ads in his sample; nor does he allow for the possibility (in his interpretation) that advertising ‘prestige’ could exhibit strong threshold effects, which could also account for the observed behavior.

23 This identification assumption excludes the possibility that a first buyer of a new product might have consumed other products of the same brand in the past, otherwise it is unclear that there is no persuasion effect for that type of buyer. Thus, while very clever, this assumption might not hold in practice.
Review of Similar Models of Advertising. We conclude this Section with a review of the three papers which deploy models of price and advertising competition that are close to ours.24

Gasmi, Laffont, and Vuong [1992] propose an empirical methodology for studying various types of collusive behavior in pricing and advertising. They derive two first order conditions (for prices and advertising) and one demand equation (for the product market, cola) for each firm and estimate them all jointly.25

Roberts and Samuelson [1988] estimate a model where demand is modeled statically, while supply is modeled dynamically. By assuming that firms have perfect foresight of future input prices, Roberts and Samuelson end up estimating a set of first order conditions for prices and advertising, as well as demand equations. Thus, even if they start from a dynamic supply model, in practice the system of equations they estimate is quite similar to the one considered by Gasmi, Laffont, and Vuong [1992].

Goeree [2008] considers a discrete choice consumer model under limited information, where advertising influences the set of products from which consumers choose to purchase, but does not enter into the utility function. She derives first order conditions for advertising and prices, as well demand functions for the products (computers).

In many ways our approach is similar to the ones used in these three papers. We also use a theoretical model to derive the first order conditions for prices and advertising. There are, however, important differences between our work and theirs. The main methodological differences are related to how we code advertising content, how we model demand, the nature of the exogenous variation that we use to identify the model, and how we estimate the parameters of the model.

First, all three look at the total advertising expenditure, while we distinguish between comparative and non-comparative advertising expenditures.

Second, our demand (as well as Goeree’s [2008]) is derived from a discrete choice model, while Gasmi, Laffont, and Vuong [1992] and Roberts and Samuelson [1988] postulate a set of residual demand functions. We have in common with Roberts and Samuelson [1988] a market expansion effect and a share effect, although we do not have the possibility that rivals’ demands can rise with own advertising.

Third, we use a combination of exogenous shocks and firm-specific generic prices to construct sources of exogenous variation in the data. Instead, Gasmi, Laffont, and Vuong [1992] use aggregate variables (e.g. the price of sugar). Roberts and Samuelson [1988]) use aggregate variables (e.g. cost of capital) and the number of own and rival brands.26 Goeree [2008] uses type of instrumental variables introduced by Bresnahan

24 Other papers (e.g. Shum [2004] or Nevo [2000,2001]) that use static models assume that advertising is exogenous, though they justify that assumption in their contexts. Clearly, these papers do not include first order conditions for advertising.
25 Kadiyali [1996] proposes an empirical methodology to investigate strategic entry and deterrence, where firms compete in prices and advertising. Since she closely follows Gasmi, Laffont, and Vuong [1992], the methodological differences between her paper and ours are the same as those between our paper and Gasmi, Laffont, and Vuong [1992].
26 The numbers of own and rival brands are valid instruments as long as these numbers are determined prior to price and advertising choices. This type of instrument was first proposed by Bresnahan [1987] and has been widely used since. We cannot use such brand numbers, since these are constant over the time period.
[1987]: the characteristics of the products produced by the competitors. Because we look at brands and not products, such instrumental variables cannot be used in an obvious way (brands have many differentiated products). We will discuss this more in Section (8).

Finally, our estimation methodology is different from those in the other papers. While they estimate a full set of simultaneous equations, we use the first order conditions for prices to solve the prices out of the advertising first order conditions. Thus, we fully exploit the theoretical model in the same way that they do, but we reduce the number of equations to be estimated. If the model is correctly specified (which is the maintained assumption in their studies, as in ours), then the estimation results should be the same under the two approaches.27

3 The Model

The theoretical model suggests certain regularities between market shares and both non-comparative and comparative advertising. Notice that the predictions for non-comparative advertising hold without the more specific functional form restrictions imposed later for the comparative advertising case. These size-advertising relations therefore hold in more general settings and also even when there is no comparative advertising, and so they constitute a contribution to the understanding of the size-advertising relation which is broader than the particular comparative advertising application developed in the sequel.

We first describe the demand side assumptions and then we derive the equilibrium predictions from the model. These take the form of advertising intensities as a function of market shares, and they form the basis of the estimation which follows. As we will see, the key predictions are all supported by the data.

We assume that each product is associated to a quality index and demand depends on the quality indices of all firms, in a manner familiar from, and standard in discrete choice analysis. These quality indices are influenced positively by own advertising (both non-comparative and comparative) and negatively by competitors’ comparative advertising. They are also influenced by medical news shocks which unexpectedly indicate good news or bad news about the health effects of the product(s).

3.1 Demand

Suppose that Firm $j = 1, \ldots, n$ charges price $p_j$ and has perceived quality $Q_j(\cdot), j = 1, \ldots, n$. We retain the subscript $j$ on $Q_j(\cdot)$ because when we get to the econometrics, 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.

Firms can increase own perceived quality through both types of advertising, and degrade competitors’

---

27In our small sample, estimating the full model would likely to lead to more precise estimates. We leave the estimation of the full model to future work.
quality through comparative advertising. Comparative advertising, by its very nature of comparing, both raises own perceived quality and reduces the perceived quality of rival products. The corresponding arguments of $Q_j(\cdot)$ are advertising expenditure by Firm $j$ which directly promotes its own product, denoted by $A_{jj}$; “outgoing” advertising by Firm $j$ targeted against Firm $k$, $A_{jk}$, $k \neq j$, which has a direct positive effect; and “incoming” comparative advertising by Firm $k$ targeting Firm $j$, $A_{kj}$, $k \neq j$, which has a negative (detraction) effect on Firm $j$’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.\(^{28}\)

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 individual choice data; we simply use it to frame the structure of the demand system. Preferences are described by a (conditional indirect) utility function:

$$U_j = \delta_j + \mu \varepsilon_j, \quad j = 0, 1, ..., n,$$

in standard fashion, where

$$\delta_j = Q_j(\cdot) - p_j$$

is the “objective” utility, and where we let the “outside option” (of not buying a painkiller) be associated to an objective utility $\delta_0 = V_0$. The parameter $\mu$ expresses the degree of horizontal consumer/product heterogeneity.\(^{29}\)

The structure of the random term determines the form of the corresponding demand function. At first, we do not impose further structure, but we later specialize (for the comparative advertising analysis) to the logit model to get a sharper set of benchmark properties. The corresponding market shares are denoted $s_j$, $j = 0, ..., n$, and each $s_j$ is increasing in its own objective utility, and decreasing in rivals’ objective utilities.\(^{30}\)

Assume that there are $M$ consumers in the market, so that the total demand for product $j$ will be $Ms_j$, $j = 0, ..., n$.

### 3.2 Profits

Assume that product $j$ is produced by Firm $j$ at constant marginal cost, $c_j$.

Firm $j$’s profit-maximizing problem is:

\(^{28}\)Throughout, we assume sufficient concavity that the relevant second order conditions hold.

\(^{29}\)As in Anderson, de Palma, and Thisse (1992). This parameter is especially needed whenever we specialize the model to the multinomial logit. Note that econometric specifications often set a marginal utility of money parameter (often $\alpha$) before the price term, and they normalize $\mu = 1$. This is therefore effectively setting $\alpha = 1/\mu$: we do not do this here because we shall shortly substitute out price term anyway, and the intuitions are cleaner without carrying around this $\alpha$.

\(^{30}\)For example, in the standard logit model, we have $s_j = \frac{\exp[\delta_j/\mu]}{\sum_{k=0}^{n} \exp[\delta_k/\mu]}$, $j = 0, ..., n$. 

12
Max \pi_j = M(p_j - c_j)s_j - A_{jj} - \gamma \sum_{k \neq j} A_{jk} \quad j = 1, \ldots, n. \quad (3)

Here \gamma > 1 reflects that comparative advertising may be intrinsically more costly because of the risk involved that a competitor might challenge the ad and it will have to be withdrawn and replaced with a less suitable one.\(^{31}\)

The advertising quantities (the \(A\)'s) are dollar expenditures.\(^{32}\) The idea is that advertising expenditures will be optimally allocated across media (and times of day in the case of radio/TV). Then market prices for access to eyeballs (and eyeballs of different value to advertisers) should embody the condition that there should be no systematically better/cheaper way to reach viewers. The strong form of this (efficient markets) hypothesis implicitly assumes that there are enough advertiser types, and there is no great difference in the values of consumers to OTC analgesics advertising compared to other sectors.\(^{33}\)

We assume in what follows that pricing and advertising levels are determined simultaneously in a Nash equilibrium.

### 3.3 Firms’ Optimal Choices

#### 3.3.1 Pricing

Recalling that shares, \(s_j\), depend on all the \(\delta\)'s, the price condition is determined in the standard manner by:

\[
\frac{d\pi_j}{dp_j} = Ms_j - M(p_j - c_j) \frac{ds_j}{d\delta_j} = 0, \quad j = 1, \ldots, n,
\]

which yields a solution \(p_j > c_j\): firms always select strictly positive mark-ups.

\(^{31}\)Hosp (2007) from Goodwin Procter LLP notes that “Comparative advertising is a useful tool to promote an advertiser’s goods and to tout the superior quality of the advertiser’s goods over those of its competitors. Comparative advertising, however, is also the form of advertising that is most likely to lead to disputes. In undertaking comparative advertising a company should be cognizant of the potential risks and pitfalls that can lead to costly disputes and litigation. The competitor will scrutinize the advertising, and is more likely to be willing to bear the expense of litigation or dispute resolution in an instance where the competitor itself has been targeted.”

More formally, suppose that a comparative ad is successfully challenged with probability \(P\), and that when withdrawn it must be replaced with an ad of lower effectiveness, and the effectiveness is a fraction \(\beta\) of that of the preferred ad. Let \(p^A\) per denote the cost of airing a non-comparative (on a particular channel at a particular time). Then the cost of airing the comparative ad is \(p^A ((1-s_j) + s_j/\beta)\). If we normalize the cost of the non-comparative advertising by setting \(p^A = 1\), then we have the effective comparative ad cost as \(\gamma = ((1-s_j) + s_j/\beta) > 1\).

\(^{32}\)They therefore need to be deflated by an advertising price index: as long as the price per viewer reached has not changed in a manner systematically different from the general inflation rate, the CPI is a decent proxy, and will be used below.

\(^{33}\)For example, suppose that each ad aired at a particular time on a particular channel cost \(\bar{p}\) and delivered \(H\) “hits.” (where the hit is measured in dollars). Then the equilibrium price of an ad delivering \(H/2\) hits should be \(\bar{p}/2\), etc.: the price per hit ought to be the same. Factoring in hits of different worth (the audience composition factor) follows similar lines. Notice though that such arbitrage arguments require sufficient homogeneity in valuations of at least some sub-set of advertising agents. The second caveat is that the arbitrage argument most directly applies to numbers of viewers hit, whereas here we deploy a demand form where ads enter a representative utility. It remains to be seen how consistent this is with an approach where heterogeneous individuals (who see different numbers of ads) are aggregated up to give a market demand function (see for example Goeree (2008) for an empirical application, albeit in the context of informative ads / consideration sets).
3.3.2 Non-Comparative Advertising

The following analysis covers persuasive advertising generally, and is not confined to the specifics of the comparative advertising approach which follows.

Non-comparative advertising expenditures are determined by:

\[
\frac{d\pi_j}{dA_{jj}} = \frac{d\pi_j}{d\delta_j} \frac{\partial Q_j}{\partial A_{jj}} - 1 = M(p_j - c_j) \frac{d\pi_j}{d\delta_j} \frac{\partial Q_j}{\partial A_{jj}} - 1 \leq 0, \text{ with equality if } A_{jj} > 0 \quad j = 1, ..., n, \tag{5}
\]

where the partial derivative function \( \frac{\partial Q_j}{\partial A_{jj}} \) may depend on any or all of the arguments of \( Q_j(.) \). The pricing first-order condition (4) can be substituted into the advertising one (5) to give the equilibrium conditions:

\[
M\frac{\partial Q_j}{\partial A_{jj}} \leq 1, \quad \text{with equality if } A_{jj} > 0, \quad j = 1, ..., n. \tag{6}
\]

The interpretation is the following. Raising \( A_{jj} \) by \$1 and raising price by \$\frac{\partial Q_j}{\partial A_{jj}} \) too leaves \( \delta_j \) unchanged. This change therefore increases the revenue by \$\frac{\partial Q_j}{\partial A_{jj}} \) on the existing consumer base (i.e., \( M\frac{\partial Q_j}{\partial A_{jj}} \) consumers). This extra revenue is equated to the \$1 marginal cost of the change, the RHS of (6). We term the relation in (6) the non-comparative advertising quasi-reaction function. It is a function of whatever advertising variables are in \( Q_j \) (note that they all involve firm \( j \) as either emitter or target), along with \( j \)'s share. This differs from a full reaction function because it still may include \( j \)'s other advertising choices, and because it includes the market share, which in turn includes all prices and advertising.

The relationship in (6) already gives a strong prediction for markets where there is no comparative advertising (e.g., when comparative advertising is barred). Indeed, suppose that the perceived quality changes with advertising in the same (concave) manner for all firms. Then the firms with larger market shares will advertise more.36 The intuition is that the advertising cost per customer is lower for larger firms. This is a useful characterization result for advertising in general: note (as per the discussion in the introduction) that it is not a causal relationship. The fundamental parameters of the model determine which firms will be large and advertise more. For example, if firms differ by intrinsic "quality" which is independent of the marginal benefit from advertising (this is the case for our parameter \( W_j \) in the econometric specification below in Section 5), then one might expect that firms with higher such quality will be those advertising more.37 The same relation holds in the presence of comparative advertising, given some strong separability properties on \( Q_j(.) \).

34 These conditions can be written in the form of elasticities. This yields Dorfman-Steiner conditions for differentiated products oligopoly; the comparative advertising conditions below can also be written in such a form.

35 If \( \frac{\partial Q_j}{\partial A_{jj}} \) were constant (which would arise if ads entered perceived quality linearly), then it is unlikely that the system of equations given by (6) has interior solutions. Below we (implicitly) invoke sufficient concavity of \( Q_j \) for interior solutions.

36 In this case, \( M\frac{\partial Q_j}{\partial A_{jj}} = 1 \), is the first order condition, with (temporarily) \( Q(.) \) the production of quality from advertising. Clearly, the larger is the share, the smaller must be \( Q \), and hence the higher must be ads. Note we did not use any symmetry property of the share formula: what did all the work was the same \( Q \) function.

37 This indeed can be shown to be the case in some specifications of the model.
Proposition 1 (Non-Comparative Advertising levels) Let \( Q_j(\cdot) \) be additively separable, and let the function \( \frac{\partial Q_j}{\partial A_{jj}} \) be the same decreasing function of \( A_{jj} \) for all firms, \( j = 1, ..., n \). Then, in equilibrium, firms with larger market shares will use more non-comparative advertising.

**Proof.** From the relation (6), any firm which is active in non-comparative advertising will set its corresponding advertising level to satisfy \( Ms_j \frac{\partial Q_j}{\partial A_{jj}} = 1 \). Since \( \frac{\partial Q_j}{\partial A_{jj}} \) is decreasing in \( A_{jj} \), firms for which \( s_j \) is larger will advertise more (choose a higher value of \( A_{jj} \)) than those with smaller market shares. For firms with low enough market shares, from (4) 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 (5) is negative when \( \frac{\partial Q_j}{\partial A_{jj}} \) is evaluated at \( A_{jj} = 0 \). □

Although we will not impose the strong separability in our estimation below (for reasons elucidated in Section ), the Proposition is still a useful benchmark (and indeed covers the case of no comparative advertising), even though the conditions given are strong. For the model we estimate, the Proposition holds, without imposing additive separability, as long as other advertising levels are constant. However, Proposition 2 works against this effect, while Proposition 3 works in its favor. This is because a larger firm tends to attack more (Proposition 2), which decreases the marginal efficiency of non-comparative ads, but on the other hand, it also tends to be attacked more, thus increasing the marginal efficiency of non-comparative ads. The net effect cannot be determined \textit{a priori}.

We now turn to comparative advertising levels, employing a further restriction on demands.

### 3.3.3 Comparative Advertising

The general problem is more opaque than for own ads, so we use a logit formulation. Then, assuming the idiosyncratic match terms are i.i.d. with the Type 1 Extreme Value Distribution, the market share for Firm \( j \) (fraction of consumers buying from Firm \( j \)) will be given by the logit formulation as:

\[
s_j = \frac{\exp[\delta_j/\mu]}{\sum_{k=0}^{n} \exp[\delta_k/\mu]}, \quad j = 0, ..., n, \tag{7}
\]

This formulation has important properties (readily proved by simple differentiation) useful to the subsequent development. First, cross effects are given as:

\[
\frac{ds_j}{d\delta_k} = -\frac{s_j s_k}{\mu}, \quad j = 0, ..., n, \quad j \neq k, \tag{8}
\]

which is also the expression for \( \frac{ds_k}{d\delta_j} \) (such symmetry is a general property of linear random utility models: see Anderson, de Palma, and Thisse, 1992, Ch. 2, for example).

\[38\] There is a further caveat here that larger firms do not necessarily comparatively advertise the most, because they do not attack themselves.
Second, the own effect is readily derived as:

\[ \frac{ds_j}{d\delta_j} = \frac{s_j(1 - s_j)}{\mu}, \quad j = 0, \ldots, n, \tag{9} \]

Using this expression, the price first-order condition (4) under the logit formulation is now

\[ \frac{d\pi_j}{dp_j} = Ms_j - M(p_j - c_j)\frac{s_j(1 - s_j)}{\mu} = 0, \quad j = 1, \ldots, n. \tag{10} \]

Recalling that the perceived quality is \( Q_j(A_{jj}, \{ A_{jk}\}_{k \neq j}, \{ A_{kj}\}_{k \neq j}) \), we can determine the advertising spending against rivals by differentiating (3) to get (for \( k = 1, \ldots, n, \ j = 1, \ldots, n, \ k \neq j \)):

\[
\frac{d\pi_j}{dA_{jk}} = \frac{\partial Q_j}{\partial A_{jk}} + M(p_j - c_j)(\frac{s_j}{\mu} - \lambda) - \gamma \leq 0, \quad \text{with equality if } A_{jk} > 0.
\]

Inserting the price first-order conditions (10) gives (for \( k = 1, \ldots, n, \ j = 1, \ldots, n, \ k \neq j \)):

\[
\frac{d\pi_j}{dA_{jk}} = Ms_j \frac{\partial Q_j}{\partial A_{jk}} - M\frac{s_j s_k}{1 - s_j} \frac{\partial Q_k}{\partial A_{jk}} \leq \gamma. \tag{11}
\]

The relation between market share and comparative advertising takes a particularly clean form when the quality function embodies a perfect substitutability relation. This formulation includes the Net Persuasion form used below in the estimation (as well as the semi-separable form of earlier versions of the paper).

Suppose therefore that the quality function can be written as

\[
Q_j(A_{jj}, \{ A_{jk}\}_{k \neq j}, \{ A_{kj}\}_{k \neq j}) = Q_j(A_{jj} + \lambda \sum_{k \neq j} A_{jk}, \{ A_{kj}\}_{k \neq j}, \ j = 1, \ldots, n), \quad \text{where } 0 < \lambda < 1 \text{ reflects the idea that comparative advertising should not have a stronger DIRECT effect than non-comparative advertising.} \]

Suppose for the present argument that the solution for non-comparative ads is interior. Then, the non-comparative advertising condition

\[
(Ms_j \frac{\partial Q_j}{\partial A_{jj}} = 1) \]

implies that \( Ms_j \frac{\partial Q_j}{\partial A_{jj}} = \lambda \), and hence, using equation (11), we can write:

\[
(0 < \lambda) - M\frac{s_j s_k}{1 - s_j} \frac{\partial Q_k}{\partial A_{jk}} \leq \gamma - \lambda. \tag{12}
\]

---

39 These properties are related to the IIA property of the Logit model: as an option becomes more attractive, it draws customers from other products in proportion to the product of its own and their market shares.

40 When the (pure) non-comparative advertising level is positive, its condition gives (as before):

\[ Ms_j \frac{\partial Q_j}{\partial A_{jj}} = 1, \quad j = 1, \ldots, n. \]

Hence we can write the comparative advertising first-order condition (for positive \( A_{jk} \)) as:

\[ \frac{\partial Q_j}{\partial A_{jk}} = \frac{s_k}{1 - s_j} \frac{\partial Q_j}{\partial A_{jj}}, \quad k = 1, \ldots, n, \ j = 1, \ldots, n, \ k \neq j. \]

The first term on the LHS can naturally be interpreted as the marginal rate of substitution of the two ad types into perceived quality, the second term reflects the additional benefit from denigration, while the RHS is the relative price.

41 The Net Persuasion form used below has

\[ Q_j(A_{jj} + \lambda \sum_{k \neq j} A_{jk}, \{ A_{kj}\}_{k \neq j}) = Q_j(A_{jj} + \lambda \sum_{k \neq j} A_{jk} - \omega \sum_{k \neq j} A_{jk}, \{ A_{kj}\}_{k \neq j}). \]
The intuition is as follows. Raising $A_{jk}$ by $1$ is equivalent to brand $k$ raising its price by $\frac{\partial Q_k}{\partial A_{jk}}$ (since the same $\delta_k$ is attained). Such a rival price change (which $j$ thus effectuates through comparative advertising) causes $j$’s market share to rise by $\frac{s_j\delta_k}{\mu}$. This increment is valued at $M(p_j - c_j)$. By the price first-order condition, $p_j - c_j = \frac{1}{\mu(1-s_j)}$, and (12) follows. This relation (12) generates two strong results that relate comparative advertising to market share. A sufficient condition for these results to hold is that the quality function takes one of the two following forms:

Q1. Let the quality function be $Q_j(A_{jj} + \lambda \sum_{k \neq j} A_{jk}, \{A_{kj}\}_{k \neq j})$, with $Q_j(\cdot)$ additively separable in incoming comparative ads, $\{A_{kj}\}_{k \neq j}$, with $\frac{\partial Q_j}{\partial A_{kj}}$ the same increasing function of $A_{kj}$ for all firms, $j,k = 1,\ldots,n$.

Q2. Let the quality function be $Q_j(A_{jj} + \lambda \sum_{k \neq j} A_{jk} - \omega \sum_{k \neq j} A_{kj}, \{A_{kj}\}_{k \neq j})$, with $Q_j(\cdot)$ additively separable in Net Persuasion, $A_{jj} + \lambda \sum_{k \neq j} A_{jk} - \omega \sum_{k \neq j} A_{kj}$, and incoming comparative ads, $\{A_{kj}\}_{k \neq j}$. Denote the marginal effect of $A_{kj}$ on $Q_j$ that does NOT come through Net Persuasion as $\frac{\partial Q_{PUSH}^j}{\partial A_{kj}} < 0$, and assume this is the same increasing function of $A_{kj}$ for all firms, $j,k = 1,\ldots,n$.

We are now ready to state the targeting share results.

**Proposition 2 (Larger target more)** Let the quality function satisfy either Q1 or Q2. Then, in equilibrium, for all firms using a strictly positive level of non-comparative advertising, larger firms will use more comparative advertising against each target.

**Proof.** Consider first firms using a strictly positive level of comparative advertising against target $k$. Then (12) holds with equality, i.e.,

$$-M \frac{s_j s_k}{1 - s_j} \frac{\partial Q_k}{\partial A_{jk}} = \gamma - \lambda.$$  

We now consider the two different Q specifications.

Q1. For any given target $k$, note that the ratio $\frac{s_j}{(1-s_j)}$ on the LHS above is decreasing in market share, $s_j$. Hence $\frac{\partial Q_k}{\partial A_{jk}} (< 0)$ must be higher the larger is $s_j$, and the corresponding $A_{jk}$ must be larger since $\frac{\partial Q_k}{\partial A_{jk}}$ is increasing and the same for all firms. For firms with low enough market shares, from (4) the term $(p_j - c_j) \frac{ds_j}{\partial A_{jk}}$ is small enough that (12) holds with strict inequality when $\frac{\partial Q_k}{\partial A_{jk}}$ is evaluated at $A_{jk} = 0$.

Q2. We can break down the term $\frac{\partial Q_k}{\partial A_{jk}}$ into two parts, the one through Net Persuasion, and the other through the direct Push effect. The former is equal to $-\omega \frac{\partial Q_k}{\partial A_{kk}}$ while the latter is $\frac{\partial Q_{PUSH}^k}{\partial A_{jk}}$, which is assumed to be negative. Then we have, by substitution, $-M \frac{s_j s_k}{1 - s_j} \left[ -\omega \frac{\partial Q_k}{\partial A_{kk}} + \frac{\partial Q_{PUSH}^k}{\partial A_{jk}} \right] = \gamma - \lambda$; recalling that $M s_k \frac{s_j}{\partial A_{kk}} = 1$ when $k$ engages in non-comparative advertising, then this equation which determines comparative advertising becomes

$$s_j \left[ \omega - M s_k \frac{\partial Q_{PUSH}^k}{\partial A_{jk}} \right] = \gamma - \lambda.$$  

This yields the comparative advertising quasi-reaction function for the case at hand. For any given target $k$, the ratio $\frac{s_j}{(1-s_j)}$ on the LHS is decreasing in market share, $s_j$. Hence $\frac{\partial Q_{PUSH}^k}{\partial A_{jk}} (< 0)$ must be higher the
larger is $s_j$, and the corresponding $A_{jk}$ must be larger since $\frac{\partial Q_{push}}{\partial A_{jk}}$ is increasing and the same for all firms.

This follows from the logit property that the fall-out is greater from peeling off consumers from a larger rival. This suggests that the largest brands will also be those attacked most (Tylenol in our industry context.) The property also extends to the case when the quality function depends on net persuasion and incoming attacks.

Looking from the perspective of attack targets as a function of attacker size, we have:

**Proposition 3 (Larger targeted more)** Let the quality function satisfy either $Q1$ or $Q2$. Then, considering attacks from firms with positive levels of non-comparative advertising, in equilibrium, larger firms suffer more attacks from each rival.

**Proof.** For $Q1$, the proof is analogous to that of Proposition 2, noting that for any given rival $j$, the LHS of (12) is increasing in market share of the firm attacked, $s_k$. For $Q2$, the result follows from (12) by noting (on the LHS) that the larger is $s_k$, then the smaller must be $-\frac{\partial Q_{push}}{\partial A_{jk}}$, which in turn means that $A_{jk}$ must be larger.

Before turning to the econometric specifications, we first discuss the data: note in particular that Table 2 below roughly supports the two preceding Propositions.

### 4 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, Acetaminophen, Ibuprofen, and Naproxen Sodium. The nationally advertised brands are such familiar brand names as Tylenol (acetaminophen), Advil and Motrin (ibuprofen), Aleve (naproxen sodium), Bayer (aspirin or combination), and Excedrin (acetaminophen or combination). **Table 1** summarizes market shares, ownership, prices and advertising levels in this industry.\(^{42}\)

We use three different data-sets: (1) sales (2) advertising, and (3) medical news data. Sales and advertising data were collected by AC Nielsen and TNS - Media Intelligence respectively, and we coded the advertising content. We constructed the medical news data-set from publicly available news archives.

#### 4.1 Product

The product level data consist of 4-weekly observations of average prices, dollar sales, and dollar market shares (excluding Wal-Mart sales) of all OTC pain relievers sold in the U.S. national market during the 5

\(^{42}\)We exclude Midol and Pamprin from the sample because they are both aimed more narrowly at the menstrual pain-relief market and they both have small market shares.
years from March of 2001 through December of 2005 (a total of 58 monthly observations).\textsuperscript{43} We have data on essential product attributes noted on the packages and the fraction of products sold of each such type: active ingredient, strength (regular, extra strength, etc. - as regulated by the FDA), pill type (caplet, tablet, gelcap, etc.), number of pills contained in the product, and purpose (menstrual, migraine, arthritis, general, children, etc.), although in the end we did not use these data. There were no brand introductions during the period analyzed.

To convert sales quantities to the market shares to be used in the estimation, we need a measure of the total market size, which also defines the demand for the outside good. We used the Census data on the number of adults (18 years or older) in the U.S. multiplied by the average number of pain days an individual has,\textsuperscript{44} and by the maximum FDA-allowed number of pills for 24 hours. This we used to define a "serving" below: therefore, a "price per serving" is the price to the consumer of a day’s worth of pain relief at maximum FDA dosage. Each brand’s individual share was computed as the fraction of total pain killed by that drug.

4.2 Advertising Data

The advertising data contain monthly advertising expenditures on each ad, and video files of all TV advertisements for the 2001-2005 time period for each brand advertised in the OTC analgesics category. The vast majority of the advertising budgets (at least 88%) were spent on broadcast television advertising, and we ignore here other forms of advertising (chiefly magazines). The major novelty of this data-set is that it enables us to include advertising content (focusing on comparative advertising) in the analysis of this market.

The advertising data-set includes 4503 individual commercials. Out of 4503 commercials, 346 had missing video files. Each individual video was aired multiple times: the total number of commercials shown over the 5 year period in all types of TV media was 595,216. All the included ads were watched, and coded according to their content. Specifically, we recorded whether the commercial had any comparative claims – whether the product was explicitly compared to any other products. If a commercial was comparative, we also recorded which brand (or class of drugs) it was compared to (e.g. to Advil or Aleve; or to Ibuprofen-based drugs). The coding enables us to split the advertising expenditure into comparative advertising against each of the mentioned rivals. If an ad had no comparative claims, it was classified as a non-comparative ad. If it was comparative, we divided the expenditures equally across all brands targeted. Table 2 presents the complete picture of cross targeting and the advertising expenditure on each of the rival brand targeting. This table shows every nationally advertised brand used comparative advertising during the sample period. However, the brands against which comparisons were made are only a subset of the nationally advertised brands. The targets are the "big Three:" Tylenol, Advil, Aleve, plus Excedrin.\textsuperscript{45}

\textsuperscript{43}4-week product level data was normalized to monthly frequency to match the advertising data frequency.

\textsuperscript{44}Source: Centers for Disease Control and Prevention, http://www.cdc.gov/

\textsuperscript{45}Motrin does not attack Tylenol because the parent company is the same; likewise, Bayer does not attack Aleve for the same reason. However, we have effectively ignored these multi-product firm relations in the data.
Notice that these data provide some informal support for Propositions 2 and 3. The entries on the diagonal are zeroes through not attacking oneself: a couple of the others are explained by cross-ownership (which we discuss further in the conclusions).

4.3 News Shocks

We follow an approach similar to Chintagunta, Jiang and Jin (2007) to collect the data on these shocks. We used Lexis-Nexis to search over all articles published between 2001 and 2005 on topics related to the OTC analgesics industry. The keywords that we used consisted of brand names, such as "Aleve," "Tylenol," "Advil," "Vioxx," and the names of their active ingredients, such as "Naproxen," or "Acetaminophen." Then we made searches using generic terms such as "pain killers" or "analgesics." We recorded article name, source and date. From a data-set of articles we then constructed a data-set of news shocks. First, multiple articles reporting the same news were assigned to a unique shock ID. Second, we checked whether a news shock was associated with any new medical findings that were published in major scientific journals. As a result of this data cleaning, our news shock data-set includes 15 news shocks between March of 2001 and December of 2005. Finally, we classified the shocks by their impact. If a news shock was reported in a major national newspaper (USA Today, Washington Post, Wall Street Journal, New York Times), then we classified it as a major shock. Otherwise we classified it as a minor shock. This classification is useful to verify whether our identification strategy is robust to changes in the way we define news shocks. Table 3 reports the news shocks, by their title, date, scientific publication, and impact (Major or Minor).

For each major shock that happened during period $t$ we constructed a dummy variable which is equal to 1 in all the periods after and including $t$: $t$, $t + 1$, ..., $T$.\textsuperscript{46} In the empirical analysis below, we interacted each of the major shocks listed in Table 3 with brand dummies. This approach enables us to let the data determine whether a medical news shock affected the demand (instead of us arbitrarily assigning which shock affected which brand in which way), and, if it did, whether a shock had a positive or negative effect on that brand.\textsuperscript{47} We assume in the model below, that news shocks surprise both consumers and firms and they are treated as brand characteristics that affect the underlying quality of any given brand, along with the marginal efficiency of advertising.

Figure 1 presents the occurrence of the 8 major shocks, highlighting the reaction of sales and advertising to those medical shocks.

\textsuperscript{46}We experimented with allowing shocks to depreciate over time at varying rates, but found out that the version without depreciating had a better explanatory power. Also, allowing shocks to affect brands only in the short term (varying number of periods after the shock happened) did not prove to be an effective strategy as well.

\textsuperscript{47}We first included all the 15 shocks listed in Table 4, but quickly discovered that only the major shocks had consistent impact on analysed brands. Hence our analysis focuses only on those 8 major shocks.
5 Econometric Analysis

Here we discuss the quality function upon which we base the structural empirical analysis. We first recall that non-comparative ads are given by (6):

$$M_{s_j} \frac{\partial Q_j}{\partial A_{jj}} \leq 1,$$

with equality when $A_{jj} > 0$. Comparative advertising is given from (11) as

$$M_{s_j} \frac{\partial Q_j}{\partial A_{jk}} - M \frac{s_j s_k}{(1 - s_j)} \frac{\partial Q_k}{\partial A_{jk}} \leq \gamma,$$

with equality when $A_{jk} > 0$. We want to estimate these equations, and to do so we use a particular formulation for $Q_j(\cdot)$. We first present it, and discuss its properties, and then we give some background supporting the choice of the particular function.

In the following, we will separate out the advertising contribution to perceived quality from the intrinsic, or “base quality.” That is, we will write

$$Q_j(\cdot) = \bar{Q}_j(\cdot) + \bar{W}_j$$

where only $\bar{Q}_j(\cdot)$ depends on advertising levels, and $\bar{W}_j$ is a variable specific to Firm $j$ which affects quality with no interaction with $j$’s advertising.

The base quality function we use is:

$$\bar{Q}_j \left( A_{jj} + \lambda \sum_{k \neq j} A_{jk}, \{A_{kj}\}_{k \neq j} \right) = \alpha_D \ln \left( A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj} - \omega \sum_{k \neq j} A_{kj} \right) - \beta \sum_{k \neq j} \ln \left( A_{kj} + \bar{A}_{kj}^{in} \right),$$

(14)

By contrast to the $\bar{W}_j$, the $A$ variables with overbars do interact with their corresponding advertising levels, and determine the marginal efficiency of non-comparative and comparative advertising. For example, the higher is $\bar{A}_{jj}$, the lower is the marginal efficiency of non-comparative advertising; while the higher is $\bar{A}_{kj}^{in}$, the lower the marginal efficiency of attacks by $k$ against $j$, in the sense of less incremental Pull-down.

In the econometric specification, both types of variables will depend on some of the observed variables (for example news shocks) as well as some of the random shocks. The last terms are attacks, and have decreasing

---

48 The key elements in the structure of the model which tie together to get these strong results are that we use a one-stage game (ads and prices are set simultaneously), a discrete-choice-based demand structure, and a simple form for $Q_j$ below. These results form a benchmark which generates strong and simple predictions, which are picked up in the empirical investigation.

49 In an earlier version of the paper we used the functional form $\bar{Q}_j(\cdot) = \alpha_D \ln \left( A_{jj} + \lambda \sum_k A_{jk} + \bar{A}_{jj} \right) - \beta \sum_{k \neq j} \ln \left( A_{kj} + \bar{A}_{kj} \right)$. This functional form leads to the following first order condition for non-comparative advertising: $A_{jj} = \max \left\{ \alpha_D M_{s_j} - \lambda \sum_{k \neq j} A_{jk} - \bar{A}_{jj}, 0 \right\}$. It leads to the following first order condition for comparative advertising: $A_{jk} = \max \left\{ -\bar{A}_{jk} + M \frac{s_{jk}}{1 - s_j} - \frac{\bar{A}_{jj}}{1 - s_j}, 0 \right\}$. Notice that these first order conditions are very similar to the ones that we estimate. We use the specification in the main text of the paper because we find that $\omega$ is statistically and economically different from zero.
marginal impact: an attack from several directions hurts more than the same spending from one direction. (Think consumer perception.) Below we refer to the \( \tilde{W} \) variables as base quality, while the \( \tilde{A} \) variables are called advertising base allure.

We describe parameters by reference to the logit formulation. In that case, the demand numerator of (7) becomes:

\[
\left( A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \tilde{A}_{jj} - \omega \sum_{k \neq j} A_{kj} \right)^{\alpha_D/\mu} \Pi_{k \neq j} \left( A_{kj} + \tilde{A}_{kj} \right)^{-\beta/\mu} \exp \left( \frac{\tilde{W}_j - p_j}{\mu} \right).
\]

Interpretation is best done in terms of the two constituent effects, here Net Persuasion and Pull, although matters are complicated by the fact that the incoming attacks enter both of these. In terms of perceived quality, \( Q_j(\cdot) \), \( \beta \) is (minus) the elasticity due to the Pull effect of the incoming comparative advertising attack from \( k \). Hence, the \( \beta \) parameter is loosely (\( \mu \) times) the elasticity of demand with respect to incoming attacks in the Pull effect (loosely because this is just the denominator of demand, and also it is more specifically the elasticity with respect to \( A_{kj} + \tilde{A}_{kj} \)). Incoming attacks also enter via Net Persuasion: \( \omega \) tells us how substitutable incoming attacks are in Net Persuasion. Likewise, the parameter \( \lambda \) tells us how substitutable are own outgoing ads for non-comparative ads in the Net Persuasion component, and is a key parameter of interest below. The parameter \( \alpha_D \) gives (\( \mu \) times) the elasticity of demand (again with the loose proviso) with respect to Net Persuasion.

5.1 The Equations to Be Estimated

The derivative with respect to \( A_{jj} \), corresponding to own non-comparative advertising choice, is (see (6)):

\[
M_{s_j} \left( A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \tilde{A}_{jj} - \omega \sum_{k \neq j} A_{kj} \right)^{\alpha_D/\mu} \leq 1, \text{ with equality if } A_{jj} > 0.
\]

Therefore we run the non-comparative ad equations as:

\[
A_{jj} = \max \left\{ -\tilde{A}_{jj} + \alpha_D M_{s_j} - \lambda \sum_{k \neq j} A_{jk} + \omega \sum_{k \neq j} A_{kj}, 0 \right\}, \quad j = 1, \ldots, n. \tag{15}
\]

These equations enable us to determine the \( \tilde{A}_{jj} \), \( \alpha_D \), \( \lambda \), and \( \omega \) parameters. We expect \( \tilde{A}_{jj} \) to be positive so that there is a positive base allure, and the marginal efficiency of the first dollar of persuasion is not infinite. Proposition 1 suggests that \( \alpha_D \) should be positive, so that for firms with the same \( \tilde{A}_{jj} \), the higher market share goes together with the higher advertising level. We expect \( \lambda \in (0, 1) \) so that outgoing attacks aid Net Persuasion, although less effectively than non-comparative ads. We expect \( \omega > 0 \) so that attacks reduce net persuasion. Since an attack also has a direct impact through the Pull effect, we might expect that the effect of an attack in Net Persuasion can be undone by a non-comparative ad, so \( \omega < 1 \). The latter property though is not predicted from the model strictu sensu. Since the econometric analysis does not restrict the parameters to lie within the suggested bounds, we would view parameter estimates within the
suggested bounds as quite a strong confirmation of the model, and especially if they lay in the middle of the suggested bounds.

Now consider comparative ads. From (11), we have:

\[
M_{sj} = \alpha D \lambda \left( A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj} - \omega \sum_{k \neq j} A_{kj} \right) + M_{sj} s_{k} \omega \alpha D \lambda \left( A_{kk} + \lambda \sum_{l \neq k} A_{kl} + \bar{A}_{kk} - \omega \sum_{l \neq k} A_{lk} \right) + M_{sj} s_{k} \beta \left( 1 - s_{j} \right) \left( 1 - s_{j} \right) \omega A_{jk} \leq \bar{A}_{jk}
\]

with equality if \( A_{jk} > 0 \). Note that we are not going to get simple linear equations to estimate if non-comparative advertising is zero (i.e., the equation above is not linear).\(^{50}\) However, assuming that own non-comparative advertising is positive, we can replace the first term by substituting \( j \)'s non-comparative advertising quasi-reaction function, and, if the rival has positive non-comparative advertising too, we can also replace the second term by substituting \( k \)'s non-comparative advertising quasi-reaction function, and get (see also (13))

\[
M_{sj} s_{k} \frac{\beta}{(1 - s_{j})} \gamma - \lambda - \frac{s_{j}}{(1 - s_{j})} \omega A_{jk} \leq \bar{A}_{jk}
\]

This yields an equation to be estimated as:

\[
A_{jk} = \max \left\{ -\bar{A}_{jk} + M_{sj} s_{k} \frac{s_{k}}{(1 - s_{j})} - \omega \beta, 0 \right\}.
\]

This is a simple, though non-linear, equation in market shares only: note that other ad types are not included here (contrast (13) above).

The basic idea of the paper is to write the perceived quality function as a function of Push-Me-Up and Pull-You-Down effects. Given this form, in the Appendix we look at some restrictions on plausible functional forms, in terms of some second order considerations, and some restrictions from the data.

### 6 Identification

We estimate the equations (15) and (16)). There are two main concerns that we need to address: left-censoring of non-comparative and comparative advertising and endogeneity of market shares. To begin with, in some periods some brands do not engage in non-comparative or comparative advertising (there are corner

\(^{50}\)In the data there are less than 15 percent of the observations for which non-comparative advertising is zero. We tried the analysis with and without those observations and the results are unchanged. So we decided to keep those observations in the dataset.
solutions), hence the variables $A_{jjt}, A_{jkt}, j, k = 1, \ldots, n$, are left-censored.\textsuperscript{51} We control for the left-censoring by running Tobit regressions.

**The Nature of Endogeneity.** The endogenous variables are $A_{jjt}, A_{jkt}, A_{kjt}, s_{jt}, j, k = 1, \ldots, n$.\textsuperscript{52} To clarify the nature of the endogeneity in our analysis, we start from equation (15), which is the simplest equation to deal with. To further simplify the discussion we assume, just for the sake of exposition, that $\lambda = 0$ and $\omega = 0$. We will drop these two assumptions at the end of this section. Then (15) becomes, with the appropriate time subscripts:

$$A_{jjt} = \max \{ -\bar{A}_{jjt} + \alpha_D M s_{jt}, 0 \}.$$

The term $\bar{A}_{jjt}$ captures the advertising base allure of a brand, which we write as follows:

$$\bar{A}_{jjt} = Z_0 j t \Phi + \xi_{jt},$$

where $Z_{jt}$ are observable determinants of the advertising base allure. In this paper, these are the news shocks. The $\xi_{jt}$ are unobservable shocks to the base allure, so $\xi_{jt}$ is a structural error. Notice that $\xi_{jt}$ is here assumed to be observed by firms, but not by the econometrician.

Next, recall that the market share for brand $j$ is written as:\textsuperscript{53}

$$s_{jt} = \frac{\exp[\delta_{jt}/\mu]}{\sum_{k=0}^{n} \exp[\delta_{kt}/\mu]}, \quad j = 0, 1, \ldots, n$$

where

$$\delta_{jt} = \bar{Q}_{jt} (.) - p_{jt} + W_{jt}. \quad (17)$$

Because firms observe $\xi_{jt}$ when they choose advertising and because shares are a function of advertising (through $Q$, the perceived quality), then shares are a function of $\xi_{jt}$, and thus we will get inconsistent estimates of $\alpha_D$ and $\Phi$ if we run the following simple Tobit regression:

$$\begin{cases}
A_{jjt}^* = -\bar{A}_{jjt} + \alpha_D M s_{jt} - \xi_{jt}, \\
A_{jjt} = \max (A_{jjt}^*, 0) \\
\xi_{jt} \sim N \left(0, \sigma^2 \right)
\end{cases} \quad (18)$$

**Top Brands vs. Other Brands.** The first step to address the endogeneity of the market shares is to exploit the panel structure of our data to account for time-constant differences across brands. Essentially, we model the unobservable $\xi_{jt}$ as follows:

$$\xi_{jt} = \bar{\xi}_{jt} + \Delta \xi_{jt},$$

\textsuperscript{51} As noted above, there are two brands, Pamprin and Midol, which are primarily menstrual formulations, and that we exclude them from the empirical analysis because of their negligible market shares. Interestingly, they never engage in non-comparative advertising, only in comparative advertising. Generic brands never engage in any type of advertising.

\textsuperscript{52} Notice that prices, which are also endogenous, have been substituted out in the equations to be estimated.

\textsuperscript{53} Notice that the generics are included here: abusing notation, generic drugs can be funneled into multiple options $0$. However, as we will show later, we do not need to estimate the demand functions to estimate the relevant structural parameters.
where \( \bar{\xi}_j \) is a brand fixed effect, while \( \Delta \xi_{jt} \) are time specific idiosyncratic shocks. We have investigated various specifications for the fixed effects, and concluded that a specification where there are two fixed effects, one for the top brands (Advil, Aleve, Tylenol), and one for the other brands (Excedrin, Motrin, Bayer) fits our data best. We provide in Figure 2 a graphical description of the relationship between non-comparative advertising and market shares for all brands and months.

Figure 2 shows that there are two types of brands in the market. Aleve, Advil, and Tylenol (the ‘Top Brands’) control large market shares compared to Excedrin, Bayer, and Motrin. This is consistent with the reported weighted market share descriptive statistics in Table 1. This observation parallels the economic intuition that ‘Top Brands’ have a larger advertising base allure which translates into larger inherent quality, \( \bar{A}_{jj} \). Additionally, the linear fit between shares and non-comparative advertising has the same slope for the ‘Top Brands’ and the rest of the brands. We use the evidence from this figure to justify the construction and use of a dummy variable ‘Top Brand’.

Formally, we then have \( \bar{\xi}_j = \bar{\xi}_T \) for \( j \in \{ \text{Advil, Aleve, Tylenol} \} \) and \( \bar{\xi}_j = \bar{\xi}_O \) for \( \{ \text{Motrin, Excedrin, Bayer} \} \). Given our relatively small sample, 342 observations in some specifications, it helps to reduce the number of brand fixed effects. Another useful advantage of having such group-type fixed effects is that we avoid the incidental parameter problem that would have been there with the nonlinear Tobit regression and individual brand-specific fixed effects.\(^{54}\) The remaining source of endogeneity in our regressions then comes from any potential correlation between \( \Delta \xi_{jt} \) and \( s_{jt} \).

One route is then simply to specify conditions under which there is no remaining correlation, and proceed directly to the estimates. This is the essence of Assumption 1. If this is untenable, various exclusion restrictions can remove residual endogeneity. These are described in Assumptions 2 and 3 below. In our regressions, we will start with estimates under the simple Assumption 1, and then proceed to deploy the other 2 Assumptions. (Note that Assumption 1, if correct, obviates the others, while the other two are not mutually exclusive).

**Using Timing to Identify the Parameters.** The parameters of the regression (18) can be identified when \( \Delta \xi_{jt} \) and \( s_{jt} \) are uncorrelated by estimating a variant of (18) where the \( \xi_{jt} \) are allowed to have different means corresponding to the brand-group fixed effects. The (non-)correlation condition can be given a justification, paralleling a standard assumption in a large part of the literature estimating production functions with a particular assumption on the timing of the realizations of the errors.\(^{55}\) More specifically, a sufficient condition is the following:

\(^{54}\) One advantage of having this group fixed effects is that we avoid the incidental parameter problem that would have been there with individual brand-specific fixed effects. Notice, however, that even with individual brand specific fixed effects that incidental parameter problem would be marginal for two reasons. First, the time dimension grows over time, while the number of brands remains equal to six. Second, the incidental parameter problem is less important with a Tobit than with a Probit [here cites].

\(^{55}\) See Griliches and Mairesse [1999] for an illuminating review of the literature on the estimation of production functions.
Assumption 1 After controlling for the news shocks, which we assume to enter directly through $Z_{jt}$, and after including brand fixed effects, the time specific idiosyncratic error $\Delta \xi_{jt}$ is uncorrelated with $s_j$, that is $E (\Delta \xi_{jt} | s_{jt}, Z_{jt}) = 0$.

Clearly, the news shocks are exogenous since they require new medical discoveries, which ‘surprise’ both the consumers and the firms. Here, variation in the knowledge of the health properties of the products is captured by the news shocks. One standard interpretation for this maintained assumption is that we are basically able to observe all the variables that the firms take into account when taking their decisions, including the news shocks (e.g. the information that consumers and firms have at any point in time). This means that neither the econometrician nor the firms observe $\Delta \xi_{jt}$ before taking their advertising and pricing decisions. When this assumption is untenable, identification can be achieved using exclusion restrictions.

We now discuss the crucial identification assumption of this paper.\(^5\)

**Exclusion Restrictions.** We need variables that affect advertising only through shares, but not directly. We seek variables that affect shares through prices, $p_{jt}$, but do not affect perceived quality (such the cost of producing a pill).\(^6\) To this end we make the following identification assumption:

Assumption 2 The prices of the generic products are set equal to their marginal costs, which are assumed to be constant. The prices of the generics enter into each branded product’s market share but are excluded from the equation (18).

First, the marginal cost of production of a generic product must be constant; otherwise, the price of the generic would depend on the quantity produced by the branded products, and so it would not be exogenous.\(^7\) Second, Bertrand competition among generic producers of the drugs with the same active ingredient leads to pricing at marginal cost.\(^8\) If, as to be expected, the cost of producing generic products is highly correlated with the cost of producing branded products, then generic prices have an additional indirect impact on branded products’ market shares through branded prices.

\(^5\) Notice that one could assume that the news shocks affect the utility derived by consuming that product (and its demand) but do not affect the advertising base allure, which is then assumed to be independent of the clinical properties of the active ingredients of a product. More formally, the news shocks enter into $W_{jt}$ but do not enter into $Z_{jt}$. Essentially, the advertising base allure is a function of the image or reputation of a brand, and the image and reputation is independent of the medical properties of a product. This would be the case if we believed that the consumer has a full knowledge of the medical properties of a product, and thus advertising cannot change the value of such properties to the consumer. Under this interpretation, the perceived quality of a product is not a function of its medical properties and the news shocks could be used as instrumental variables. However, we do find that news shocks play an important role as predictors of the advertising decisions in our first order conditions. Thus, the evidence is against using this identification assumption.

\(^6\) Notice that the fact we have been able to substitute out prices from the advertising first-order conditions means that we need not worry about changes in prices affecting advertising. By substituting out prices, the impact of price on advertising goes through market share.

\(^7\) The marginal cost for pharmaceuticals is reasonably constant, in the sense that there are not increasing returns to scale.\(^\text{CITE (fixed costs?)}\)

\(^8\) Notice that we can allow generic brands to charge prices that are higher than marginal costs as long as this is explained by local conditions that national brands do not take into account when they set their prices.
In practice, there are two basic instrumental variables for each share $s_{jt}$: the price of the generic product that uses the same active ingredient as the brand $j$; and the sum of the prices of the analogous instrumental variables for its five competitors.\(^{60}\) In addition, we include the interaction of the first (the generic price) and the second one (the sum of the prices of the other generics); and the squared terms of the first and the second.

Then, we interact these two instrumental variables with the news shocks. While the news shocks enter directly in the equation (18), their interactions with prices are clearly excluded from that equation.

To implement our estimation in our non-linear models, we use control functions (Heckman and Robb \[1985,1986\]).\(^{61}\) Our methodology follows Blundell and Smith (1986) and Rivers and Vuong (1988).

**Generalizing the Identification Strategy.** In the above discussion we have focused on the first order condition (15) under the assumptions that $\lambda = 0$ and $\omega = 0$. It is quite clear that even if we let that $\lambda$ and $\omega$ to be different from zero, we can use the same instrumental variables. This is exactly what we do. Essentially, we use variation in the generic prices and their interactions with the news shocks to identify the effect of all of our endogenous variables.\(^{62}\)

### 7 Results

#### 7.1 Non-Comparative Advertising

We estimate the following statistical specification for non-comparative advertising (see (15)):

$$
\begin{align*}
A_{jjt}^* &= -Z_{jt}^T \Phi + \alpha_D M s_{jt} - \lambda \sum_{k \neq j} A_{jk} + \omega \sum_{k \neq j} A_{kj} - \bar{\xi}_j - \Delta \xi_{jt}, \\
A_{jjt} &= \max \left( A_{jjt}^*, 0 \right).
\end{align*}

\text{(19)}
$$

Notice that $\alpha_D$, $\lambda$, and $\omega$ are identified here. Clearly, one of the great advantages of using the functional

---

\(^{60}\)For example, the instrumental variables for the share of Tylenol at time $t$ are the (average) price of its generic version that uses the same active ingredient, Acetaminophen; and the sum of the (average) prices of the the generic price of Advil (Ibuprofen), Bayer (Aspirin), Aleve (Naproxen), Motrin (Ibuprofen), and Excedrin (Excedrin).

\(^{61}\)In practice, the estimation is made in two steps. First, we run the LHS endogenous variables (here market shares) on all exogenous variables, including those excluded from the second stage relationship. Then, we run the second stage regression (advertising levels here) now including the residuals from the first regression as an additional explanatory variable (the “Control Function”) to all the second stage explanatory variables. For example, if we want to estimate the parameters of the non-comparative advertising first order condition (ads on sales), we first run shares on generic prices and news shocks, and compute the residuals. Then we run a Tobit where ads are explained by market share, news shocks (if not excluded) and the residuals.

\(^{62}\)Thus, there are no exogenous variables that identify shares but not the other advertising variables. We know that advertisers must meet the Federal Trade Commission (FTC) standard of truthful and not misleading advertising claims. All material claims must be substantiated by a reasonable basis of support and firms need to evaluate whether their promotional message is likely to be challenged by a competitor or ad monitoring institution. Failure to have robust substantiation for a commercial may result in serious and costly consequences among which are failure to gain network approval and high litigation costs. The most common consequence is the publicized disruption of the ad campaign, sunk costs invested in the ad campaign and negative press related to the brand name. Over the five year period, we observe 15 OTC analgesics advertising claims challenged by the FTC, National Advertising Division (NAD), a competitor or a consumer. The problem with using these data is that the challenges are a function of the amount of advertising expenditures. So they cannot be considered exogenous in our regressions. This problem is not different from the one that it is encountered when we estimate market power and we do not have information on the marginal cost. Adding more equations (the first order condition for price and the demand equation) would let us identify $\alpha_I$, $\gamma$, and $\beta$. 

---

27
form (14) is the transparency and simplicity of the first order condition above.63 We have a linear relationship between expenditures on non-comparative advertising \((A_{jjt})\) and shares \((Ms_{jt})\), outgoing comparative advertising \(\left(\sum_{k \neq j} A_{jk}\right)\), and incoming comparative ad attacks \(\left(\sum_{k \neq j} A_{kj}\right)\).

Recall that \(\lambda\) is a substitutability parameter of outgoing comparative ad with (outgoing, of course) non-comparative ads. In other words, \(\lambda\) measures how much must be spent on non-comparative advertising to replace $1 spent on comparative advertising to generate the same "push" in (own) perceived. For example, \(\lambda = 0.75\) means that the firm can raise its perceived quality by the same amount if it spends 1.33 dollars in comparative advertising or 1 dollar in non-comparative advertising. This parameter does not represent the full effect of comparative advertising relative to non-comparative advertising, as there is also the Pull effect which is directly denigrating the perceived quality of targeted competitors’ brands. Were we to find \(\lambda = 1\) then comparative and non-comparative advertising would have the same effect on the perceived quality of a brand. This would provide support to the coding choice of all the previous studies, which looked at the total expenditure in advertising.64 If \(\lambda \neq 1\), then we should conclude that comparative and non-comparative advertising have different effects and should be coded separately.

\(\omega\) is a substitutability parameter of incoming comparative ad attacks with non-comparative ads. It measures how much a firm should spend on non-comparative advertising to restore $1 worth of detraction from the comparative ad attack of a competitor. For example, \(\omega = 0.5\) means that the firm can keep its perceived quality unchanged if it spends 50 cents on non-comparative advertising or 66 cents \((\omega/\lambda)\) dollars in outgoing comparative advertising. Clearly, \(\omega = 0\) implies that incoming attacks have no effect on perceived quality, which would imply that comparative ads have only a push effect, and no pull effect.

Finally, the fixed effect \(\bar{\xi}_j = \bar{\xi}_{TB}\) if \(j\) is a Top Brand, and \(\bar{\xi}_j = \bar{\xi}_{OB}\) if \(j\) is not a Top Brand (for obvious collinearity reasons, only the fixed effect for Top Brand will be reported).

**Column 1 of Table 4** provides the estimates of \(\alpha_D\), \(\lambda\), and \(\omega\) when we run the simple Tobit regression (15) for non-comparative advertising. The news shocks are not included in this first estimation (so we are setting \(\Phi = 0\)).

The coefficient \(\alpha_D\) is estimated very precisely, both here and in the other specifications, with a parameter estimate of 0.157. This confirms the suggestion of Proposition 1, that larger firms engage in more non-comparative advertising to push up perceived quality and demand. To provide an economic interpretation of the coefficient \(\alpha_D\) we compute the elasticity of non-comparative advertising to shares (market size is essentially constant over time and the same across brands):

\[
e_{A_{jj}, s_j} = \frac{dA_{jj}}{ds_j} \frac{s_j}{A_{jj}}
\]

63 These properties stem from the separability of the Net Persuasion term from the Pull terms, and the logarithmic form it takes.
64 Although this does not matter in industries in which little comparative advertising is used.
We find the median elasticity to be equal to 2.098, which means that a 10 percent increase in market share, $s_j$, implies a 21% increase in non-comparative advertising. This is clearly a strong relationship.

The substitutability parameter, $\lambda$, is estimated to be 0.497. This means that each dollar spent on comparative ad increases the perceived quality of the attacking brand by the same amount as 49.7 cents spent on non-comparative ad. This is gratifyingly close to saying that a comparative ad is half Push - in the comparison one brand is promoted and the other pulled down. The estimation suggests a clean 50-50 split.

The substitutability parameter, $\omega$, is estimated to be 0.347. This suggests that incoming attacks do have a sizeable negative effect on the perceived quality of the attacked firm. Every dollar spent on incoming attacks requires 35 cents to mitigate.

The dummy variable for the Top Brands’ advertising base allure advantage is estimated to be $-0.471$. It has a negative sign, which means that the larger firms, Aleve, Tylenol and Advil have inherently higher advertising base allure than the other brands. The constant is equal to $-0.039$. Because this constant is close to zero and statistically insignificant, the other brands have minimal (although still positive) advertising base allure.

As a measure of fitness, we compute the value of the likelihood function at the maximum, which is equal to 56.576. We will use this measure of fit to compare the results across columns.

**Column 2** adds on the news shocks vector $Z$. Thus, we estimate their effects on the amount spent on non-comparative advertising by getting estimates for $\Phi$. Under Assumption 1, we get consistent estimates of the parameters of the model.

The way we deal with news shocks is the following. We interact each news shock with brand dummies for all brands. This leads to six (brands) times ten (shocks) variables to include in the regression. This way to deal with the shocks lets the data pick up which shocks had an impact on the firms’ decisions and, also, it allows the shocks to have different effects on different brands. Because of the large number of variables, we do not report the results for the shocks, but one can look at the graphs in Figure 1 to get a sense of which shocks had an effect on which firm.

We estimate $\alpha_D$ to be equal to 0.123. The share elasticity of non-comparative advertising is now down to 1.652 from 2.098. The dummy variable TopBrand is now equal to $-0.292$, and is not statistically significant, indicating a lack of evidence that the top brands have better advertising base allure. The estimates of $\lambda$ and $\omega$ are essentially unchanged.

The likelihood function is now equal to 163.729, which says, as one would expect, that adding exogenous variation in the news shocks improves the fit.

---

65 Many of them turn out to be statistically and economically insignificant, but we do not drop them out, since they do not increase the computational time. When we run non-linear tobits (Table 5) we will drop the ones that are neither statistically nor economically significant.

66 Notice that what we are doing is different from using time dummies. Here we are using each news shock as a natural experiment which is allowed to have a different effect on the utilities of each of the six brands.
Column 3 includes shocks as controls and uses generic prices and their interactions with the news shocks as instrumental variables. Essentially we use Assumption 2 to identify the parameters of the model. Column 3 reports the main set of results for the first order condition (19). This specification achieves the best fit among the ones we report in Table 4.

We estimate $\alpha_D$ to be 0.214. The corresponding median elasticity is 2.874. Both the coefficient and the median elasticity are substantially larger than those we estimate in Columns 1 and 2. There is clear evidence that $s_{jt}$ is an endogenous variable in the regression (the coefficient of the control function is $-0.169$ and statistically significant), and the bias is not negligible.

We estimate $\lambda$ equal to 0.532. The estimate of $\lambda$ is not very different from the one in Column 1, and this might be because the first stage IV regression might not explain too much of the variation in $\sum_{k \neq j} A_{jk}$. We think that this is the case because the coefficient of the control function is small (0.026) and not statistically significant. That is, the result for the coefficient of the control function would suggest that $\sum_{k \neq j} A_{jk}$ is exogenous, and this is clearly not plausible (we do estimate the first order conditions for comparative advertising later on). Thus, we should interpret the result of 0.532 as a lower bound on the true value of $\lambda$.

We estimate $\omega$ equal to 0.438. The coefficient is significantly larger than the one we estimated in Columns 1 and 2. Not surprisingly, the control function that we use to address the endogeneity of $\sum_{k \neq j} A_{kj}$ has a large and statistically significant coefficient ($-0.268$).

Finally, the dummy variable for the Top Brands’ advertising base allure advantage is estimated to be $-0.642$. As in Column 1, this provides evidence that the larger firms, Aleve, Tylenol, and Advil, have inherently higher advertising base allure than the other brands.

We summarize our empirical analysis of the first order condition (19) as follows. First, Proposition 1’s suggestion that higher shares, ceteris paribus, are associated with higher non-comparative advertising is confirmed. Second, we find that there are important differences in the advertising base allure of the largest brands versus the other brands. The largest brands have an advertising base allure that substantially and significantly larger than that of the other brands. Third, as we expected, we find evidence of a clear endogeneity of market shares (and other advertising variables) in the advertising first order conditions, which creates a substantial downward bias on the coefficient of market shares. Most importantly, the estimates of the components of the Net Persuasion function lie within the expected ranges. Outgoing attacks are half as powerful as direct non-comparative ads in raising perceived quality. Incoming attacks draw down a brand by around 40 cents, in terms of the non-comparative ads that restore Net Persuasion: however, attacks have an additional negative effect, through the Pull effect. This we estimate next.
7.2 Comparative Advertising

The second relation that we test is comparative ad relation (??). The unit of observation now is a pair of brands, as we study attacks of one brand, \( j \), on another brand, \( k \). Formally, we estimate the following regression, which includes pair specific group-type fixed effects:

\[
A_{jk} = \max \left\{ -\Gamma Z^C_{jkt} - \xi_{TB,TB} - \tilde{\xi}_{TB,OB} - \tilde{\xi}_{OB,TB} + \left[ \frac{(\gamma - \lambda)}{\beta} \frac{1}{M s_{jt} s_{kt}} \frac{1}{M s_{kt}} \left( \frac{\gamma - \lambda}{\beta} + \omega \right) \right]^{-1}, 0 \right\}
\]  
(20)

where \( \tilde{\xi}_{jk} = \tilde{\xi}_{TB,TB} \) if \( j \) and \( k \) are both Top Brands, \( \tilde{\xi}_{jk} = \tilde{\xi}_{TB,OB} \) if \( j \) is a Top Brand (i.e., Advil, Aleve, Tylenol) and \( k \) is an Other Brand, and likewise for \( \tilde{\xi}_{OB,TB} \) and \( \tilde{\xi}_{OB,OB} \) (one is omitted because we include a constant term in the regression). For example, \( \tilde{\xi}_{TB,TB} \) is the pairwise group-fixed effect (to be estimated) if both the ‘attacker’, \( j \), and the ‘attacked’, \( k \), are top brands.

Notice that \( \frac{\gamma - \lambda}{\beta} \) and \( \frac{\omega}{\beta} \) are identified from the above equation, but \( \gamma \), \( \lambda \), \( \omega \), and \( \beta \) are not separately identified. We can recover each of them by using the results from the estimation of (19). In particular, we will be able to recover the deep structural parameters \( \gamma \) and \( \beta \). Recall that the parameter \( \gamma \) measures the relative cost of comparative advertising versus non-comparative advertising. The parameter \( \beta \) is proportional to (minus) the demand elasticity of the pull effect.

Testing the Predictions of the Model. Before embarking on the estimation of the non-linear Tobit (20), we run a simpler test of Propositions 2 and 3. We run the following regression, which can be interpreted as a linearized version of (20):

\[
A_{jkt} = \max \left\{ -\Gamma Z^C_{jkt} - \tilde{\xi}_{TB,TB} - \tilde{\xi}_{TB,OB} - \tilde{\xi}_{OB,TB} + \theta_1 M s_{kt} + \theta_2 M s_{jt} s_{kt}, 0 \right\}.
\]  
(21)

Propositions 2 and 3 predict that both \( \theta_1 \) and \( \theta_2 \) should be positive. That is, the expenditures on comparative advertising by firm \( j \) attacking firm \( k \) are increasing in the share of firm \( k \), \( s_{kt} \), and in the interaction of the shares of \( j \) and \( k \). \( \theta_1 \) and \( \theta_2 \) cannot be related to the structural parameters of the model in a straightforward way. However, we can still use this simple approach to test whether the model’s predictions are consistent with the data.

Columns 1-2 of Table 5 present the estimation results for (21).

**Column 1** presents results from the simple Tobit. We estimate \( \theta_1 \) equal to 0.073 and \( \theta_2 \) equal to 6.139. They are both positive and thus provide evidence in support of the theoretical model developed in Section (3). In particular, these results say that firms have a greater incentive to attack larger firms, and this incentive is increasing in the share of the attacker.

To provide an economic interpretation of these parameters, we can again compute elasticities. The median elasticity with respect to \( s_{jt} \), \( e_{A_{jt}, s_{jt}} = \frac{dA_{jt}}{ds_{jt}} \frac{s_{jt}}{A_{jt}} \), is equal to 1.445 and the one with respect to \( s_{kt} \) is
equal to 5.089. Clearly, the main determinant of comparative ads is the market size of the attacked brand. The larger that is, the larger the comparative ads. In particular, a 10% higher market share implies that the comparative ads against that brand are higher by 50.89 percent.

Column 2 uses generic prices as instruments and includes medical news as control variables. We estimate \( \theta_1 \) equal to 0.099 and \( \theta_2 \) equal to 14.197. These results, along with the coefficient estimates of the control functions suggest that both \( s_{kt} \) and \( s_{jt} \) are endogenous in this regression. In particular, the coefficient for the control function of \( s_{kt} \) is equal to \(-0.158\), and statistically very significant. Given the larger values of \( \theta_1 \) and \( \theta_2 \), we also find larger elasticities. The elasticity with respect to \( s_j \) is now equal to 2.647 while the elasticity with respect to \( s_k \) is equal to 8.053.

The Structural Equation. [THIS SECTION IS IN PROGRESS]

8 Robustness

In our view, the main issue that we have to deal with is whether by omitting dynamic effects, we introduce a bias in the estimation of the relationships between the main variables of the model. There are two related dynamic features that our static model might be missing. First, \( \tilde{A}_{jj} \) and \( \tilde{A}_{jk} \) might be related to the goodwill of a firm, and that goodwill might depend on past advertising decisions of the firm. We can check the importance of this aspect by adding lags in our regressions. Second, as Dube, Hitsch, Manchanda [2005] show in their descriptive analysis, pulsing might play an important role in advertising decisions depending on the industry that we look at. In this section, we look at these two features and check, indirectly, whether omitting them from the analysis might bias our results. Because we are just checking for the robustness of the results, we only look at the non-comparative advertising first order condition.

Goodwill. Advertising goodwill is the idea that past advertising is like an investment over time which creates a stock at any moment. This stock, in turn, is subject to depreciation as the consumer "forgets" past ads. If there are strong stock effects (depreciation is not quick), then firms are engaged in a dynamic game. Solving such a game and writing the appropriate structural model would be substantially more involved than the simple static model characterized above.

Here we essentially estimate the regression (19) after including the one month lagged value of \( A_{jjt} \). That is, we estimate the following:

\[
\begin{align*}
A^*_{jjt} &= \varphi A_{ij,t-1} - Z_{jt}' \Phi + \alpha_D M s_{jt} - \lambda \sum_{k \neq j} A_{jk} + \omega \sum_{k \neq j} A_{kj} - \bar{\xi}_j - \Delta \xi_{jt}, \\
A_{jjt} &= \max (A^*_{jjt}, 0),
\end{align*}
\]

One way to read this equation is to notice that past expenditures in non-comparative advertising enter into the term \( \tilde{A}_{jjt} \), Firm \( j \)'s time \( t \) advertising base allure.
Table 6 presents the results, which should be compared to those in Table 4.

**Column 1 of Table 6** shows the results when we do not control for the endogeneity of shares and advertising expenditures, and when we include the news shocks as control variables. The point is to compare the results in this column with those in **Column 2 of Table 4**. The noticeable difference is that the coefficients are all smaller, but have the same signs. For example, \( \alpha_D \) is now equal to 0.049 instead of 0.123, \( \lambda \) is equal to 0.417 instead of 0.514 and \( \omega \) is equal to 0.226 instead of 0.316. Thus, omitting measures of past advertising choices does not seem to affect our estimation results in a fundamental way.

**Column 2 of Table 6** shows the results when we instrument for the shares, the outgoing comparative advertising, and the incoming comparative advertising. Notice that we do not instrument for the past period non-comparative advertising \( (A_{ij,t-1}) \). We take it as predetermined and uncorrelated with the current values of the unobservables. This column should be compared with Column 3 of Table 4. Again, we do not find dramatic differences. The coefficient estimates are quite different, but the differences are quite small.

Finally, **Column 3 of Table 6** shows the results when we instrument for the shares, the outgoing comparative advertising, the incoming comparative advertising, and the lagged values of non-comparative advertising. The main takeaway from this last column is that our instrumental variables are just unable to identify all four of the endogenous variables. We can see that by the extent to which all coefficients are imprecisely estimated. Still, the magnitudes and signs of the coefficients are not very different from those in **Column 2**.

Overall, we find that adding the lag value does not seem to affect our estimation results in a fundamental way.

**Pulsing.** Pulsing is the phenomenon of uneven advertising levels over time. A campaign will have a specific start date, and a series of ads will be run at quite high intensity. In many industries, there is a considerable lag (or at least a lull) until the next campaign starts up (a new "media blitz"). This pattern is thought more effective than running ads at a steady level, in part because of attention thresholds for individuals' perception, etc.

One very simple way to test whether pulsing occur in this industry is the following: We compare how the results change if we use quarterly instead of monthly data. Dube, Hitsch, Manchanda [2005] show very irregular episodes of advertising to test their theory of pulsing. Clearly, the more one aggregates the data over time, the less irregular the episodes of advertising become. So our idea is that if there is pulsing in our monthly data, and if accounting for pulsing would affect our results radically, then we should see sizeable differences in the estimates that we get by using quarterly instead of monthly data.

Hence we estimate regression (19), but with quarterly instead of monthly data.

The results are striking. With the exception of the third column, to which we will return later, it is clear
that the estimates are basically the same as in Table 4.

Column 1 of Table 7 presents the results when we run a simple Tobit, without including the news shocks. The estimate of $\alpha_D$ looks much bigger but if we look at the elasticity we notice that it is essentially the same as in the Column 1 of Table 4. The estimates of $\lambda$ and $\omega$ are also essentially the same in the first columns of Table 4 and Table 7. Only the estimate of the dummy variable TopBrand is much larger than in Table 4.

The results in Column 2 of Table 7 are also very similar to those in Column 2 of Table 4. Only $\lambda$ is estimated a bit larger.

Finally, the results in Column 3 of Table 7 are statistically very imprecise. We do not have enough variation in the data to run an instrumental variable regression with so few observations. So the comparison of Columns 3 in Tables 4 and 7 is not very fruitful.

Overall, not accounting for pulsing does not seem to lead to any type of bias in our estimates.

9 Conclusions

The paper proposes a novel oligopoly model of advertising, based on persuasive advertising which shifts ("pulls up") perceived product qualities. The model also introduces comparative advertising as having both a pull up effect on own perceived quality, and a "pull-down" effect on a targeted rival’s quality. The empirical results for the non-comparative advertising are very clean. First, half of a comparative ad constitutes pure push, insofar it has the same effect on own perceived quality as half the dollar amount spent on a non-comparative ad. What happens to the other half is damage inflicted on the target of the comparative ad. First, it takes 40 cents of non-comparative ads to offset every incoming dollar of attack, and that is just in terms of the net persuasion part of advertising. The other part of the harm to a rival is in the Pull-Down effect, which involves a further loss (otherwise, if there were a 50 cents direct again, and only a 40 cent harm to a specific rival, of which the attacker can only expect a fraction of the fall-out, comparative advertising would not be deployed).

The (linearized) comparative ads estimates indicate that there is a strong positive effect of larger size in comparative advertising, and a much stronger (in terms of elasticity) positive effect of larger size of the target. This concurs with the theoretical predictions of the Push-Pull model, and it is apparent in the raw data that the largest target is the largest firm (Tylenol).

The effects of advertising in this 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 between brands to and fro just to stay afloat. This feature is reminiscent of the Zero-Sum Game critique of advertising; that it serves solely to reshuffle demand and firms are better off if they could agree not to do it (they would save the

---

67 One notable context of comparative ads is the case of negative political ads, which sometimes emphasize the negative features of rivals more than the positive features of the candidate on whose behalf the ad is aired.
expense). The critique is a fortiori true of comparative advertising, at least as modeled though the Push-Pull model. Firms would be better off if they could agree not to do it. This reason might partially explain why it is not prevalent in many industries. As a form of quasi-collusion, firms do not begin the process because they realize it might trigger responses. It is noteworthy in this regard that comparative advertising is being used more and more, coinciding with a recession, when collusion typically has more trouble surviving.

When consumers have different tastes over different characteristics, comparative advertising (done by different parties in different characteristics directions) may serve to enhance the perceived horizontal differentiation between products. This effect is closed down in the current model, but introducing it would likely give both better estimates as well as an improved perspective on the social benefits of the practice, at least insofar as the advertising informs heterogenous consumers about true product performance differences.

Thus far, we have not introduced into the estimation the fact that there are two two-brand firms in the sample. Bayer owns Aleve and its Aspirin, while McNeil owns Tylenol and Motrin: hence the corresponding zeroes in the Attack Matrix of Table 2. It is though straightforward to write the price and advertising first-order conditions for multiproduct firms, and these yield quite intuitive results. First, non-comparative ads per brand for a multi-product firm tend to be lower than for a single product firm with the same market share. This is because the multi-product advertiser internalizes the deleterious effect of increased persuasion on sibling brands. Second, comparative ads per brand for a multi-product firm tend to be higher than for a single product firm with the same market share. A multi-product firm benefits more from attacks because the lost customer base of the attacked product can be picked up by both its own brands. These effects are quite noticeable in the data.

References


[42] Liaukonyte, Jura (2009), Is comparative advertising an active ingredient in the market for pain relief? Mimeo, University of Virginia.


10 Appendix - Theory

10.1 Multiple products per firm

In the OTC analgesics industry, as indeed in most others, each firm sells several variants of its product (gelcaps, liquid; arthritis variants, children’s...). There are also multiple bottle sizes, ranging from 2-pill packs to sometimes 1000-pill bottles. Yet the analysis above has retained the fiction of a single price per brand. We now show that the same advertising result holds by the same technique of substituting out the pricing equations.

Suppose then that there were two variants of brand \( j \) (or indeed two bottle sizes: the argument that follows applies to any number and combination). Denote these with superscripts \( a \) and \( b \) for the two variants.

Then we would write the profit function (simplifying for the moment to suppress comparative ads):

\[
Max \ \pi_j = M(p^a_j - c_j)s^a_j + M(p^b_j - c_j)s^b_j - A_{jj}
\]

The advertising first order condition is

\[
\frac{d\pi_j}{dA_{jj}} = M(p^a_j - c_j)\frac{\partial Q^a_j}{\partial A_{jj}} + M(p^b_j - c_j)\frac{\partial Q^b_j}{\partial A_{jj}}
\]

Using the price first order conditions:

\[-M(p^a_j - c_j)\frac{\partial s^a_j}{\partial \delta^a_j} - M(p^b_j - c_j)\frac{\partial s^b_j}{\partial \delta^b_j} + Ms^j = 0
\]

and the analogous condition for \( p^b_j \) yields

\[
\frac{d\pi_j}{dA_{jj}} = Ms^a_j\frac{\partial Q^a_j}{\partial A_{jj}} + Ms^b_j\frac{\partial Q^b_j}{\partial A_{jj}} - 1 = 0.
\]

Under the (strong) assumption that \( \frac{\partial Q^a_j}{\partial A_{jj}} = \frac{\partial Q^b_j}{\partial A_{jj}} = Q'(A_{jj}) \) we have

\[
Ms_jQ'(A_{jj}) = 1,
\]

where we have defined \( s_j = s^a_j + s^b_j \), so the same ad relationship holds as when there is but a single product type (and correspondingly a single price to be chosen).

10.2 Multiple Consumer Types

The theoretical analysis still goes through with multiple consumer types provided one is prepared to make some assumptions. Indeed, if there are 2 consumer types (say headache sufferers, \( A \), and arthritis sufferers, \( B \)) then denote their demands by subscripts and write profits for firm \( j \) as Firm \( i \)'s profit-maximizing problem is:

\[40\]
Max \( \pi_j = (p_j - c_j) \left( N^A s_j^A + N^B s_j^B \right) - A_{jj} - \gamma \sum_{k \neq 1}^n A_{jk} \quad j = 1, \ldots, n, \)

The crucial assumption is that the quality indices have the same derivatives across types. In that case, in parallel to the single type analysis (inserting price conditions (4) into the advertising ones (5)) we now get the relation for positive advertising as:

\[
\left( N^A s_j^A + N^B s_j^B \right) \frac{\partial Q_j}{\partial A_{jj}} = 1, \quad j = 1, \ldots, n.
\]

with a similar replacement pertaining to comparative advertising. Thus it is not necessary that the quality indices be the same across types, but instead that their derivatives are. This is still a strong requirement though. To see why, consider comparative advertising: it may have a much different effect on arthritis sufferers than headache-prone ones if it stresses relative speeds of pain relief.

11 Appendix on Functional Forms

We shall also retain the restriction that there is perfect substitutability between non-comparative advertising and outgoing attack ads so that Push is just own enhancement. Suppose then that write the quality function as

\[
Q_j = g \left( A_{jj} + \lambda \sum_{k \neq j} A_{jk}, \{A_{kj}\}_{k \neq j} \right),
\]

where the second term is the vector of incoming attacks, the Pull effect.

First, suppose that we were to write the function \( g(\cdot) \) as additively separable in the Push and Pull effects. Then, the first-order condition for non-comparative ads, which is \( M s_j \frac{\partial Q_j}{\partial A_{jj}} \leq 1 \), with equality if \( A_{jj} > 0 \), (see (6)) implies that \( A_{jj} \) should depend only on own share, \( s_j \), and outgoing attacks, \( \sum_{k \neq j} A_{jk} \).

However, preliminary data analysis suggests that incoming attacks are also strongly significant in explaining non-comparative ads, \( A_{jj} \). This leads us to reject formulations in which the \( g \) function is additively separable into Push and Pull effects.\(^{68}\)

Second, perhaps the most natural way to write the function \( g(\cdot) \) is as a function of Net Persuasion,

\[
NP_j = A_{jj} + \lambda \sum_{k \neq j} A_{jk} - \omega \sum_{k \neq j} A_{kj}
\]

which is the (weighted) total Push minus the total Pull (with \( \lambda > 0 \) and \( \omega > 0 \)). This gives the desired property for the non-comparative advertising equation that it depends (and in a simple linear manner) on the incoming attacks. However, difficulties arise when formulating the comparative ad first-order conditions. Indeed, from (12) we have \( (0 < ) - M s_j \frac{\partial Q_j}{\partial A_{jj}} \leq \gamma - \lambda \). The problem arises once we realize that the pure Net

\(^{68}\)The same argument rules out a fully separable form where we also split up the Push effect into two separate parts, non-comparative ads and outgoing comparative ones.
Persuasion form above implies that we can write the term $\frac{\partial Q_j}{\partial A_{jk}}$ as $-\omega \frac{\partial Q_j}{\partial A_{sk}}$, and so, whenever both $j$ and $k$ have positive non-comparative ads (a frequent occurrence in the data) then (12) becomes $\frac{s_j}{1-s_j} \omega = \gamma - \lambda$. This clearly cannot hold. The problem, as described in detail in the Appendix, is that the solution violates the second order condition, and actually represents a minimum in the profit function.

It helps at this point to think of the second order conditions in terms of concavity and convexity of the Push and Pull functions. At given prices, the own share, $s_j$ ought to be increasing and (at least locally) concave in non-comparative ads. The first derivative (setting $\mu = 1$ to simplify) is $\frac{\partial s_j}{\partial A_{jj}} = s_j (1-s_j) \frac{\partial Q_j}{\partial A_{jj}}$, which is positive as desired as long as $Q_j$ increases in Push. The second derivative is $\frac{\partial^2 s_j}{\partial A_{jj}^2} = (1-2s_j) s_j (1-s_j) \frac{\partial Q_j}{\partial A_{jj}} + s_j (1-s_j) \frac{\partial^2 Q_j}{\partial A_{jj}^2}$. The first term here is positive (at least in the relevant range where shares are not excessive), and so it is necessary that the second term is negative. This implies that the quality function needs to be concave in Push.

Consider now the Pull effect. From (12) we have $-M \frac{s_j s_k}{1-s_j} \frac{\partial Q_k}{\partial A_{jk}} = \gamma - \lambda$ when $A_{jj} > 0$ and $A_{jk} > 0$ (and where $\frac{\partial Q_k}{\partial A_{jk}} < 0$). For a profit maximum, the LHS ought to be decreasing in $A_{jk}$. Following the analysis in the Appendix (see (23)), a necessary condition is:

$$\frac{s_j s_k}{(1-s_j) \partial A_{jk}^2} + \frac{\partial s_j s_k}{(1-s_j) \partial Q_k} \left( \frac{\partial Q_k}{\partial A_{jk}} \right)^2 > 0,$$

which, taking out a positive factor, becomes

$$\frac{\partial^2 Q_k}{\partial A_{jk}^2} + \left(1-s_k - \frac{s_k}{(1-s_j)} \right) \left( \frac{\partial Q_k}{\partial A_{jk}} \right)^2 > 0,$$

for which a sufficient condition (given the shares are not extremely large in equilibrium) is that $Q_j$ be convex in incoming attacks.

Now consider the quality function (14) above, i.e., Net Persuasion with an additional separable pull effect:\(^{69}\)

$$\bar{Q}_j = \alpha_D \ln \left( A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj} - \omega \sum_{k \neq j} A_{kj} \right) - \beta \sum_{k \neq j} \ln (A_{kj} + \bar{A}_{kj}).$$

Clearly this is concave in own non-comparative ads, as we required in the analysis above. Regarding

\(^{69}\) An alternative perceived quality function for product $j$, which has the desired properties, takes the form:

$$Q_j = \ln \left[ (w_j + A_{jj} + \lambda \sum_{k \neq j} A_{jk})^{\theta_1} + \sum_{k \neq j} (w_j + A_{kj})^{\theta_2} + \bar{A}_j \right]$$

where we retain the property that outgoing and non-comparative ads are perfect substitutes. For $Q_j$ to be increasing and concave in own non-comparative ads (and outgoing attacks), and decreasing and convex in incoming attacks, we need some parameter restrictions. First, $\theta_1 > 0$ (and it should not be much greater than 1) and $\theta_2 < 0$. Furthermore, we do not want $\theta_1 = 1$ or else the non-comparative ad equation will be independent of incoming attacks. Second, $w_j > 0$ to allow for firms to not advertise at all (otherwise, the ad derivative is infinite, at least for $\theta_1 < 1$ so firms would always want some ads). Also we want $w_2 > 0$ in order to get possible corner solutions for comparative ads.

The first order conditions for this model are highly non-linear though.
incoming attacks, the quality derivative is

\[ \frac{\partial Q_j}{\partial A_{kj}} = \frac{-\omega \alpha}{(A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj} - \omega \sum_{k \neq j} A_{kj})} - \frac{\beta}{(A_{kj} + \bar{A}_{kj})}, \]

which is decreasing, as desired. So consider the next derivative:

\[ \frac{\partial^2 Q_j}{\partial A^2_{kj}} = \frac{-\omega^2 \alpha}{(A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj} - \omega \sum_{k \neq j} A_{kj})^2} + \frac{\beta}{(A_{kj} + \bar{A}_{kj})^2}; \]

this sign is ambiguous, and depends on the sign of

\[ -\omega^2 \alpha_D (A_{kj} + \bar{A}_{kj})^2 + \beta \left( A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj} - \omega \sum_{k \neq j} A_{kj} \right)^2. \]

This will be positive, as desired, if \( \omega \) is small enough or \( \beta \) is large enough.

### 11.1 Second order conditions for pure Net Persuasion formulation

From the text, we have the first order condition for comparative ads, after substituting out the pricing first order condition and the non-comparative ads condition (when positive) as (see (12)):

\[ \frac{d\pi_j}{dA_{jk}} = \lambda - M \frac{s_j s_k}{(1 - s_j)} \frac{\partial Q_k}{\partial A_{jk}} - \gamma, \]

Consider now the sign of the derivative of the RHS w.r.t. \( A_{jk} \) (and henceforth set \( \mu = 1 \) for simplicity). It has the sign of

\[ -\left\{ \frac{s_j s_k}{(1 - s_j)} \frac{\partial^2 Q_k}{\partial A^2_{jk}} + \frac{\partial^2 Q_k}{\partial Q_k} \left( \frac{\partial Q_k}{\partial A_{jk}} \right)^2 + \frac{\partial^2 Q_k}{\partial Q_j \partial A_{jk} \partial A_{jk}} \right\}. \]  \( \text{(23)} \)

We are going to show the term in brackets is negative, so that the expression is positive, and thus indicates a minimum in the dimension of \( A_{jk} \).

The last term in the brackets has the sign of \( \frac{\partial^2 Q_k}{\partial Q_j \partial A_{jk} \partial A_{jk}} \), and so we shall first show that \( \frac{\partial^2 Q_k}{\partial Q_j \partial A_{jk} \partial A_{jk}} > 0 \).

This is true because \( s_j \) is increasing in \( Q_j \) and \( \frac{s_k}{(1 - s_j)} \) is independent of \( Q_j \) (for the Logit model: multiply denominator and numerator by the logit denominator, so the terms in \( Q_j \) disappear). Hence we have proved the first part.

We now want to show that

\[ \frac{s_j s_k}{(1 - s_j)} \frac{\partial^2 Q_k}{\partial A^2_{jk}} + \frac{\partial^2 Q_k}{\partial Q_k} \left( \frac{\partial Q_k}{\partial A_{jk}} \right)^2 < 0. \]

To do this, note that \( Q_k \) is decreasing in \( A_{jk} \) and concave; given the logarithmic form of \( Q(.) \), the second derivative is the negative of the first one squared.
Hence we need to show that
\[- s_j s_k (1 - s_j) + \frac{\partial s_j s_k}{\partial Q_k} < 0.\]

Now note that \(\frac{\partial s_j}{\partial Q_k} = \frac{1}{s_j - 1} (1 - s_j) \frac{s_j}{(1 - s_j)^2} = \left( -s_j s_k \right) - s_j (1 - s_j) \frac{s_j}{(1 - s_j)^2} \) and so \(\frac{\partial s_j s_k}{\partial Q_k} = \left( -s_j s_k \right) - s_j (1 - s_j) \frac{s_j}{(1 - s_j)^2} s_k + \frac{s_j}{(1 - s_j)^2} s_k (1 - s_j) =
\]

Hence the desired condition
\[- \frac{s_j s_k}{(1 - s_j)} + \frac{\partial s_j s_k}{\partial Q_k} < 0\]

is
\[\frac{s_j s_k}{(1 - s_j)} \left( -1 - \frac{s_k}{(1 - s_j)} + 1 - s_k \right) < 0\]

which is clearly true.

This means that the solution is a minimum.
<table>
<thead>
<tr>
<th>Brand</th>
<th>Active Ing.</th>
<th>Price per serving</th>
<th>Sales share*</th>
<th>Brand vol. share*</th>
<th>Weighted share**</th>
<th>Max pills***</th>
<th>TA/Sales</th>
<th>CA/Sales</th>
<th>CA/TA</th>
<th>Ownership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tylenol</td>
<td>ACT</td>
<td>$1.09</td>
<td>29.17%</td>
<td>36.85%</td>
<td>30.19%</td>
<td>7.63</td>
<td>25.38%</td>
<td>12.49%</td>
<td>41.00%</td>
<td>McNeil</td>
</tr>
<tr>
<td>Advil</td>
<td>IB</td>
<td>$0.86</td>
<td>17.20%</td>
<td>23.00%</td>
<td>23.89%</td>
<td>5.89</td>
<td>25.70%</td>
<td>20.18%</td>
<td>74.55%</td>
<td>Wyeth</td>
</tr>
<tr>
<td>Aleve</td>
<td>NS</td>
<td>$0.56</td>
<td>8.27%</td>
<td>10.72%</td>
<td>22.15%</td>
<td>3.00</td>
<td>34.09%</td>
<td>31.46%</td>
<td>88.73%</td>
<td>Bayer</td>
</tr>
<tr>
<td>Excedrin</td>
<td>ACT</td>
<td>$1.11</td>
<td>8.80%</td>
<td>11.03%</td>
<td>8.16%</td>
<td>8.77</td>
<td>27.87%</td>
<td>5.15%</td>
<td>23.08%</td>
<td>Novartis</td>
</tr>
<tr>
<td>Bayer</td>
<td>ASP</td>
<td>$1.28</td>
<td>5.73%</td>
<td>10.06%</td>
<td>6.87%</td>
<td>10.09</td>
<td>35.04%</td>
<td>14.65%</td>
<td>30.63%</td>
<td>Bayer</td>
</tr>
<tr>
<td>Motrin</td>
<td>IB</td>
<td>$0.85</td>
<td>8.03%</td>
<td>6.99%</td>
<td>7.57%</td>
<td>5.85</td>
<td>30.43%</td>
<td>17.01%</td>
<td>35.02%</td>
<td>McNeil</td>
</tr>
<tr>
<td>Generic</td>
<td>ACT</td>
<td>$0.58</td>
<td>8.01%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generic</td>
<td>IB</td>
<td>$0.36</td>
<td>9.25%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generic</td>
<td>ASP</td>
<td>$0.57</td>
<td>6.08%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generic</td>
<td>NS</td>
<td>$0.31</td>
<td>1.66%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: ACT-Acetaminophen, IB-Bupropfen, NS-Naproxen Sodium, ASP-Aspirin, TA-Tot ads, CA-Comparative ads

* Inside dollar share of branded products only
** Inside share of branded products weighted by the strength of pills
*** Average maximum number of pills within 24 hrs (determined by FDA)

Table 1. Brands, market share and advertising levels of OTC analgesics market

<table>
<thead>
<tr>
<th>T &gt; A</th>
<th>Advil</th>
<th>Aleve</th>
<th>Excedrin</th>
<th>Tylenol</th>
<th>Total Direct CA</th>
<th>Total CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advil</td>
<td></td>
<td>18.69</td>
<td>5.15</td>
<td>177.03</td>
<td>200.87</td>
<td>219.00</td>
</tr>
<tr>
<td></td>
<td>[26]</td>
<td>[20]</td>
<td>[56]</td>
<td></td>
<td>[102]</td>
<td></td>
</tr>
<tr>
<td>Aleve</td>
<td></td>
<td></td>
<td>0.48</td>
<td>131.86</td>
<td>132.14</td>
<td>157.00</td>
</tr>
<tr>
<td></td>
<td>[7]</td>
<td>[58]</td>
<td></td>
<td></td>
<td>[65]</td>
<td></td>
</tr>
<tr>
<td>Excedrin</td>
<td></td>
<td>2.94</td>
<td></td>
<td>20.94</td>
<td>23.87</td>
<td>27.70</td>
</tr>
<tr>
<td></td>
<td>[6]</td>
<td></td>
<td>[14]</td>
<td></td>
<td>[20]</td>
<td></td>
</tr>
<tr>
<td>Tylenol</td>
<td>9.60</td>
<td>31.64</td>
<td></td>
<td></td>
<td>41.24</td>
<td>116.00</td>
</tr>
<tr>
<td></td>
<td>[11]</td>
<td>[28]</td>
<td>[30]</td>
<td></td>
<td>[39]</td>
<td></td>
</tr>
<tr>
<td>Motrin</td>
<td>19.10</td>
<td>19.06</td>
<td></td>
<td></td>
<td>38.17</td>
<td>38.20</td>
</tr>
<tr>
<td></td>
<td>[25]</td>
<td>[25]</td>
<td>[50]</td>
<td></td>
<td>[50]</td>
<td></td>
</tr>
<tr>
<td>Bayer</td>
<td>13.78</td>
<td></td>
<td>15.69</td>
<td></td>
<td>29.47</td>
<td>40.30</td>
</tr>
<tr>
<td></td>
<td>[24]</td>
<td></td>
<td>[37]</td>
<td></td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>42.48</td>
<td>72.33</td>
<td>5.63</td>
<td>345.31</td>
<td>345.31</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The first row in each cell shows the expenditures in millions of dollars, the second row represents the total number of periods in which a specific attack occurred. Expenditures were divided by the number of targets if there were multiple targets. If a brand ad mentioned n rivals, then the expenditure attributed to each pair is counted as (1/n) of total expenditure of that particular ad. The column "Total Direct CA" represents the total expenditures on ads that mentioned actual rival brands, whereas "Total CA" figure also includes indirect comparative ads, such as comparisons to a certain class of products.

Table 2. Comparative advertising and target pairs

45
<table>
<thead>
<tr>
<th>News Shock</th>
<th>Date</th>
<th>Source</th>
<th>Major Shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ibuprofen Interferes with Aspirin</td>
<td>12/20/2001</td>
<td>New England Journal of Medicine, 2001, Volume 345:1809-1817</td>
<td>Yes</td>
</tr>
<tr>
<td>Aspirin May Prevent Pancreatic Cancer</td>
<td>8/6/2002</td>
<td>Journal of the National Cancer Institute 2002; 94:1168-71</td>
<td>No</td>
</tr>
<tr>
<td>F.D.A. Panel Calls for Stronger Warnings on Aspirin and Related Painkillers</td>
<td>9/21/2002</td>
<td>FDA Public Health Advisory</td>
<td>Yes</td>
</tr>
<tr>
<td>Misusing acetaminophen, other painkillers can be deadly, FDA warns</td>
<td>1/23/2004</td>
<td>FDA Public Health Advisory</td>
<td>No</td>
</tr>
<tr>
<td>Vioxx Withdrawn From the Market</td>
<td>9/30/2004</td>
<td>FDA Public Health Advisory</td>
<td>Yes</td>
</tr>
<tr>
<td>Use of naproxen (Aleve) associated with an increased cardiovascular (CV) risk</td>
<td>12/23/2004</td>
<td>FDA Public Health Advisory</td>
<td>Yes</td>
</tr>
<tr>
<td>Bextra Withdrawn</td>
<td>4/7/2005</td>
<td>FDA Public Health Advisory</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 3. Medical News shocks and their descriptions
Table 4: Non-Comparative Advertising

<table>
<thead>
<tr>
<th>IV</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical news as controls</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>MS&lt;sub&gt;j&lt;/sub&gt;</td>
<td>0.157***</td>
<td>0.123***</td>
<td>0.214***</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.033)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>Σ Comparative Ads</td>
<td>-0.497***</td>
<td>-0.514***</td>
<td>-0.532***</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.062)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>Incoming Attack Ads</td>
<td>0.347***</td>
<td>0.316***</td>
<td>0.438***</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.066)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>Top Brand</td>
<td>-0.471***</td>
<td>-0.292</td>
<td>-0.642***</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.114)</td>
<td>(0.228)</td>
</tr>
<tr>
<td>Control Function [MS&lt;sub&gt;j&lt;/sub&gt;]</td>
<td></td>
<td>-0.169*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.088)</td>
<td></td>
</tr>
<tr>
<td>Control Function [Σ Comparative Ads]</td>
<td></td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.110)</td>
<td></td>
</tr>
<tr>
<td>Control Function [Incoming Attack Ads]</td>
<td></td>
<td>-0.268*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.153)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.039</td>
<td>0.042</td>
<td>-0.078</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.060)</td>
<td>(0.128)</td>
</tr>
<tr>
<td>Likelihood Function</td>
<td>56.576</td>
<td>163.729</td>
<td>171.835</td>
</tr>
<tr>
<td>Median Elasticity S&lt;sub&gt;j&lt;/sub&gt;</td>
<td>2.098</td>
<td>1.652</td>
<td>2.874</td>
</tr>
</tbody>
</table>

Number of observations: 346; Left-Censored Observations: 47

note: *** p<0.01, ** p<0.05, * p<0.1

Columns 2 and 3 include the interactions of news shocks with brand dummies.

Their coefficient estimates are not reported for sake of brevity.

Figure 1:
### Table 5: Comparative Advertising (Linearized Version)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Medical news as controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>MS&lt;sub&gt;k&lt;/sub&gt;</td>
<td>0.073***</td>
<td>0.099***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>MSjS&lt;sub&gt;k&lt;/sub&gt;</td>
<td>6.139**</td>
<td>14.197***</td>
</tr>
<tr>
<td></td>
<td>(2.490)</td>
<td>(4.685)</td>
</tr>
<tr>
<td>Top Brand - Top Brand</td>
<td>-0.090*</td>
<td>-0.523***</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>0.087</td>
</tr>
<tr>
<td>Top Brand - Other Brand</td>
<td>(0.072)**</td>
<td>-0.081*</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>0.047</td>
</tr>
<tr>
<td>Other Brand - Top Brand</td>
<td>-0.17265***</td>
<td>-0.445***</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>0.077</td>
</tr>
<tr>
<td>Control Function [MSj]</td>
<td>-0.029</td>
<td>0.027</td>
</tr>
<tr>
<td>Control Function [MS&lt;sub&gt;k&lt;/sub&gt;]</td>
<td>-0.158***</td>
<td>0.029</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.29456***</td>
<td>-0.178***</td>
</tr>
<tr>
<td></td>
<td>.03517</td>
<td>0.048</td>
</tr>
<tr>
<td>Median elasticity S&lt;sub&gt;j&lt;/sub&gt;</td>
<td>1.144</td>
<td>2.647</td>
</tr>
<tr>
<td>Median elasticity S&lt;sub&gt;k&lt;/sub&gt;</td>
<td>5.089</td>
<td>8.053</td>
</tr>
</tbody>
</table>

Number of observations: 1160; Left-Censored Observations: 663

**Note:** *** p<0.01, ** p<0.05, * p<0.1

All columns include the interactions of news shocks with brand dummies. Their coefficient estimates are not reported for sake of brevity.
<table>
<thead>
<tr>
<th>Table 6: Adding Lags</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Medical news as controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$\text{MS}_j$</td>
<td>0.049</td>
<td>0.146*</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.077)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>$\Sigma$ Comparative Ads</td>
<td>-0.417</td>
<td>-0.418***</td>
<td>-0.408***</td>
</tr>
<tr>
<td></td>
<td>-0.06</td>
<td>(0.113)</td>
<td>(0.112)</td>
</tr>
<tr>
<td>Incoming Attack Ads</td>
<td>0.226</td>
<td>0.204</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.145)</td>
<td>(0.148)</td>
</tr>
<tr>
<td>Top Brand</td>
<td>-0.107</td>
<td>-0.418</td>
<td>-0.275</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.226)</td>
<td>(0.233)</td>
</tr>
<tr>
<td>Lag Outgoing Non-Comp Ad</td>
<td>0.323</td>
<td>0.300***</td>
<td>0.543***</td>
</tr>
<tr>
<td>Control Function [MSj ]</td>
<td>(0.054)</td>
<td>(0.055)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>Control Function [$\Sigma$ Comparative Ads]</td>
<td>-0.132</td>
<td>-0.093</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.085)</td>
<td>(0.085)</td>
<td></td>
</tr>
<tr>
<td>Control Function [Incoming Attack Ads]</td>
<td>0.006</td>
<td>-0.020</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.107)</td>
<td></td>
</tr>
<tr>
<td>Control Function [Lag Non-Comparative Ads]</td>
<td>-0.030</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.155)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.075</td>
<td>-0.063</td>
<td>-0.050</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.125)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>Likelihood Function</td>
<td>184.350</td>
<td>186.298</td>
<td>189.194</td>
</tr>
<tr>
<td>Median Elasticity $\bar{s}_j$</td>
<td>0.655</td>
<td>1.963</td>
<td>1.383</td>
</tr>
</tbody>
</table>

Number of observations: 346; Left-Censored Observations: 47

Note: *** p<0.01, ** p<0.05, * p<0.1

Columns 2 and 3 include the interactions of news shocks with brand dummies.

Their coefficient estimates are not reported for sake of brevity.
Table 7: Quarterly

<table>
<thead>
<tr>
<th>IV</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical news as controls</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MS&lt;sub&gt;i&lt;/sub&gt;</td>
<td>0.400***</td>
<td>0.329**</td>
<td>0.273</td>
</tr>
<tr>
<td>(0.102)</td>
<td>(0.139)</td>
<td>(0.249)</td>
<td></td>
</tr>
<tr>
<td>Σ Comparative Ads</td>
<td>-0.471***</td>
<td>-0.613***</td>
<td>-0.746</td>
</tr>
<tr>
<td>(0.099)</td>
<td>(0.094)</td>
<td>(0.177)</td>
<td></td>
</tr>
<tr>
<td>Incoming Attack Ads</td>
<td>0.376***</td>
<td>0.360***</td>
<td>0.413</td>
</tr>
<tr>
<td>(0.090)</td>
<td>(0.086)</td>
<td>(0.146)</td>
<td></td>
</tr>
<tr>
<td>Top Brand</td>
<td>-1.195***</td>
<td>-0.714</td>
<td>-0.507</td>
</tr>
<tr>
<td>(0.323)</td>
<td>(0.457)</td>
<td>(0.755)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.020</td>
<td>0.085</td>
<td>0.483</td>
</tr>
<tr>
<td>(0.150)</td>
<td>(0.222)</td>
<td>(0.275)</td>
<td></td>
</tr>
<tr>
<td>Control Function [1] Shares</td>
<td>0.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Function [2] Sumcompads</td>
<td>0.168</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Function [3] Incoming</td>
<td>-0.079</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood Function</td>
<td>-55.339</td>
<td>5.910</td>
<td>9.687</td>
</tr>
<tr>
<td>Median Elasticity S&lt;sub&gt;j&lt;/sub&gt;</td>
<td>1.906</td>
<td>1.572</td>
<td>1.30381</td>
</tr>
</tbody>
</table>

Number of observations: 112; Left-Censored Observations: 47

Note: *** p<0.01, ** p<0.05, * p<0.1
Columns 2 and 3 include the interactions of news shocks with brand dummies.
Their coefficient estimates are not reported for sake of brevity.

Figure 4:
<table>
<thead>
<tr>
<th>Year</th>
<th>Month</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>9</td>
<td>Early Vioxx/Celebrex safety concerns</td>
</tr>
<tr>
<td>2001</td>
<td>12</td>
<td>Ibuprofen counteracts Aspirin</td>
</tr>
<tr>
<td>2002</td>
<td>9</td>
<td>FDA panel calls for stronger warnings on NSAIDs</td>
</tr>
<tr>
<td>2003</td>
<td>3</td>
<td>Aspirin prevents colorectal adenomas</td>
</tr>
<tr>
<td>2003</td>
<td>4</td>
<td>Bextra withdrawal</td>
</tr>
<tr>
<td>2003</td>
<td>9</td>
<td>NSAIDs inhibit cardioprotective benefits of Aspirin</td>
</tr>
<tr>
<td>2004</td>
<td>10</td>
<td>Vioxx withdrawal</td>
</tr>
<tr>
<td>2004</td>
<td>12</td>
<td>Aleve is associated with increased cardio risk</td>
</tr>
</tbody>
</table>

Figure 1. Timelines of Advertising Expenditures, Market Shares and Medical New Shocks
Figure 2. Relationship between Non-Comparative Ads and Market Shares