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: March 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 75 cents spent on non-comparative (purely direct) advertising: the remainder is attributable to pulling down rivals, and there is strong evidence of damage to targets.

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

JEL CLassification

*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

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

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, 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.

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. The only way to gather data on the precise extent of the practice of comparative advertising is to actually watch the ads and to code them by content, and then match the result to advertising expenditures data. Coding targets too then yields an "Attack Matrix" of how much each brand spent in comparative advertising naming which rivals, along with how much was spent on non-comparative advertising.

Non-comparative advertising involves only positive promotion. A comparative advertisement, by comparing ones 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

[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.

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. We find indeed that comparative advertising is less effective at pushing up than is non-comparative advertising, while the pulling down effect does hurt rivals quite substantially. These results are broadly consonant with the push-pull model we propose to frame the strategic interaction in the market.

The push-pull model incorporates the features noted above. It 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.[3] By hurting the rival product directly, some consumers are diverted, and the comparative advertiser succeeds in attracting some portion of those consumers.

As we discuss below in the literature review, 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),[4] 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.

This approach is both novel (and simple) theoretically, and it gives clean relations for estimation. In particular, we use the equilibrium pricing (first-order) conditions to eliminate prices from the relation between advertising and sales. The benefit of the structural modeling approach is to use the theoretical framework and relate ad levels of the different ad types to market shares, by using pricing first order conditions at the aggregate brand level to substitute out prices and so 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. This also means that we circumvent having to estimate pricing equations, and we do not need to jointly estimate the demand function (which also allows us to retain some flexibility there – in particular,

[3] 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.

[4] Notwithstanding Lord Leverhulme's fine quote concerning its unpredictability: "Half of my advertising is useless. I just do not know which half."

the non-comparative advertising equation is quite general). This method is therefore quite useful for other studies. (For example, Gasmi et al. (1992) have to jointly estimate a specific demand function; Roberts and Samuelson (1988) simply assume prices are fixed since cigarette prices do not vary much over their sample period.)

On the theoretical side, the paper delivers the positive economics of advertising spending in terms of correlations with other observable market variables, like market shares. 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.[5] The current analysis presents stand-alone results on advertising/share relations which are independent of the comparative advertising analysis.

The next contribution of the paper is to introduce comparative advertising into the equilibrium marketing mix. To derive strong predictions, we use strong functional form assumptions. We base these predictions on the benchmark case of a logit approach, wherein comparative advertising pulls down the perceived quality of targeted rivals' products. From the equilibrium solution to this larger game we can derive testable hypotheses about the attack pattern of comparative advertising. In particular, this analysis suggests that larger firms are both more likely to be attacked, and to be attacked more by larger firms. However, because firms do not attack themselves, this does not necessarily imply that the largest firms attack the most.

The way in which advertising enters the model is most simply thought of as persuasive advertising that shifts demand up. 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.[6]

Before expanding on other types of advertising that might be consistent with our approach, it is worth first discussing what is probably not consistent. Our approach has advertising which typically does not increase competitive pressure between firms. If, by contrast, advertising were of horizontal match characteristics, then such advertising would pivot demands (see also Johnson and Myatt 2004): some consumers would be

[5] 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.

[6] 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 (19xx) ["A representative consumer theory..." IER], and introduce advertising into it.

prepared to pay more for what they are (post-advertising) more sure they like, while others are prepared to pay less for what they know they like less. That is, imparting horizontal information is a two-edged sword. This indeed is the type of horizontal match informative advertising considered by Anderson and Renault (2009).

A different type of informative advertising is awareness advertising that makes consumers aware of the existence of a product. Under this formulation, consumers do not know a product exists unless they get advertised information. Informing more consumers costs more, and one typically assumes that consumers are reached randomly. Hence different consumers know about different sub-sets of products (there are different overlaps of reach), and so firms end up competing in what are effectively multiple market segments. In this setting, if firms advertise more, then they tend to compete with more rivals for consumers, and so advertising tends to be pro-competitive. This is the setting analyzed theoretically by Grossman and Shapiro (1984) building on Butters (1977).[7] A similar formulation - informative reach advertising and with a logit demand system - forms the basis for the empirical model implemented by Goeree (2008) for the US Personal Computer industry. Our context of analgesics is rather different because all are available on the shelf in the drug-store whenever the consumer wants to buy a pain-killer. Nonetheless, the persuasive advertising model we use might be usefully augmented by using a reach function to allow for different consumers to receive different amounts of persuasion from different firms.

Let us now return to our model, and the other possible interpretations of our formulation. We noted above that the model is not well construed as one of horizontal match information (whereby some consumers like features and others dislike them); but it could reasonably be one of vertical match information. That is, advertising can impart information on product characteristics that appeal to consumers (see the further discussion of persuasion games in the section on modelling comparative advertising below).

The signalling explanation of advertising can also be picked up within the current framework: insofar as higher advertising levels are a signal of higher actual quality, the perceived quality can be proxied by advertising levels. However, since the signalling model typically relies on repeat purchases to motivate firms to advertise, this suggests substantial long-term effects of advertising on sales, which we do not allow for in our static approach. The same critique applies to the vertical information back-drop suggested above; unless we accept that consumers are quite "forgetful," and indeed, our current formulation even if cast as prestige regards it as quite ephemeral. This is an extreme case, although some studies have suggested relatively short term effects (see Bagwell 2008 for a review of some evidence.) We undertake some sensitivity analysis that sheds light on this assumption.

Finally, we note that medical news shocks as we enter them in the model can be viewed as contributing to perceived qualities. From our profit maximizing problem we derive first order conditions for price and

[7] Cristou and Vettas (2008) analyse a non-localized discrete choice version of the Grossman-Shapiro model.

advertising mix levels. We are able to substitute the price first order condition to arrive at linear equations which we later bring to the data. We argue that generic prices of equivalent branded drugs are good approximations of marginal costs and use them as instruments in the estimation.

There is quite a lot of literature estimating the effects of advertising on consumer demand, although not all of it satisfactorily addresses endogeneity concerns. Most of the literature uses individual consumer demand data. There are only two papers that estimate firm choice of advertising directly from a structural model, and (like us) neither has individual consumer demand data. These two key papers are Roberts and Samuelson (1988) and Goeree (2008). Our paper and Roberts and Samuelson (1988) have in common the market expansion effect and a share effect, though we do not model the possibility that rivals' demands can rise with own advertising.

Ackerberg (2001, 2003) argues that the observed facts that "experienced" consumers (those who have previously bought the new product, 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. Dube et al. has in common with Roberts and Samuelson (1988) a dynamic theoretical model of ads, though their model is only concerned with estimating the demand side (the estimates are then used to calibrate the theoretical dynamic model).

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 empirical specification and identification of model in Sections 5 and 6. Section 7 discusses results and Section 8 concludes.

2 Literature Review

2.1 Theoretical Economics Literature on Advertising

Much of the economic theory of advertising has been concerned the mechanism by which advertising affects choice, and the welfare economics of the market outcome (see Bagwell (2007) for a fine and comprehensive survey). Moreover, much work has considered very particular market structures, most often monopoly.[8]

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

[8] 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.

thought to decrease social welfare by deterring potential competition and creating barriers for new entrants.

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)).[9] 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.

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.[10] Kihlstrom and Riordan (1984) show a role for dissipative advertising in a perfectly competitive model and . 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 (2001) show that under duopoly there must always be dissipative advertising by the high quality firm if qualities are similar enough.

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 way for advertising to affect directly consumer well-being, and so gives a way of thinking about persuasive advertising. Another tack to thinking about persuasive advertising is propounded by Dixit and Norman (1978), who view 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).[11] Regardless, they suggest that there is a tendency for too much 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

[9] Indeed, informative advertising can reduce consumers' search costs to learn about the existence of products, their prices, qualities, and specifications.

[10] 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.

[11] 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.

approach is broadly consistent with advertising as a demand shifter.

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.

Anderson and Renault (2009) model advertising as purely and directly informative revelation of horizontal match characteristics of products.[12] Revelation of such information increases product differentiation, although this does not always increase firm profits, as explained below. Comparative advertising in this context is modeled as revelation of characteristics (match information) of the rival product along with own characteristics.

In the AR (2009) framework, there are two firms, and they potentially differ by a starting advantage one may hold over the other (a demand shifter), which might be interpreted as a perceived quality advantage. This is taken as parametric. Consider first firms that are roughly symmetric in starting perceived quality, which is equivalent to considering firms of similar size (market share). Then each firm has a unilateral incentive to advertise to increase product differentiation: consumers recognize better matches and so firms rationally anticipate raising prices to take advantage of more dedicated consumers. In this context there is no role for comparative advertising (of the type described) because each firm already wants to advertise its own characteristics itself.

However, matters are very different when firms are asymmetric: refer to the larger one as the one with the higher (perceived) quality. 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

[12] Anderson and Renault (2009) build 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.

smaller one. Hence, comparative advertising is carried out by the smaller firm against its larger rival, and arises if firms are different enough.

Our model arrives at a similar finding though we treat advertising role differently. In this paper 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 not immediately evident how the AR results extend to more firms, except insofar as an industry of 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 problem with 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.

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.

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.

Shy (1995) argues that in the case of differentiated products, comparative advertising informs consumers about the difference between the brand they have purchased in the past and their ideal brand. The model explains only the 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.

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 Economics Literature on Advertising

Recently, a number of papers have built empirical models in order to gain insight into the advertising effects. Ackerberg (2003) estimates a model for Yoplait yogurt that distinguishes between advertising with informative and persuasive effects, and finds a large and significant informative effect of advertising and an insignificant prestige effect. However, his results are primarily driven by the observations of new product introductions and not dealing with the effect of advertising on demand for established brands. Assuming an informative role of advertising and utilizing highly disaggregated panel data-set, Erdem and Keane (1996) estimate a dynamic model where consumers learn about the quality of laundry detergents through past experience and advertising expenditure. However, they do not endogenously model competitors' strategic interactions. Anand and Shachar (2004) provide evidence that advertising enables consumers to better match their respective tastes with the advertised TV show attributes. Their model is estimated with individual level data, and endogeneity of choice can be taken care of by including individual fixed effects.[13]

Shum (2004) uses scanner panel data-set to show that advertising counteracts the tendencies toward repeat purchasing due to brand loyalty. Dube, Hitsch and Manchanda (2005) develop a model of dynamic advertising competition, and apply it to the problem of optimal advertising scheduling through time. They find evidence for an advertising threshold effect, which is qualitatively similar to the S-shaped advertising response function. The response function identification relies on treating the residual promised Gross Rating Point (GRP) levels as exogenous (observed at the end or the beginning of each period). A related endogeneity problem in their model concerns the relationship between advertising and the unobserved shocks to goodwill depreciation. They assume that firms observe the realization of those shocks before they make their decisions. Manchanda, Dube, Goh and Chintagunta (2006), measure the impact of banner advertising on current customers' probabilities of repeat buying while accounting for duration dependence. Their model controls for unobserved individual differences by specifying a distribution over the individual customer advertising

[13] It would be interesting to test whether advertising serves as a mechanism to match consumers with their most preferred products. However we are not able to do that due to the aggregate nature of our data. Additionally, it might be the case that advertising itself alters the valuation of the advertised characteristics, making the appreciation of product characteristics endogenous.

response parameters. Erdem, Keane and Sun (2006) use Nielsen scanner panel data on four categories of consumer goods to examine how TV advertising and other marketing activities affect the demand curve facing a brand. They find that advertising is profitable not because it lowers the elasticity of demand for the advertised good, but because it raises the level of demand for the vast majority of the analyzed brands.

All of the above mentioned models either observe individual purchases over time or rely on other type of panel data. Our data is more similar to that of Goeree (2008) where only aggregate sales and advertising data are observed. Goeree (2008) studies the effect of advertising expenditure on product awareness in the personal computer market and finds that assumptions about the choice set size are critical in estimating price and advertising elasticities as well as firm mark-ups. We assume that all consumers are aware of the existence of all available OTC analgesics in the market: all packaged consumer goods within a certain category (e.g. pain relievers) are usually displayed in the same area within the aisle, and usually within the same shelf.

Most of the previous empirical studies on comparative advertising have focused on cross-industry analysis, bundling together comparative advertising across different industries with diverse market structures (Chou et al. (1987) and Harmon et al. (1983)).

2.4 Marketing Literature

The Marketing literature documents comparative advertising and analyzes its effectiveness. Marketing researchers distinguish between two types of comparative advertising: direct and indirect. Direct comparative ads involve mentioning the competitor or explicitly revealing the competitor's brand image (e.g. "Advil is faster and stronger than Tylenol"), whereas indirect comparative advertising contains just a generic comparative claim (e.g. "Probably the best beer in the world" or "Tylenol is safer than other regular nonprescription pain relievers"). Muehling et al. quote that around 40% of all advertising is comparative. Pechmann and Stewart (1990) code TV commercials and suggest that the majority of all ads are indirectly comparative (60% vs. 20% which contain direct comparative claims, and the rest are non-comparative).[14]

Behavioral studies in marketing suggest that consumers pay more attention to, and are generally more aware of, products after viewing comparative advertising relative to generic advertising (Grewal et al. (1997)). A lot of behavioral studies find that comparative advertising provides advantages that are not associated with non-comparative advertising. Wilkie and Farris (1974) originally proposed that comparative ads may be perceived as more relevant, especially when well-known brands are used as competitive references. They argue that the comparative format is a recognizable cue which may trigger processing activity. Similarly, Wilson and Muderrisoglu (1980) suggest comparative ads produce greater mental activity. Belch (1981), Stutts (1982) and Swinyard (1981) study the cognitive response to comparative ads and provide further

[14] According to the attendees of the annual NAD conference devoted to comparative advertising (http://www.narcpartners.org/events/conference/event.aspx). approximately 2/3 of advertising is directly or indirectly comparative.

evidence on this point. Muehling et al. (1990) find that comparative ads generate more message recall than a similar non-comparative ads.

Chevins (1975) and Giges (1980a, 1980b) see comparative ads as industry's own form of consumerism and suggest they appeal to a wider audience and enhance sponsor's brand identification, message persuasiveness and market share. Diamond (1978), Krakowiecki (1977), Levine (1976), Neiman (1987) and Phillips (1980) find that comparative ads generate undesirable outcomes: increase consumer awareness of competitors' brands, decrease claim credibility, and produce confusion.

The empirical evidence of role and effects of comparative advertising role is conflicting. Whether these findings imply that comparative advertising in fact affects demand in a different way than non-comparative advertising, is not clear. Rose et al. (1993) points out that it is difficult to differentiate between whether consumers are making better informed decisions or are simply more persuaded by comparative advertising. Existing empirical studies on comparative advertising are usually of a behavioral and experimental nature, and they ignore the consequences of the rival brand's use of comparative advertising. The exception is Liaukonyte (2008) who models the demand side of OTC analgesics and finds that comparative advertising increases the perceived quality of a product more than non-comparative advertising. Additionally, comparative advertising is found to be very effective in denigrating the perceived quality of a targeted brand in turn cannibalizing competitor's market share. However, Liaukonyte (2008) does not take into account the supply side of the market and interdependent strategic firm decisions in price and advertising setting. Our model does exactly that.

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 $j = 1, ...n$ charges price p_j and has perceived quality $Q_j(.)$, $j = 1, ...n$. We retain the subscript j on $Q_j(.)$ because when we get to the 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' 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 $Q_j(.)$ 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.[15]

The demand side is generated by a discrete choice model of individual behavior where each consumer buys one unit of her most preferred good. Then preferences are described by an indirect utility function:

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

in standard fashion, where

$$\delta_j = Q_j(.) - p_j \tag{2}$$

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 μ expresses the degree of horizontal consumer/product heterogeneity.[16]

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.[17]

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

[16] 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 α) 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 α.

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

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:

$$\underset{\{p_j, A_j\}}{Max}\, \pi_j = M(p_j - c_j)s_j - A_{jj} - \gamma \sum_{k \neq j} A_{jk} \qquad j = 1, ...n. \tag{3}$$

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.[18]

The advertising quantities (the A's) are dollar expenditures.[19] 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 values of consumers in OTC advertising from other sectors.[20]

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

[18] Hosp (2007) from Goodwin Procter LLP in his publication "Weighing the Risks: A Practice Guide to Comparative Advertising" 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 $\mathbb{P}$, and that when withdrawn it must be replaced with an ad of lower effectiveness, and the effectiveness is a fraction β 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 $p^A\left((1 - s_j) + s_j/\beta\right)$. If we normalize the cost of the non-comparative advertising by setting $p^A = 1$, then we have the effective comparative ad cost as $\gamma = ((1 - s_j) + s_j/\beta) > 1$.

[19] 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.

[20] For example, suppose that each ad aired at a particular time on a particular channel cost $\hat{p}$ and delivered H "hits" (where the hit is measured in dollars). Then the equilibrium price of an ad delivering $H/2$ hits should be $\hat{p}/2$, 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 heterogenous 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 Firms' Optimal Choices

3.3.1 Pricing

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

$$\frac{d\pi_j}{dp_j} = Ms_j - M(p_j - c_j)\frac{ds_j}{d\delta_j} = 0, \qquad j = 1, ...n, \tag{4}$$

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

3.3.2 Non-comparative Advertising

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

The non-comparative advertising expenditures come from:[21]

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

where the partial derivative function $\frac{\partial Q_j}{\partial A_{jj}}$ may depend on any of the arguments of $Q_j(.)$. The first-order conditions (4) and (5) can be usefully combined to give the desired equilibrium conditions:[22] substituting the pricing condition (4) into the advertising one (5) yields the relation:

$$Ms_j\frac{\partial Q_j}{\partial A_{jj}} \leq 1, \text{ with equality if } A_{jj} > 0, \qquad j = 1, ..., n. \tag{6}$$

The interpretation is the following. Raising A_{jj} by \$1 and raising price by \$$\frac{\partial Q_j}{\partial A_{jj}}$ too leaves δ_j unchanged. This change therefore increases the revenue by \$$\frac{\partial Q_j}{\partial A_{jj}}$ on the existing consumer base (i.e., Ms_j consumers). This extra revenue is equated to the \$1 marginal cost of the change, the RHS of (6).

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.[23] 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

[21] 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.

[22] If $\frac{\partial Q}{\partial A_{jj}}$ 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.

[23] In this case, $Ms_jQ'(A_{jj}) = 1$, 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.

of the marginal benefit from advertising (this is the case for our parameter $\bar{W}_j$ in the econometric specification below in Section 5), then one might expect that firms with higher such quality will be those advertising more. This indeed can be shown to be the case in some central specifications of the model (for example, in terms of the parameters below, if $\alpha = 1$, and using a logit demand with the "fully separable" specification in Section 5.1 below).

The same relation holds, given some strong separability properties on $Q_j(.)$, which we embody in the functional forms we use below (Section 5.1) that $\frac{\partial Q_j}{\partial A_{jj}}$ is independent of outgoing and incoming comparative advertising.[24] In summary:

Proposition 1 ***(Non-Comparative Advertising levels)*** *Let $Q_j(.)$ 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 $A_{jj} = 0$. ■

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:

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

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

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

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).

[24] Under some conditions on the parameters $\bar{A}_{jj}$. Top and smaller values.

Second, the own effect is readily derived as:[25]

$$\frac{ds_j}{d\delta_j} = \frac{s_j(1-s_j)}{\mu}, \qquad j = 0, ...n, \tag{9}$$

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

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

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

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

with equality if $A_{jk} \;\; > \;\; 0$.

Inserting the price first-order conditions (10) gives (for $k = 1, ..., n, \quad j = 1, ...n, \;\; k \neq j$):[26]

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

The relation between market share and comparative advertising takes a particularly clean form when the quality function embodies a perfect substitutability relation, and this includes the semi-separable form used below in estimation. Suppose therefore that the quality function can be written as $Q_j(A_{jj}, \{A_{jk}\}_{k\neq j}, \{A_{kj}\}_{k\neq j}) = Q_j(A_{jj} + \lambda\Sigma_{k\neq j}A_{jk}, \{A_{kj}\}_{k\neq j})$, $j = 1, ..., n$, where $0 < \lambda < 1$ 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 $Ms_j\frac{\partial Q_j}{\partial A_{jk}} = \lambda$, and hence, using equation (11), we can write:

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

[25] 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.

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

$$Ms_j\frac{\partial Q_j}{\partial A_{jj}} = 1, \qquad j = 1, ..., n.$$

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

$$\frac{\frac{\partial Q_j}{\partial A_{jk}}}{\frac{\partial Q_j}{\partial A_{jj}}} - \frac{s_k}{(1-s_j)}\frac{\frac{\partial Q_k}{\partial A_{jk}}}{\frac{\partial Q_j}{\partial A_{jj}}} = \gamma, \;\; k = 1, ..., n, \quad j = 1, ...n, \quad k \neq j.$$

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.

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 δ_k is attained). Such a rival price change (which j thus effectuates through comparative advertising) causes j's market share to rise by $\frac{s_j s_k}{\mu}$. 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.

Proposition 2 ***(Larger target more)*** *Let the quality function be $Q_j(A_{jj} + \lambda \Sigma_{k \neq j} A_{jk}, \{A_{kj}\}_{k \neq j})$, with $Q_j\,(.)$ 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, $j, k = 1, ..., n$. 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. For any given target k, note that the ratio $\frac{s_j}{(1-s_j)}$ on the LHS 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 $(p_j - c_j)\frac{ds_j}{d\delta_j}$ is small enough that (12) holds with strict inequality when $\frac{\partial Q_k}{\partial A_{jk}}$ is evaluated at $A_{jk} = 0$. ■

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.)

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

Proposition 3 ***(Larger targeted more)*** *Let the quality function be $Q_j(A_{jj} + \lambda \Sigma_{k \neq j} A_{jk}, \{A_{kj}\}_{k \neq j})$, with $Q_j\,(.)$ 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, $j, k = 1, ..., n$. Then, considering attacks from firms with positive levels of non-comparative advertising, in equilibrium, larger firms suffer more attacks from each rival.*

Proof. 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. ■

Before turning to the econometric specifications, we first discuss the data.

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.[27]

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 any OTC pain reliever sold in the U.S. national market during the 5 years from March of 2001 through December of 2005 (a total of 58 monthly observations).[28] 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,[29] 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.

[27] 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.
[28] 4-week product level data was normalized to monthly frequency to match the advertising data frequency.
[29] Source: Centers for Disease Control and Prevention, http://www.cdc.gov/

Brand	Active Ing.	Price per serving	Sales share*	Brand vol. share*	Weighted share**	Max pills***	TA/Sales	CA/Sales	CA/TA	Ownership
Tylenol	ACT	$1.09	29.17%	36.85%	30.19%	7.63	25.38%	12.49%	41.00%	McNeil
Advil	IB	$0.86	17.20%	23.00%	23.89%	5.89	25.70%	20.18%	74.55%	Wyeth
Aleve	NS	$0.56	8.27%	10.72%	22.15%	3.00	34.09%	31.46%	88.73%	Bayer
Excedrin	ACT	$1.11	8.80%	11.03%	8.16%	8.77	27.87%	5.15%	23.08%	Novartis
Bayer	ASP	$1.28	5.73%	10.06%	6.87%	10.09	35.04%	14.65%	30.63%	Bayer
Motrin	IB	$0.85	5.83%	6.99%	7.57%	5.85	30.43%	17.01%	35.02%	McNeil
Generic	ACT	$0.58	8.01%							
Generic	IB	$0.36	9.25%							
Generic	ASP	$0.57	6.08%							
Generic	NS	$0.31	1.66%							

Notes: ACT-Acetaminophen, IB-Ibuprofen, NS-Naproxen Sodium, ASP-Aspirin, TA-Total 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

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 3** 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.[30]

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

T > / A v	Advil	Aleve	Excedrin	Tylenol	Total Direct CA	Total CA
Advil	-	18.69 [26]	5.15 [20]	177.03 [56]	**200.87 [102]**	**219.00**
Aleve	-	-	0.48 [7]	131.66 [58]	**132.14 [65]**	**157.00**
Excedrin	-	2.94 [6]	-	20.94 [14]	**23.87 [20]**	**27.70**
Tylenol	9.60 [11]	31.64 [28]	-	-	**41.24 [39]**	**116.00**
Motrin	19.10 [25]	19.06 [25]	-	-	**38.17 [50]**	**38.20**
Bayer	13.78 [24]	-	-	15.69 [37]	**29.47 [61]**	**40.30**
Total	**42.48 [60]**	**72.33 [85]**	**5.63 [27]**	**345.31 [165]**		

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

4.3 News Shocks

Between 2001 and 2005, the OTC analgesics market endured several major medical news related “shocks”. The most notable 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).

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 a "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

reason. However, we have effectively ignored these multi-product firm relations in the data.

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).

News Shock	Date	Source	Major Shock
Risk of Cardiovascular Events Associated With Selective COX-2 Inhibitors	8/21/2001	Journal of the American Medical Association, 2001; 286:954-959.	Yes
Ibuprofen May Prevent Alzheimer's	11/8/2001	Nature, 8 November 2001	No
Ibuprofen Interferes with Aspirin	12/20/2001	New England Journal of Medicine, 2001, Volume 345:1809-1817	Yes
Aspirin May Prevent Prostate Cancer	3/12/2002	Mayo Clinic Proceedings 2002 Mar; 77(3): 219-225.	No
Aspirin May Prevent Pancreatic Cancer	8/6/2002	Journal of the National Cancer Institute 2002; 94:1168-71	No
F.D.A. Panel Calls for Stronger Warnings on Aspirin and Related Painkillers	9/21/2002	FDA Public Health Advisory	Yes
Aspirin Prevents Colorectal Adenomas	3/6/2003	New England Journal of Medicine, 2003, Volume 348:891-899.	Yes
Aspirin Could Reduce Breast Cancer Risk	4/8/2003	Journal of the American Medical Association, 2004;291:2433-2440.	No
NSAIDs May Offer Protection Against Alzheimer's Disease	4/2/2003	American Academy Of Neurology (2003, April 2)	Yes
Anti-Inflammatory Pain Relievers Inhibit Cardioprotective Benefits of Aspirin	9/9/2003	Circulation, 9/9/2003	Yes
Misusing acetaminophen, other painkillers can be deadly, FDA warns	1/23/2004	FDA Public Health Advisory	No
Elevated risk of acute myocardial infarction associated with Vioxx	4/19/2004	Circulation. 2004;109:2068-2073	No
Vioxx Withdrawn From the Market	9/30/2004		Yes
Use of naproxen (Aleve) associated with an increased cardiovascular (CV) risk	12/23/2004	FDA Public Health Advisory	Yes
Bextra Withdrawn	4/7/2005		Yes

Table 3. Medical News shocks and their descriptions

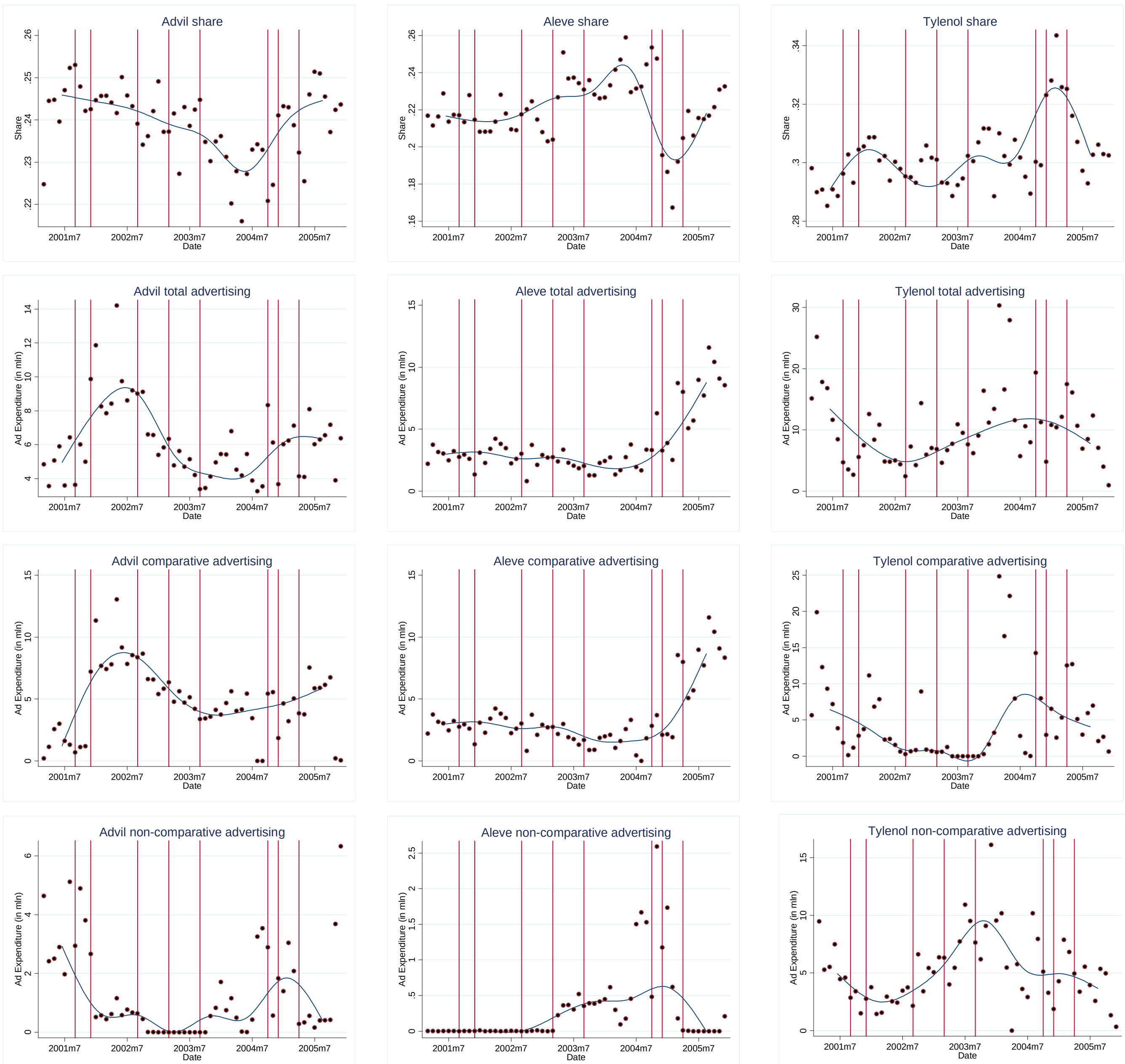
For each *major* shock that happened during period t we constructed a dummy variable which is equal to 1 in *all* the periods after and including t: t, $t+1$, ..., T.[31]. In the empirical analysis below, we interacted each of the major shocks listed in **Table 3** with brand dummies. This approach enables 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.[32] We assume in the model below, that news shocks surprise both consumers and firms and they

[31] 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.

[32] 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.

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.



Year	Month	Description	Year	Month	Description
2001	9	Early Vioxx/Celebrex safety concerns	2003	9	NSAIDs inhibit cardioprotective benefits of Aspirin
2001	12	Ibuprofen counteracts Aspirin	2004	10	Vioxx withdrawal
2002	9	FDA panel calls for stronger warnings on NSAIDs	2004	12	Aleve is associated with increased cardio risk
2003	3	Aspirin prevents colorectal adenomas	2005	4	Bextra withdrawal

Figure 1. Timelines of Advertising Expenditures, Market Shares and Medical New Shocks

5 Econometric Analysis

Here we derive the two special cases upon which we base the structural empirical analysis.[33] We first recall that non-comparative ads are given by (6):

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

with equality when $A_{jj} > 0$.

Similarly, whenever comparative advertising is positive we have from (11) that

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

with equality when $A_{jk} > 0$.

These are the equations we want to test under various specifications of Q. In what follows we draw out the implications for two different formulations for $Q_j(.)$, both mixes of logs and perfect substitutes. If we enter comparative and non-comparative ads as perfect substitutes (in the sub-valuation function, $Q_j(.)$) then we can use the non-comparative advertising first-order condition (11) to simplify the comparative advertising one. However, then we also find we will need to instrument for own comparative ads in the non-comparative advertising equation. We start with this case, although when we come to the estimation, we present the cases in the opposite order.

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

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

where only $\bar{Q}_j(.)$ 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. We will consider two alternative specifications for $\bar{Q}_j$.

[33] It is important to reiterate some key elements in the structure of the model which tie together to get these strong results. In particular, we have used a one-stage game (ads and prices are set simultaneously), a discrete-choice-based demand structure, and simple forms for Q_j below. Nonetheless, these results form a benchmark case, and are noteworthy for generating strong and simple predictions, which are picked up in the empirical investigation.

5.1 Fully separable specification

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

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 advertising (of the type denoted). For example, the higher is $\bar{A}_{jj}$, the lower is the marginal efficiency of non-comparative advertising. In the econometric specification, both types of variables will depend on some observed variables (for example news shocks) as well as some random shocks. The last terms are attacks, and have decreasing marginal impact: an attack from several directions hurts more than the same spending from one direction. (Think consumer perception.) Below we refer to the $\bar{W}$ variables as base quality, while the $\bar{A}$ variables are called advertising base allure.

The first order conditions are quite straightforward. Each type of ad has its own special kick, though the combination of $\bar{A}_{jk}$ and α_I may mean that some are not deployed because they are not very effective. We have the interpretation of the parameters as elasticities of the sub-parts of Q, which is not that useful, and we also have them as the demand elasticities, as per the semi-separable analysis below.

The non-comparative advertising equation (6) becomes

$$A_{jj} = \max\left\{-\bar{A}_{jj} + \alpha_D M s_j, 0\right\} \tag{13}$$

This version implies that non-comparative advertising is not directly related to other firms' market shares, nor to comparative ad levels.

The comparative advertising equation (11) becomes (after some algebra):

$$A_{jk} = \max\left\{-\bar{A}_{jk} + M s_j \frac{\alpha_I}{\gamma} + M \frac{s_j s_k}{1 - s_j} \frac{\beta}{\gamma}, 0\right\}. \tag{14}$$

This therefore basically runs comparative advertising on a constant, own market share, and $\frac{s_j s_k}{1-s_j}$ (and allows us to estimate $\frac{\alpha_I}{\gamma}$ and $\frac{\beta}{\gamma}$: the non-comparative advertising equation allows us to estimate α_D, but we cannot identify λ).

Proposition 1 says that α_D should be positive, so that for firms with the same $\bar{A}_{jj}$, the higher market share goes together with the higher advertising level. Proposition 3 suggests that $\frac{\beta}{\gamma}$ should be positive - although this specification does not satisfy the assumptions of the Proposition, the same argument goes through.

5.2 Semi-separable specification

This is the following:

$$\bar{Q}_j(.) = \alpha_D \ln\left(A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj}\right) - \beta \sum_{k \neq j} \ln\left(A_{kj} + \bar{A}_{kj}\right). \tag{15}$$

This has comparative and non-comparative ads as perfect substitutes in the own push up effect (the attack feature of comparative advertising comes in the Pull-Down part on others).

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

$$\left(A_{jj} + \lambda \sum\nolimits_{k \neq j} A_{jk} + \bar{A}_{jj}\right)^{\alpha_D/\mu} \Pi_{k \neq j} \left(A_{kj} + \bar{A}_{kj}\right)^{-\beta/\mu} \exp\left(\bar{W}_j - p_j/\mu\right),$$

so that the β parameter is loosely (μ times) the elasticity of demand with respect to incoming attacks (loosely because this is just the denominator of demand, and also it is more specifically the elasticity with respect to $A_{kj} + \bar{A}_{kj}$). In terms of perceived quality, $Q_j(.)$, β is (minus) the elasticity of the incoming comparative advertising attack from k. The parameter λ tells us how substitutable are own outgoing ads for non-comparative ads, and is a key parameter of interest below.

The first equation (6) gives:

$$\frac{\alpha_D M s_j}{\left(A_{jj} + \lambda \sum_{k \neq j} A_{jk} + \bar{A}_{jj}\right)} \leq 1,$$

with equality if $A_{jj} > 0$.

Therefore we run the non-comparative ad equation as:

$$A_{jj} = \max\left\{\alpha_D M s_j - \lambda \sum\nolimits_{k \neq j} A_{jk} - \bar{A}_{jj}, 0\right\}. \tag{16}$$

We have to instrument for own outgoing ads. These equations enable us to determine both the α_D and λ parameters.

The comparative advertising equations, (11), noting that $\frac{\partial Q_j}{\partial A_{jk}} = \lambda \frac{\partial Q_j}{\partial A_{jj}}$ when $A_{jj} > 0$, becomes simply $\lambda - M \frac{s_j s_k}{1-s_j} \frac{\partial Q_k}{\partial A_{jk}} - \gamma \leq 0$, or

$$A_{jk} = \max\left\{-\bar{A}_{jk} + M \frac{s_j s_k}{1 - s_j} \frac{\beta}{\gamma - \lambda}, 0\right\} \tag{17}$$

for $A_{jj} > 0$. This means we can find the parameter $\frac{\beta}{\gamma - \lambda}$, but we cannot find β or γ alone.

5.2.1 Total Comparative Advertising

We can also add these up over all attack targets: to get the total comparative ad spending by j as:

$$\sum\nolimits_{k \neq j} A_{jk} = -\sum\nolimits_{k \neq j} \bar{A}_{jk} + M \frac{\beta}{\gamma - \lambda} \frac{s_j(1 - s_j - s_0)}{1 - s_j}.$$

One important feature of this last equation is that it increases and then decreases in firm size: there is a hump. Hence the push-pull model of advertising predicts a non-monotone inverted U relation between size and comparative advertising. The next figure sets $s_0 = 3/5$, and plots comparative advertising against firm size:[34]

[34] The plot is of $\frac{x}{(1-x)}\left(\frac{3}{5} - x\right)$.

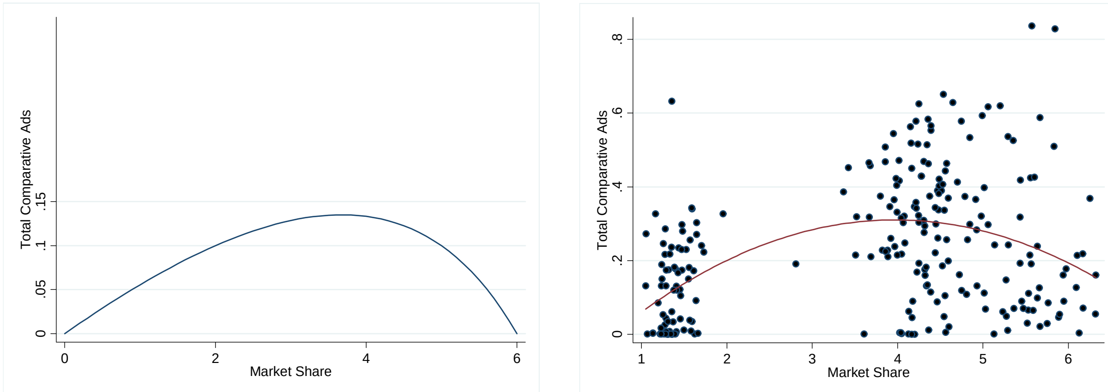


Figure 2. Theoretical and Empirical Total Comparative Advertising and Firm Size Relationship

This relationship is borne out in the data to some extent and is presented in the following diagram: the smaller firms spend less in absolute terms on comparative advertising; the largest one (Tylenol) also does not use it as much as the medium size ones.[35] Note we have fitted a hump relation here: the property above (and indeed the one before draws a similar comment) does not necessarily mean that the largest firm attacks the most, because it does not attack itself.

6 Identification

We estimate the equations (13) and (14) and then (16) and (17). 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 solutions), hence the variables A_{jjt}, A_{jkt}, $j,k = 1,...,n$, are left-censored.[36] 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}, $j,k = 1,...,n$.[37] To clarify the nature of the endogeneity in our analysis, we start from equation (13), which is the simplest equation to deal with. We rewrite it here with the appropriate time subscripts:

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

[35] Direct CA absolute: (from highest to lowest CA expenditures)Advil (2) Aleve (4)Tylenol (1)Motrin (5)Midol (7)Bayer (6)Excedrin (3)Pamprin (8). The outlier is Excedrin (the others fit the right ranking pattern almost perfectly, in the sense we can put a hump through them). Excedrin is though ACT, like Tylenol: there are Excedrin-Tylenol attacks, though small.

CA/TA (from highest to lowest): Midol (7), Aleve (4), Advil (2), Motrin (5), Pamprin (8), Bayer (6), Excedrin (3), Tylenol (1). Recall though that current market shares are dollar market shares and not what we use in estimation (Federico knows those and I cannot compute them unless I have the data that he used in empirical part). Though, they are usually very strongly positively correlated.

[36] As noted above, there are two brands, Pamprin and Midol, which are primarily menstrual formulations, and we exclude them from the empirical analysis. Interestingly, they never engage in non-comparative advertising, only in comparative advertising. Generic brands never engage in any type of advertising.

[37] Notice that prices, which are also endogenous, have been substituted out in the equations (13), (14), (16), and (17).

The term $\bar{A}_{jjt}$ captures the advertising base allure of a brand. We can write this advertising base allure as follows:

$$\bar{A}_{jjt} = Z_{jt}^{'}\Phi + \xi_{jt},$$

where Z_{jt} are observable determinants of the advertising base allure [** which are the news shocks below; anything else?], while ξ_{jt} are unobservable shocks to the base allure. Thus, the unobservable ξ_{jt} is a structural error. Notice that ξ_{jt} is here assumed to be observed by the firms and by the consumers, but not by the econometrician.

Next, recall that the market share for brand j is written as:

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

[** we need to note that the generics are entries here, and NOT in the OG; so we need a notation for strategic firms vs. fringe. Also, we do not need all of this demand structure for what follows; could we, e.g., get a nested version? as illustration. Make sure we say v early that we have little restriction when it comes to just non-comp ads eq.; this means our price foc substitution device can have useful generality, though it does mean we aren't using all the data (see Intro discussion)] where

$$\delta_{jt} = \bar{Q}_{jt}\left(.\right) - p_{jt} + W_{jt}'\Psi + \eta_{jt}, \tag{18}$$

where we have made the substitution

$$\bar{W}_{jt} = W_{jt}'\Psi + \eta_{jt},$$

Here, W_{jt} are observable determinants of the 'true' quality of a brand. [** news shocks enter here, anything else?] η_{jt} is another structural error, which measures any unobserved determinant of the true quality.

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

$$\begin{cases} A_{jjt}^{*} = -Z_{jt}'\Phi + \alpha_D M s_{jt} - \xi_{jt}, & \xi_{jt} \sim N\left(0, \sigma^2\right) \\ A_{jjt} = \max\left(A_{jjt}^{*}, 0\right). \end{cases} \tag{19}$$

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 ξ_{jt} as follows:

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

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 1 a graphical description of the relationship between non-comparative advertising and market shares for all brands and months.

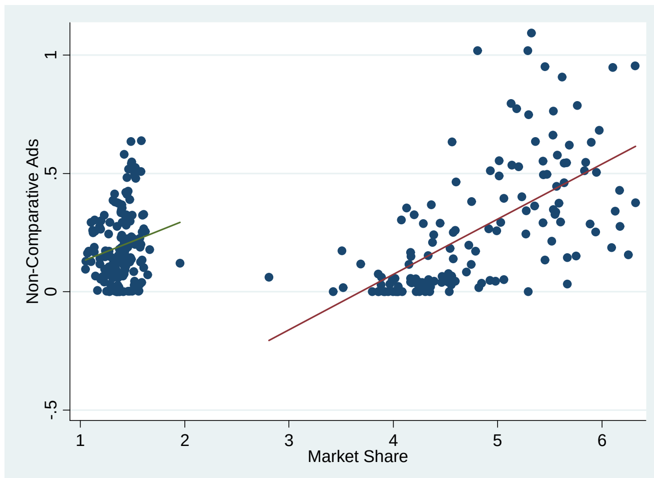


Figure 3. Relationship between Non-Comparative Ads and Market Shares

Figure 3 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[38] 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 $j \in \{Advil,\ Aleve,\ Tylenol\}$ and $\bar{\xi}_j = \bar{\xi}_O$ for $\{Motrin,\ Excedrin,$ $Bayer\}$.[39] 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

[38] More precisely, $Market\ Share = Ms_j/10^7$.

[39] 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].

individual brand-specific fixed effects.[40] 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 (19) can be identified when $\Delta\xi_{jt}$ and s_{jt} are uncorrelated by estimating a variant of (19) where the ξ_{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 (starting from [cite] to the more recent work of []) with a particular assumption on the timing of the realizations of the errors. More specifically, a sufficient condition is the following:

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 $E\left(\Delta\xi_{jt}|s_{jt}, Z_{jt}\right) = 0$.*

One standard interpretation for such a condition is that we are basically able to observe all the variables that the firms take into account when taking their decisions. This means that neither the econometrician nor the firms observe $\Delta\xi_{jt}$ before taking their advertising and pricing decisions.[41] When this assumption is untenable, and identification can be achieved using exclusion restrictions. We now discuss two possible such restrictions.

Exclusion Restrictions. We need variables that affect advertising only through shares, but not directly. One route is to seek variables that enter W_{jt} but do not enter into Z_{jt}. An alternative route is to seek variables that affect shares through prices, p_{jt}, but do not affect perceived quality (such the cost of producing a pill).[42] We consider both types of variables.

[40] Notice, however, that even with individual brand specific fixed effects the 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].

[41] However, the equations that we estimate here were based on assuming that firms *do* observe $\Delta\xi_{jt}$ perfectly at each period: if we did believe the timing assumption, the equations to be estimated would be modified by taking due account of the uncertainty that the firm is facing.

[42] 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.

First, we introduce the following:

Assumption 2 *The news shocks enter into W_{jt} but are not part of Z_{jt}. That is, the news shocks affect the base quality of a product directly, but do not affect its advertising base allure, $\bar{A}_{jjt}$.*

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. Thus, the idea behind this assumption is that the news shocks are associated to the health properties of a product and affect the utility derived by consuming that product (and its demand) directly. However, the news shocks 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. 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.

Another exclusion condition that may be used in place of or alongside that in Assumption 2 is the following:

Assumption 3 *The prices of the generic products are set equal to their marginal costs, which are assumed to be constant. Thus, the prices of the generics enter into each branded product's market share but are excluded from the equation (19).*

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.[43] Second, Bertrand competition among generic producers of the drugs with the same active ingredient leads to pricing at marginal cost.[44] 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.

Control Functions. To implement our estimation in our non-linear models, we use control functions [Heckman, Vella, and co. here]. In practice, the estimation is made in two steps. First, we run the LHS endogenous variables (here market shares, and later also outgoing comparative ads too) on ALL exogenous variables, including those excluded from the second stage relationship. Then, we run the second stage

[43] The marginal cost for pharmaceuticals is reasonably constant, in the sense that there are not increasing returns to scale. CITE (fixed costs?)

[44] 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.

regression (advertising levels here) now including the residuals from the first regression as an additional explanatory variable (the "Control Function") [**right terminology?**] 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.

Our methodology follows [**** needs refs** :: Blundell and Smith and Vuong and ?].[45]

A Look at the other First Order Conditions. In the above discussion we have focused on the first order condition (13). The identification assumptions for the other first order conditions are similar, only the set of instruments changes.

For example, we use the price of the generics using acetaminophen as an excluded variable in the advertising first order condition (13) of Tylenol. We use the prices of the generics using acetaminophen, the prices of the generics using ibuprofen, and their interactions in the first order condition (14) when j is Tylenol and k is Advil. The only difference worthy of mention is that when we run the Tobit regression on the first order condition (16) we need to instrument $\sum_{k \neq j} A_{jk}$, so that the first stage is a Tobit regression as well.

7 Results

For each specification that we run we will first provide some graphical illustration of stylized patterns in the data, and then we will provide the estimation results for the first order conditions (13) and (14), and (16) and (17).

7.1 Fully-Separable Specification

7.1.1 Non-comparative Advertising

The first equation that we estimate is the first order condition (13). Modifying (19) to include fixed effects gives

$$\begin{cases} A^*_{jjt} = -Z'_{jt}\Phi + \alpha_D M s_{jt} - \bar{\xi}_j - \Delta\xi_{jt}, \qquad \Delta\xi_{jt} \sim N\left(0, \sigma^2\right) \\ A_{jjt} = \max\left(A^*_{jjt}, 0\right), \end{cases} \tag{20}$$

where 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.

Table 4 presents the results of four different specifications. All of the specifications are Tobits, where the dependent variable is the total amount spent in non-comparative advertising by the brand j in tens of millions of dollars. The market shares (multiplied by market size and divided by 10^7) are always included

[45] The tobit regressions are all run using *tobit* and *ivtobit* in the Stata package. In specifications where the first stage regression is a tobit itself we compute the standard errors using a boostrap technique.

as an explanatory variable. We include the "outside good" and generics, so that this variable is the total number of pain-hours per month killed by the drug.[46]

(1) FULLY SEPARABLE: NONCOMPARATIVE ADS	**(1)**	**(2)**	**(3)**	**(4)**
IV Tobit	No	No	Yes	Yes
Generic prices as instruments	No	No	No	Yes
Medical news as instruments	No	No	Yes	No
Medical news as controls	No	Yes	No	Yes
MS_j [α_D]	0.260***	0.174***	0.284***	0.281***
	(0.021)	(0.030)	(0.023)	(0.061)
Top Brand [ζ_T]	-0.827***	-0.517***	-0.908***	-0.857***
	(0.073)	(0.106)	(0.080)	(0.201)
Constant [-?$_{jj}$]	-0.201***	-0.077	-0.235***	-0.238**
	(0.033)	(0.057)	(0.035)	(0.099)
Control Function [MS_j]			-0.136**	-0.139**
			(0.053)	(0.070)
First Stage R2 [MS_j]			0.816	0.041
Root Squared MSE	0.198***	0.150***	0.196***	0.149***
Median Elasticity S_j	**3.481**	**2.336**	**3.811**	**3.767**

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

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

Table 4. Fully Separable Specification: Non-Comparative Ads

The first row shows the estimates for the coefficient of Ms_j, α_D. The theoretical model predicts $\alpha_D > 0$: an increase in non-comparative advertising should increase the perceived quality of the product. The second and third rows pick up the impact of the advertising base allure fixed effects. For a Top Brand, the fixed effect is the sum of the two; for non-top brands it given as the value in the third row. Finally, the fourth row reports the estimated coefficient for the control function in cases where we instruments to control for endogeneity of market shares.

Column 1 provides the estimate of α_D when we run the simple Tobit regression (20) of non-comparative advertising, when news shocks are not included in the estimation (so we are setting $\Phi = 0$). Since we are not instrumenting market shares, this will be true if market shares are exogenous so $E\left(\Delta\xi_{jt}|s_{jt}\right) = 0$, as per Assumption 1.

We find the coefficient α_D to be estimated very precisely, both here and in the other specifications. It is positive, which implies that we cannot reject the advertising model that we constructed in Section (3).

To provide an economic interpretation of the coefficient α_D we compute the elasticity of non-comparative advertising to shares (market size is essentially constant over time and the same across brands):

$$e_{A_{jj},s_j} = \frac{dA_{jj}}{ds_j}\frac{s_j}{A_{jj}}.$$

[46] Basically, we are dividing the left and right hand sides by 10^7.

We find the median elasticity to be equal to 3.481, which means that a 10 percent increase in market share, s_j, implies a 34.81 percent increase in non-comparative advertising. This is clearly a strong relationship.

We also find that the constant is equal to -0.201. The fact that we find this significant negative number coheres with our theory insofar as the advertising base allure ($\bar{A}_{jj}$) must be positive in order for the log advertising specification in our model to be defined for all advertising values. This is coherent with the idea that the linear equation indeed is consistent with the first order condition from our model.

The dummy variable for the Top Brands' advertising base allure advantage is estimated to be -0.827. 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. This though means that the marginal efficiency of advertising, in terms of quality increase is LOWER for these brands, although of course they garner the increase over a larger customer base.

As a measure of fitness, we compute the root squared mean standard error of the regression (RMSE), which is equal to 0.198. We will use this measure of fit to compare the results across columns.

Column 2 adds on the Z vector in the form of news shocks. Thus, we estimate their effects on the amount spent on non-comparative advertising by getting estimates for Φ. We are still relying on Assumption 1 for identification ($E\left(\Delta\xi_{jt}|s_{jt}, Z_{jt}\right) = 0.$)

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. Many of them turn out to be statistically and economically insignificant, thus we drop them out. 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 3 to get a sense of which shocks had an effect on which firm.

We estimate α_D to be equal to 0.174. The elasticity is now down to 2.336 from 3.481, a 32 percent drop in its value. The dummy variable *TopBrand* is now equal to -0.517. Thus, the top brands still have an advertising base allure. Instead, because the constant is now almost equal to zero and statistically insignificant, the other brands do not have any advertising base allure. This is quite surprising, and suggests that further investigation in the role of news shocks is needed. This low value might be explained by endogeneity of market shares, which we control for in the next two treatments.

The measure of fit, the RMSE, is now equal to 0.150, suggesting that exogenous variation in the news shocks improves the fit by approximately 20 percent.

Column 3 presents the results when rely on the second identification assumption, but not the third one: the news shocks affect the quality function directly, but do not affect the advertising base allure. Thus, news shocks enter into W_{jt} of equation (18) but not in Z_{jt} of the equation (20).

Here, the coefficient α_D is equal to 0.284. The elasticity is now equal to 3.811, not significantly different from the value of 3.481 that we report in Column 1. Notice, moreover, that the fits of the specifications in Column 1 and 3 are not much different, suggesting that using shocks as instruments does not provide a better fit. These results suggest that news shocks should enter into the advertising first order conditions directly, through their effect on the advertising base allure of a firm. This leads us to the last specification for this table.

Finally, we find that the coefficient of the control function is equal to -0.136 and is statistically significant. This provides evidence that the s_{jt} is endogenous. However, from a purely empirical standpoint, the endogeneity of s_{jt} only leads to negligible bias in its estimated coefficient.

Column 4 includes shocks as controls and uses generic prices as instrumental variables. Essentially we use the third and last identification assumption. This is the main specification for the first order condition (20).

We estimate α_D to be 0.281. The corresponding median elasticity is 3.77. Both the coefficient and the median elasticity are substantially larger than those we estimate in **Column 2**. Notice that we again find evidence that s_{jt} is an endogenous variable in the regression (the coefficient of the control function is -0.139 and statistically significant), but the bias is negligible. This specification achieves the best fit among the ones we report in **Table 4**.

We summarize our empirical analysis of the first order condition (13) as follows. First, we find evidence that non-comparative advertising raises the perceived quality of a brand, since α_D is estimated to be positive. Thus, we cannot reject the theoretical model developed in Section (3) and the corresponding Proposition 1. 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. Finally, as we expected, we find evidence of a clear endogeneity of market shares in the advertising first order conditions, which creates a substantial downward bias on the coefficient of market shares.

7.1.2 Comparative Advertising

The second relation that we test is the first order condition (14). 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:

$$\begin{cases} A^*_{jkt} = -Z'_{jkt}\Gamma - \bar{\xi}_{jk} + Ms_j\frac{\alpha_I}{\gamma} + M\frac{s_j s_k}{(1-s_j)}\frac{\beta}{\gamma} - \Delta\xi_{jkt}, \qquad \Delta\xi_{jkt} \sim N\left(0,\sigma^2\right) \\ A_{jkt} = \max\left(A^*_{jkt}, 0\right). \end{cases} \tag{21}$$

where $\bar{\xi}_{jk} = \bar{\xi}_{TB,TB}$ if j and k are both Top Brands, $\bar{\xi}_{jk} = \bar{\xi}_{TB,OB}$ if j is a Top Brand (i.e., Advil, Aleve, Tylenol) and k is an Other Brand, and likewise for $\bar{\xi}_{OB,TB}$ and $\bar{\xi}_{OB,OB}$. For example, $\bar{\zeta}_{TB,TB}$ is the pairwise group-fixed effect (to be estimated) if both the 'attacker', j, and the 'attacked', k, are top brands. As in Section (7.1.1), we will allow news shocks to enter directly as control variables in the determination of (here, comparative) advertising and indirectly as instrumental variables. The variables that are constructed from the news shocks and that enter into Z_{jkt} are constructed as follows. We interact the ten news shocks with pair specific brand dummies (there are thirty of them). We then drop the shocks that did not have an economically or statistically significant effect (the parameter estimates are still too many to report.)

Finally, before we start discussing the results, recall that we can only identify the ratios $\frac{\alpha_I}{\gamma}$ and $\frac{\beta}{\gamma}$. Recall that α_I measures the effect that outgoing comparative advertising has on the perceived quality of a brand. The parameter β measures the effect that an incoming comparative ad has on the perceived quality of a brand. The parameter γ measures the relative cost of comparative advertising versus non-comparative advertising.

The economic reason why we cannot identify separately α_I, γ, and β is the following. We only have data on advertising expenditures, but we do not have data on the costs of advertising. Thus, we cannot say whether the firm changes the ads because of higher costs or lower effectiveness of the ads. Only if we could use price focs and, possibly, demand functions, we could identify these three parameters separately from each other. One unsuccessful way to identify these parameters that we considered was the following. We know that advertisers must meet the Federal Trade Commission (FTC) standard of truthful and not misleading advertising claims.[47] 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.[48] 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.[49]

[47] 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 serious consequence is the publicized disruption of the ad campaign, sunk costs invested in the ad campaign and negative press related to the brand name.

[48] The National Advertising Division (NAD) is a not-for-profit institution aimed at providing inexpensive, quick, and private process to review ad campaigns. If a firm believes that a competitor's ad campaign is making misleading claims, then it can file a complaint with NAD (which only gets involved when a competitor files a complaint). NAD reviews only national advertisements. The advertising may be placed on broadcast or cable television, in radio, magazines and newspapers, on the Internet or commercial on-line services, or provided direct to the home or office. Product performance claims, superiority claims against competitive products and all kinds of scientific and technical claims in national advertising are the types of cases accepted by the NAD.

Firms prefer to file the complaint with NAD, as that can save the large amounts of money typically spent seeking reparation through the courts. The Federal Trade Commission and courts have a much slower process, and the disputed ads can continue to run during litigation (and possibly continue to steal market share from the challenger). Probably, the most important FTC case is Bayer v. FTC (2000). The case revolves around Bayer's claims that daily aspirin use is an appropriate measure to prevent heart disease, a claim that is not scientifically substantiated. Bayer settled the case by agreeing to a $1 million community education program, and also a $60,000 fine. This case was significant because of the harsh penalty the FTC pushed for due to misleading advertising, and might have caused companies to increasingly rely on and respect the decision of NAD as a means of self-regulation without the dangers of serious fines from the FTC. See *www.nadreview.org* for more information.

[49] This problem is not different from the one that it is encountered when we estimate market power and we do not have

Column 1 of **Table 5** presents the results when we run the above regression (21), and we do not include the news shocks (e.g. $\Gamma = 0$).

We estimate the coefficient of s_j, $\frac{\alpha_I}{\gamma}$, equal to -0.054. This is a small number, quite close to zero. However, it is precisely estimated, and we cannot reject the hypothesis that it is different from zero. This implies that outgoing comparative advertising has a *negative* effect on the perceived quality of the firm that is airing the comparative ad. This result is quite surprising. We would think that $\frac{\alpha_I}{\gamma}$ should be positive, though possibly small, as we expect comparative and non-comparative advertising to be substitutes. We will delve further on this in other specifications of the model. Clearly this finding would lead to a rejection of the theoretical model developed in Section (3).

The coefficient of $\frac{s_j s_k}{(1-s_j)}$, $\frac{\beta}{\gamma}$, is 23 and precisely estimated. This implies that comparative ads have a strong negative effect on the perceived quality of the brand that is being attacked by the ad. This result is fully expected and provides further support to the theoretical model developed in Section (3). Thus, as of now, we have contradictory results from our estimation. On the one hand, the results in **Table 4**,

which we discussed in Section (7.1.1) and those for $\frac{\beta}{\gamma}$ suggest that the theory model is consistent with the observed data. On the other hand, the result for $\frac{\alpha_I}{\gamma}$ suggest that the results are instead rejecting the theory.

We can then interpret the results using the elasticity of comparative advertising with respect to s_{jt} and to s_{kt}. This tells us whether comparative advertising is driven more by the market share of the attacker or by the market share of the attacked. We estimate $e_{A_{jk},s_j} = \frac{dA_{jk}}{ds_j}\frac{s_j}{A_{jk}}$ to be equal to 1.117 and e_{A_{jk},s_k} to be equal to 4.438. 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 44.38 percent.

The results for the group-type fixed effects are also interesting. To understand them, we start with the constant term, which picks up the pairwise group fixed effect when neither j nor k are Top Brands. The constant is equal to -0.153. An interpretation of this as advertising base allure is less intuitive here, since the two brands are of the same type (not a Top Brand).

We rather propose to interpret this result as evidence that any comparative ad is always positively affecting the perceived quality of the attacker, regardless of the amount spent. We estimate Top Brand-Top Brand pair specific dummy variable, to be equal to 0.007, which means that $\bar{\zeta}_{TB,TB} = \bar{\zeta}_{OB,OB}$. The Top Brand-Other Brand dummy is equal to 0.181, thus, $\bar{\zeta}_{TB,OB}$ is essentially equal to zero, since -0.153 and 0.181 delete each other. This has an intuitive interpretation: the comparative ad is not always positively affecting the perceived quality of the Top Brand when the attack is against a brand that is not a Top Brand. Its effect depends on the amount spent in the ad. Finally, the Other Brand-Top Brand is equal to -0.008,

information on the marginal cost. Adding more equations (the first order condition for price and the demand equation) would let us identify α_I, γ, and β.

(2) FULLY SEPARABLE: COMPARATIVE ADS	(1)	(2)	(3)
IV Tobit	No	No	Yes
Generic prices as instruments	No	No	Yes
Medical news as instruments	No	No	No
Medical news as controls	No	Yes	Yes
$MS_j * S_{k/}(1-S_j)$ [β/γ]	23.028***	15.700***	39.835***
	(2.443)	(2.331)	(4.565)
S_j [α_I/γ]	-0.054***	-0.016	-0.079***
	(0.010)	(0.013)	(0.020)
Top Brand - Top Brand [ζ_{TT}]	0.007	-0.058	-0.303***
	(0.052)	(0.049)	(0.084)
Top Brand - Other Brand [ζ_{TO}]	0.181***	0.038	0.120*
	(0.045)	(0.049)	(0.069)
Other Brand - Top Brand [ζ_{OT}]	-0.008	-0.049	-0.171***
	(0.036)	(0.030)	(0.036)
Constant [$-?_{jk}$]	-0.153***	-0.112***	-0.081**
	(0.036)	(0.033)	(0.040)
Control Function [1]			-34.071***
			(5.079)
Control Function [2]			0.107***
			(0.026)
First Stage R2 [1]			0.191
First Stage R2 [2]			0.125
Root Squared MSE	0.152***	0.116***	0.114***
Median elasticity S_j[1]	**1.117**	**1.625**	**2.263**
Median elasticity S_k	**4.438**	**3.026**	**7.677**

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

note: *** p<0.01, ** p<0.05, * p<0.1; [1]range includes negative elasticity

Figure 1: Table 5. Fully Separable Specification: Comparative Ads

which implies that $\bar{\zeta}_{OB,TB} = \bar{\zeta}_{OB,OB}$.

Column 2 adds news shocks in the Tobit regression (21) under the assumption that $E\left(\Delta\xi_{jt}|s_{jt}, Z_{jt}\right) = 0$. We estimate $\frac{\beta}{\gamma}$ equal to 15.700 and $\frac{\alpha_I}{\gamma}$ to be approximately equal to zero (-0.016). Thus, adding news shocks directly in the first order condition has a strong effect on the estimated coefficients. The fit of the regression is much better, as the RMSE drops by approximately 25 percent, from 0.152 to 0.116. As it was the case for **Table 4**, we can conclude that news shocks affect the advertising base allure directly.

The result for $\frac{\alpha_I}{\gamma}$ is interesting because we now cannot reject the theoretical model. For $\frac{\alpha_I}{\gamma}$ to be equal to zero, α_I has to be equal to zero since we should not expect γ is not many order of magnitudes larger than α_I. This implies that comparative ads do not increase the perceived quality of the attacker. While this is still surprising, it is not as counter-intuitive as finding that α_I is negative.

Column 3 uses generic prices as instruments and includes medical news as control variables. We estimate $\frac{\beta}{\gamma}$ to be equal to 39.835 and $\frac{\alpha_I}{\gamma}$ to be -0.079. While $\frac{\beta}{\gamma}$ is three times as large in 3 than in Column 2, we do not find much difference for $\frac{\alpha_I}{\gamma}$. It is still very small and negative.

The large difference between the estimated values of $\frac{\beta}{\gamma}$ and $\frac{\alpha_I}{\gamma}$ in Column 2 and Column 3 suggest that both market shares s_{jt} and the interaction terms $\frac{s_{jt}s_{kt}}{1-s_{jt}}$ are endogenous variables, and their coefficients are underestimated when we do not account for that. The coefficients of the control functions are also statistically significant, confirming that the variables are not exogenous in the regression.

As one would expect, the elasticities are also quite different in Column 3 from those that we estimated in Column 1. We now find e_{A_{jk},s_j} to be equal to 2.263 and e_{A_{jk},s_k} to be equal to 7.677. These results are in the same spirit as those in Column 1, but the magnitudes are essentially doubled. Thus, we find evidence of a very strong positive relationship between comparative ads and the shares of both the attacker and the attacked.

Our conclusions for the analysis of the first order condition (14) are not as clear-cut as were the conclusions for the analysis of (13). On one hand, it is clear that there is a strong positive relationship between comparative advertising and the market shares of the attacker and the attacked. On the other hand, our specifications do not clarify whether $\frac{\alpha_I}{\gamma}$ should be set equal to zero (as suggested by the results in Column 2) or should be allowed to be in the regression and to take negative values. To further investigate the relationship between comparative ads and market shares, we next consider the case where the utility takes a different functional form, as formalized by equation (15).

7.2 Semi-Separable Specification

In this section we consider the first order conditions that we derive when we assume a semi-separable specification (15) for the perceived quality function. The discussion of the results follow the same lines as in the previous Section (7.1).

7.2.1 Non-Comparative Advertising

To begin with, we estimate the following statistical specification for the first order condition for the non-comparative advertising (16):

$$\begin{cases} A^*_{jjt} = -Z'_{jt}\Phi + \alpha_D M s_{jt} - \lambda \sum_{k\neq j} A_{jk} - \bar{\xi}_j - \Delta\xi_{jt}, \quad \Delta\xi_{jt} \sim N\left(0,\sigma^2\right), \\ A_{jjt} = \max\left(A^*_{jjt}, 0\right). \end{cases}$$

where 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. Notice that in this specification both λ and α_D are identified. Recall, that λ can be interpreted as a substitutability parameter of comparative ad vs. non-comparative ads. In other words, λ measures how much one should spend on non-comparative advertising to replace $1 spent on comparative advertising to generate the same change in the "push" part of perceived quality function (raising of your own perceived quality). 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. Note, that this parameter does not represent the "true" effectiveness of comparative advertising relative to non-comparative advertising, as there is the second, "pull", component to comparative advertising, which is directly denigrating the perceived quality of targeted competitors' brands.

The identification issues are analogous to those that we discussed when estimating the first order condition (13) so we omit them here.

Column 1 provides the results of simple Tobit regression on market shares and the non-comparative advertising advertising. The parameter α_D is estimated to be 0.236. This confirms that non-comparative advertising pushes the firm up, or increases its perceived quality and demand.

The substitutability parameter, λ, is estimated to be 0.487. This suggests that comparative ads have a positive effect on the perceived quality of the attacking firm. This effect is quite sizeable, since every dollar spent on comparative ads has the value of half a dollar of a non-comparative ad. This large economic effect and the statistical precision of the estimate is the first piece of evidence that the non-separable model considered in the previous Section (7.1) is mis-specified.

The median elasticity is estimated to be 3.158, which is a number in the same range as the elasticity we estimated in Column 1 of **Table 4**. The interpretation of the other parameters (the constant and the Top Brand) is analogous to that provided in Section (7.1) and we omit it here for sake of brevity.

(3) SEMI SEPARABLE: NONCOMPARATIVE ADS	(1)	(2)	(3)
IV Tobit	No	No	Yes
Generic prices as instruments	No	No	Yes
Medical news as instruments	No	No	No
Medical news as controls	No	Yes	Yes
MS_j [$1 / \alpha_D$]	0.236***	0.163***	0.245***
	(0.020)	(0.030)	(0.063)
Σ Comparative Ads [α_I / α_D]	-0.487***	-0.476***	-0.756***
	(0.067)	(0.060)	(0.210)
Top Brand [ζ_T]	-0.644***	-0.366***	-0.511***
	(0.072)	(0.108)	(0.181)
Constant [$-?_{jj} / \alpha_D$]	-0.140***	-0.023	-0.137
	(0.032)	(0.055)	(0.107)
Control Function [1]			-0.113
			(0.070)
Control Function [2]			0.261
			(0.214)
First Stage R2 [1]			0.072
First Stage R2 [2]			0.033
Root Squared MSE	0.184***	0.135***	0.139***
Median Elasticity S_j	**3.158**	**2.182**	**3.279**

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

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

Table 6. Semi-Separable Specification:: Non-Comparative Ads

Column 2. We add medical news shocks as control variables under the assumption that $E\left(\Delta\xi_{jt}|s_{jt}, Z_{jt}\right) = 0$. The results change a bit, but generally confirm the findings in Column 1.

Column 3. provides results when we use generic prices as instrumental variables and we include medical news shocks as control variables. Interestingly, we do not find evidence that market shares and the comparative ads are endogenous variables, since the coefficients of the Control Functions are not statistically significant. On the other hand, the coefficient estimates in Columns 1, 2 and 3 are quite different suggesting that instrumenting with the generic prices is necessary to estimate the coefficients consistently. We interpret these findings as the result of the fact that generic prices can explain only 7 percent of the variation in market shares and 3 percent in the variation in comparative ads in the first stage regressions.

The crucial result in Column 3 is surely the once concerning λ, which is estimated equal to -0.756. Recall that this means that each dollar spent on comparative ad increases the perceived quality of the attacking brand by the same amount as 75 cents spent on non-comparative ad. This is a very large number, and suggest that there is very large degree of substitutability between comparative and non-comparative ads, at least as far as their effect on the perceived quality of the attacker is concerned.

These findings suggest that the amount of money spent on comparative ads should enter into the first order for non-comparative ads. Further evidence, even though not particularly strong, is given by our measures of fit, which are better in **Table 5** than in **Table 4**.

7.2.2 Comparative Advertising

Lastly, we estimate the following statistical specification for the first order condition of semi separable specification for comparative ads (17):

$$\begin{cases} A^*_{jkt} = -Z'_{jkt}\Gamma - \bar{\xi}_{jk} + M\frac{s_j s_k}{(1-s_j)}\frac{\beta}{\gamma-\lambda} - \Delta\xi_{jkt}, \qquad \Delta\xi_{jkt} \sim N\left(0,\sigma^2\right) \\ A_{jkt} = \max\left(A^*_{jkt}, 0\right). \end{cases}$$

where $\bar{\xi}_{jk} = \bar{\xi}_{TB,TB}$ if j and k are both Top Brands, $\bar{\xi}_{jk} = \bar{\xi}_{TB,OB}$ if j is a Top Brand (i.e., Advil, Aleve, Tylenol) and k is an Other Brand, and likewise for $\bar{\xi}_{OB,TB}$ and $\bar{\xi}_{OB,OB}$. For example, $\bar{\zeta}_{TB,TB}$ is the pairwise group-fixed effect (to be estimated) if both the 'attacker', j, and the 'attacked', k, are top brands. As in Section (7.1.1), we will allow news shocks to enter directly as control variables in the determination of (here, comparative) advertising and indirectly as instrumental variables. The variables that are constructed from the news shocks and that enter into Z_{jkt} are constructed as follows. We interact the ten news shocks with pair specific brand dummies (there are thirty of them). We then drop the shocks that did not have an economically or statistically significant effect (the parameter estimates are still too many to report.)

The results are presented in **Table 7**. Because of its similarity with the analysis presented in **Table 5**, we will focus here on the differences. Notice that from a purely statistical point of view, the results in **Tables 5** and **7**, fit the data in comparable fashion.

Column 1 presents results from the simple Tobit. We estimate $\frac{\beta}{\gamma-\lambda}$ equal to 16.228. When used to compute the elasticities, we find that the elasticities are pretty much the same for the market shares of the attacker and of the attacked brand. The median elasticity with respect to s_{jt} is equal to 3.131 and equal to 3.127 with respect to s_{kt}.[50]

Columns 2 and 3 show the results when we add the news shocks and when we also use the generic prices as instrumental variables. There is clear and strong evidence that the interaction term $\frac{s_j s_k}{(1-s_j)}$ is endogenous. First, the coefficients of the Control Function in Column 3 is large and statistically very significant. Second, $\frac{\beta}{\gamma-\lambda}$ is way underestimated in Column 1 relative to Column 3. This results to elasticities which are twice the size in Column 3 (around *six*) than in Column 1 (around *three*). These elasticities suggest a strong relation for Propositions 2 and 3.

[50] These elasticities are almost the same, although they were quite different with the fully separable version of the comparative ad relation. The latter has an extra term in it, which is affected only by own share. Moreover, the common term begets almost the same elasticity for the median share because that share is very small (and the elasticities differ by a factor of 1-share).

(4) SEMI SEPARABLE: COMPARATIVE ADS	(1)	(2)	(3)
IV Tobit	No	No	Yes
Generic prices as instruments	No	No	Yes
Medical news as instruments	No	No	No
Medical news as controls	No	Yes	Yes
$MS_j * S_{k/}(1-S_j)$ [$\beta / (\gamma - \alpha_I / \alpha_D)$]	16.228***	14.723***	35.020***
	(2.106)	(2.004)	(4.250)
Top Brand - Top Brand [ζ_{TT}]	-0.029	-0.088*	-0.449***
	(0.054)	(0.046)	(0.080)
Top Brand - Other Brand [ζ_{TO}]	0.038	-0.003	-0.100***
	(0.037)	(0.032)	(0.036)
Other Brand - Top Brand [ζ_{OT}]	0.032	-0.041	-0.136***
	(0.036)	(0.030)	(0.034)
Constant [$-?_{jk}$]	-0.230***	-0.138***	-0.198***
	(0.035)	(0.028)	(0.030)
Control Function			-27.206***
			(4.980)
First Stage R2			0.177
Root Squared MSE	0.157***	0.116***	0.114***
Median elasticity S_j*	**3.131**	**2.841**	**6.757**
Median elasticity S_k	**3.127**	**2.837**	**6.749**

Number of observations: 1160; Left-Censored Observations: 663
note: *** p<0.01, ** p<0.05, * p<0.1

Table 7. Semi-Separable Specification: Comparative Ads

8 Conclusions

We allow comparative advertising to have two effects on consumer choice probabilities, and we empirically estimate the strength of these two effects. The "push" component of comparative advertising improves the direct perceived quality of the product being "pushed", and non-comparative advertising is solely push. The "pull" component of comparative advertising pulls down the perceived quality of the firm targeted by the comparative advertising.

References

[1] Ackerberg, D. (2001): Empirically Distinguishing Informative and Prestige Effects of Advertising. *RAND Journal of Economics* 32 , 100–18.

[2] Ackerberg, D. (2003): Advertising, learning, and consumer choice in experience good markets: an empirical examination. *International Economic Review,* 44, 1007-.

[3] Albion M. and P. W. Farris (1981): *The Advertising Controversy: Evidence on the Economic Effects of Advertising*. Boston, MA, Auburn House Publishing Company.

[4] Aluf, Yana, and Oz Shy (2001): Comparison advertising and competition. Mimeo, University of Haifa.

[5] Anand, B. and R. Shachar (2001), "Advertising, the Matchmaker," *Harvard Business School Working Paper* No. 02-057.

[6] Anderson, Simon P. , Federico Ciliberto and Jura Liaukonyte (2009): Getting into Your Head(ache): Advertising Content for OTC Analgesics. Mimeo, University of Virginia.

[7] Anderson, Simon P. , Federico Ciliberto, Jura Liaukonyte and Régis Renault (2008): Sheening Characteristics in Advertising. *Mimeo*, University of Virginia.

[8] Anderson, Simon P. and André de Palma (1988): Spatial Price Discrimination with Heterogeneous Products. *Review of Economic Studies*, 55, 573 592.

[9] Anderson, Simon P. and André de Palma (2001): Product Diversity in Asymmetric Oligopoly: Is the Quality of Consumer Goods Too Low? *Journal of Industrial Economics*, 49, 113-135.

[10] Anderson, Simon P. and André de Palma (1992): Multiproduct Firms: A Nested Logit Approach," *Journal of Industrial Economics*, 40, 261 276.

[11] Anderson, Simon P. and André de Palma, and Jacques-Francois Thisse (1989): Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach. *Review of Economic Studies* 56: 21-35.

[12] Anderson, S, A. de Palma and J.-F. Thisse (1992): *Discrete choice of product differentiation*. Cambridge, MA: The MIT Press.

[13] Anderson, Simon P. and Renault, Régis (1999): Pricing, Product Diversity and Search Costs: a Bertrand-Chamberlin-Diamond Model. *RAND Journal of Economics*, 30, 719-735.

[14] Anderson, Simon P. and Renault, Régis (2000): Consumer Information and Firm Pricing: Negative Externalities from Improved Information. *International Economic Review*, 31, 721-741.

[15] Anderson, Simon P. and Renault, Régis (2006): Advertising content. *American Economic Review*, 96, 93-113.

[16] Anderson, Simon P. and Renault, Regis (2009): Comparative Advertising: disclosing horizontal match information. Forthcoming, *RAND Journal of Industrial Economics*.

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

[18] Baker, B. and T. F. Bresnahan (1985): The Gains from Merger or Collusion in Product-Differentiated Industries. *The Journal of Industrial Economics, .*

[19] Barigozzi, Francesca, Paolo Garella, and Martin Peitz (2006): With a little help from my enemy: comparative vs. generic advertising. Mimeo, University of Bologna.

[20] Barigozzi, Francesca and Martin Peitz (2006): Comparative advertising and competition policy. Chapter 8, 215-263, in *Recent Developments in Antitrust: Theory and Evidence*, edited by Jay Pil Choi, MIT Press.

[21] Becker, Gary S., and Kevin M. Murphy (1993): A Simple Theory of Advertising as a Good or Bad. *Quarterly Journal of Economics* 108, 942–64.

[22] Belch, George E. (1981): An Examination of Comparative and Noncomparative Television Commercials: The Effects of Claim Variation and Repetition on Cognitive Response and Message Acceptance. *Journal of Marketing Research*, 18 (August), 333-49.

[23] Ben-Akiva, Moshe (1973): Structure of passenger travel demand models. *Transportation Research Board Record No. 526,* Washington D.C.

[24] Ben-Akiva, Moshe and Steven R. Lerman (1985): *Discrete Choice Analysis: Theory and Applications to Travel Demand*. Cambridge, MA: MIT Press.

[25] Berry, S. (1994), "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics* 25(2): 242-262.

[26] Berry, S, J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica* 63(4): 841-890.

[27] Besanko, D., S. Gupta and D. Jain (1998), "Logit Demand Estimation under Competitive Pricing Behavior: An Equilibrium Framework," *Management Science* 44 (Part 1 of 2), 1533–1547.

[28] Bresnahan, T. F. (1989), "Empirical Studies of Industries with Market Power," In: R. Schmalensee and R. D. Willig (eds.), *Handbook of Industrial Organization*, Amsterdam: Elsevier Science Publisher B. V., 1011-1057.

[29] Bresnahan, Timothy F., Scott Stern, and Manuel Trajtenberg (1997), "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *The RAND Journal of Economics,* Vol. 28, No. 0, S17-S44.

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

[31] Crawford, G. and M. Shum (2000), "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, vol. 73, issue 4, pages 1137-1173.

[32] Chintagunta, P., (2001), "Endogeneity and Heterogeneity in a Probit Demand Model: Estimation Using Aggregate Data," *Marketing Science*.

[33] Chintagunta, P., Jiang, R. and Jin, G. (2007), "Information, Learning, and Drug Diffusion: the Case of Cox-2 Inhibitors" Working paper, University of Chicago GSB Marketing.

[34] Dixit, A. K. and Norman, V. D. (1978), "Advertising and Welfare," *Bell Journal of Economics*. 9, 1-17.

[35] Donthu, Naveen (1998): A cross-country investigation of and attitude toward comparative advertising. *Journal of Advertising,* 27, 111-133.

[36] Dorfman R. and P. O. Steiner (1954), "Optimal Advertising and Optimal Quality," *American Economic Review*, 44(5): 826-836.

[37] Droge, Cornelia and Rene Y. Darmon (1987), "Associative Positioning Strategies Through Comparative Advertising: Attribute Versus Overall Similarity Approaches," *Journal of Marketing Research,* 24 (November), 377-88.

[38] Dube, Jean-Pierre, Günter Hitsch, and Puneet Manchanda (2005): An Empirical Model of Advertising Dynamics. *Quantitative Marketing and Economics*, 3(2), 2005, 107-144.

[39] Dube, Jean-Pierre, Puneet Manchanda Kim Yong Goh and Pradeep Chintagunta (2006): The Effects of Banner Advertising on Internet Purchasing. *Journal of Marketing Research*, 43(1), 98-108.

[40] Erdem, T. and M. Keane (1996), "Decision-Making under Uncertainty: Capturing Dynamic Brand Choices in Turbulent Consumer Goods Markets," *Marketing Science,* 15, 1–20.

[41] Erdem, Tülin, Michael P. Keane, and Baohong Sun (2008): The impact of advertising on consumer price sensitivity in experience goods markets. *Quantitative Marketing and Economics*, 6(2), 139-176.

[42] Hosp, D. R. (2007): Weighing The Risks: A Practical Guide To Comparative Advertising. excerpted chapter in the practice manual *Advertising and Marketing Litigation Best Practices*. Aspatore Books.

[43] Gasmi, F., J. J. Laffont, and Q. Vuong (1992) "Econometric Analysis of Collusive Behavior in a Soft-Drink Market", *Journal of Economics & Management Strategy,* 1, 277–312.

[44] Goeree, M. S. (2008), "Limited Information and Advertising in the US Personal Computer Industry," *Forthcoming Econometrica*.

[45] Golden, Linda L. (1979), "Consumer Reactions to Explicit Brand Comparisons in Advertisements," *Journal of Marketing Research*, 16 (November), 517-32.

[46] Gonul, Fusun and Srinivasan, Kannan (1993), "Consumer Purchase Behavior in a Frequently Bought Product Category: Estimation Issues and Managerial Insights from a Hazard Function Model with Heterogeneity" *Journal of the American Statistical Association,* Vol. 88, No. 424, pp. 1219-1227

[47] Goodwin, Steven and Michael Etgar (1980), "An Experimental Investigation of Comparative Advertising: Impact of Message Appeal, Information Load, and Utility of Product Class," *Journal of Marketing Research*, 17 (May), 187-202.

[48] Grewal, Dhruv, Sukumar Kavanoor, Edward F. Fern, Carolyn Costley, and James Barnes (1997): Comparative versus noncomparative advertising: a meta-analysis. *Journal of Marketing*, 61, 1-15.

[49] Grossman, Gene M. and Shapiro, Carl (1984): Informative Advertising and Differentiated Products. *Review of Economic Studies*, 51, 63-81.

[50] Grossman, Sanford J. (1981): The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24(3), 461-83.

[51] Imbens, Guido W. and Lancaster, Tony (1994) "Combining Micro and Macro Data in Microeconometric Models," *The Review of Economic Studies,* Vol. 61, No. 4 , pp. 655-680

[52] Jain, D. C., N. J. Vilcassim, and P. K. Chintagunta (1994). "A Random-Coefficients Logit Brand-Choice Model Applied to Panel Data." *Journal of Business and Economic Statistics* 12(3):317–328.

[53] Jin, Ginger, Pradeep Chintagunta, and Keanna Jiang (2009): Information, Learning, and Drug Diffusion: the Case of Cox-2 inhibitors. Mimeo, University of Maryland.

[54] Kaldor, N., (1950), "The Economic Aspects of Advertising," *Review of Economic Studies*, 18, 1-27.

[55] Kihlstrom, R. E. and Riordan, M. H. (1984), "Advertising as a signal," *Journal of Political Economy.* 92, 427-450.

[56] Kirmani, Amna (1990), "The Effect of Perceived Advertising Costs on Brand Perceptions," *The Journal of Consumer Research,* 17(September), 160-171.

[57] Kirmani, Amna (1997), "Advertising Repetition as a Signal of Quality: If It's Advertised So Often, Something Must Be Wrong," *Journal of Advertising*, XXVI (3), 77-86.

[58] Kirmani, Amna and Peter Wright (1989), "Money Talks: Perceived Advertising Expense and Expected Product Quality," *The Journal of Consumer Research*, 16 (December), 344-353.

[59] Kirmani, Amna and Valarie Zeithaml (1993), "Advertising, Perceived Quality and Brand Image," in David A. Aaker and Alexander Biel, eds., *Brand Equity and Advertising,* Erlbaum: NJ, 143 162.

[60] Lancaster, K.(1971), "Consumer Demand: a new approach," New York: *Columbia University Press.*

[61] Levine, Philip (1976), "Commercials That Name Competing Brands," *Journal of Advertising Research*, 16 (December), 7-14.

[62] Liaukonyte, Jura (2008): Is comparative advertising an active ingredient in the market for pain relief? Mimeo, University of Virginia.

[63] Lilien, G.L., Ph. Kotler, K.S. and Moorthy, K.S., (1992), "Marketing Models", Prentice-Hall, Englewood Cliffs, NJ.

[64] Mediamark Internet Reporter. Mediamark Research, Inc. (Spring 2001-Fall 2006). Non-Prescription Home Remedies. Retrieved January 6, 2008 from MRI Plus database)

[65] McFadden D. (1974), "Conditional Logit Analysis of Quantitative Choice Behavior," In P. Zarembka, ed., *Frontiers in Econometrics.* New York: Academic Press.

[66] McFadden, D., (1978), "Modeling the choice of residential location," In Transportation Research Record 672. TRB, National Research Council, Washington, DC, pp. 72–77

[67] McFadden, D. (1981), "Modelling the choice of residential location," *Spatial Interaction Theory and Residential Location,* pages 75–96. North-Holland (Amsterdam).

[68] McFadden, Daniel (1981), "Econometric Models of Probabilistic Choice," *Structural Analysis of Discrete Data with Econometric Applications*, ed. by Charles Manski and Daniel McFadden. Cambridge, MIT Press.

[69] McFadden, D. (1995), "Computing Willingness-to-Pay in Random Utility Models," Mimeograph, Department of Economics, University of California, Berkeley.

[70] McFadden, Daniel and Train, Kenneth (2000). "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 15(5), pages 447-470.

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

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

[73] Milgrom, P. R., and Roberts, J. M. (1986), "Prices and Advertising Signals of Product Quality," *Journal of Political Economy*. 94, 796-821.

[74] Miniard, Paul W., Randall L. Rose, Michael J. Barone, and Kenneth C. Manning (1993), "On the Need for Relative Measures When Assessing Comparative Advertising Effects," *Journal of Advertising*, 22 (September), 41-57.

[75] Moorthy, S. and Zhao, H., (2000) "Advertising spending and perceived quality," Marketing Letters, pp. 221–233

[76] Nelson, Phillip J. (1970), "Information and Consumer Behavior," *Journal of Political Economy*. 78, 311-329.

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

[78] Nevo A. (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2): 307-342.

[79] Pechmann C. and D. Stewart, 1990, "The Effects of Comparative Advertising on Attention, Memory, and Purchase Intentions," *Journal of Consumer Research*, 17, 180-191.

[80] Pechmann, C. and D. W. Stewart (1991), "How Direct Comparative Ads Promoting Low, Moderate and High Share Brands Affect Brand Choice," *Journal of Advertising Research*, 31, 47-55.

[81] Pechmann, Cornelia and S. Ratneshwar (1991), "The Use of Comparative Advertising for Brand Positioning: Association Versus Differentiation," *Journal of Consumer Research*, 18, 145-60.

[82] Petrin, A. (2002), "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110, 705–729.

[83] Roberts M.J. and Samuelson, L. (1988), "An Empirical Analysis of Dynamic, Nonprice Competition in an Oligopolistic Industry," *Rand Journal of Econometrics*, 19: 200-220.

[84] Rojas, C. and Peterson, E. B. "Demand for differentiated products: Price and advertising evidence from the U.S. beer market," *International Journal of Industrial Organization*, 26, 288-307.

[85] Rose R.L., Miniard, P.W., Barone, M.J., Manning, K.C. and Till, B.D (1993), "When Persuasion Goes Undetected: The Case of Comparative Advertising," *Journal of Marketing Research*, 30 315-330

[86] Schmalensee R. (1972), "The Economics of Advertising, Contributions to Economic Analysis," Amsterdam, North-Holland Publishing Co.

[87] Stigler, G.J. and Becker, G.S. (1997), "*De Gustibus Non Est Disputandum,"* American Economic Review.

[88] Shin, Hyun (1994): The Burden of Proof in a Game of Persuasion. *Journal of Economic Theory*, 64, 253-263.

[89] Shum, M. (2004), "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market," *Journal of Economics and Management Strategy* 13:2, 241-272.

[90] Shy, O., (1995), "Industrial Organization: Theory and Applications," Chapter 11. Cambridge and London MIT Press.

[91] Sutton, J. (1992), "Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration," Cambridge, MA: *The MIT Press.*

[92] Swinyard, William R. (1981), "The Interaction Between Comparative Advertising and Copy Claim Variation," *Journal of Marketing Research,* 18 (May), 175-86.

[93] Telser, L.G. (1964), "Advertising and Competition," *Journal of Political Economy* 72 , 537–62.

[94] Tirole, J. (1988), "The Theory of Industrial Organization," Cambridge, MA: *MIT Press.*

[95] Thompson, Debora Viana and Rebecca W. Hamilton (2006): The effects of information processing mode on consumers' responses to comparative advertising. *Journal of Consumer Research,* 32, 530-540.

[96] Tremblay, V.J. and Polasky, S. (2002), "Advertising with Subjective Horizontal and Vertical Product Differentiation," *Review of Industrial Organization,* 20, 253-265

[97] Villas-Boas, J.M., and Winer, R.S. (1999), "Endogeneity in Brand Choice Models," *Management Science,* Vol. 45, No. 10.

[98] Wolinsky, Asher (1986): True Monopolistic Competition as a Result of Imperfect Information. *Quarterly Journal of Economics*, 101, 493-511.

[99] *International Journal of Clinical Practice*, March 2008, 62, 3, 444–449

[100] Aspirin Could Reduce Breast Cancer Risk 4/8/2003 *Journal of the American Medical Association*, 2004;291:2433-2440.

9 Appendix - Theory

9.1 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):

$$\underset{\{p_j^a p_j^b, A_j\}}{Max} \pi_j = M(p_j^a - c_j)s_j^a + M(p_j^b - c_j)s_j^b - A_{jj}$$

The advertising first order condition is

$$\begin{aligned}
\frac{d\pi_j}{dA_{jj}} &= \frac{d\pi_j}{d\delta_j^a} \cdot \frac{\partial Q_j^a}{\partial A_{jj}} + \frac{d\pi_j}{d\delta_j^b} \cdot \frac{\partial Q_j^b}{\partial A_{jj}} \\
&= M(p_j^a - c_j)\left[\frac{\partial s_j^a}{\partial \delta_j^a}\frac{\partial Q_j^a}{\partial A_{jj}} + \frac{\partial s_j^a}{\partial \delta_j^b}\frac{\partial Q_j^b}{\partial A_{jj}}\right] + M(p_j^b - c_j)\left[\frac{\partial s_j^b}{\partial \delta_j^b}\frac{\partial Q_j^b}{\partial A_{jj}} + \frac{\partial s_j^b}{\partial \delta_j^a}\frac{\partial Q_j^a}{\partial A_{jj}}\right] - 1 = 0
\end{aligned}$$

Using the price first order conditions:

$-M(p_j^a - c_j)\frac{\partial s_j^a}{\partial \delta_j^a} - M(p_j^b - c_j)\frac{\partial s_j^b}{\partial \delta_j^a} + Ms_j^a = 0$

and the analogous condition for p_j^b yields

$$\frac{d\pi_j}{dA_{jj}} = Ms_j^a \frac{\partial Q_j^a}{\partial A_{jj}} + Ms_j^b \frac{\partial Q_j^b}{\partial A_{jj}} - 1 = 0.$$

Under the (strong) assumption that $\frac{\partial Q_j^a}{\partial A_{jj}} = \frac{\partial Q_j^b}{\partial A_{jj}} = Q'(A_{jj})$ we have

$$Ms_j Q'(A_{jj}) = 1,$$

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

9.2 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:

$$\underset{\{p_j, A_j\}}{Max}\, \pi_j = (p_j - c_j)\left(N^A s_j^A + N^B s_j^B\right) - A_{jj} - \gamma \sum_{k \neq 1}^{n} A_{jk} \qquad j = 1, ...n, \tag{22}$$

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:

$$\left(N^A s_j^A + N^B s_j^B\right) \frac{\partial Q_j}{\partial A_{jj}} = 1, \qquad j = 1, ..., n. \tag{23}$$

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.

9.3 Total ad spending

As we saw above, brand j's total comparative ad spending is $\frac{M}{\gamma} \frac{s_j}{(1-s_j)} (1 - s_j - s_0) - \sum_{k \neq j} \bar{A}_{jk}$. We now add its positive ad spending, which is (from (**??**)) $A_{jj} = M s_j - \bar{A}_{jj}$. This yields:

$$\begin{aligned} TA_j &= \sum\nolimits_k A_{jk} \\ &= \frac{M s_j}{\gamma (1 - s_j)} \left[(1 + \gamma)(1 - s_j) - s_0\right] - const. \end{aligned} \tag{24}$$

Notice that total ads are maximized (as a function of size) at a value of For example, suppose that $\gamma = 1$ nonetheless. Then we have

$\frac{s_j}{(1-s_j)} (2 - s_0 - 2s_j)$. This is plotted for various values in the next Figure. Black: $s_0 = 0.1$.

Red:$s_0 = 0.05$; green then with $\gamma = 2$.

Loosely, a dip now only occurs with low s_0 and low γ. The powerful effect of the positive ads wins out. Is this worth noting, though not in a full (sub-)section?

What we might want to stress is that we could not get a dip if all ads were push-only. That is, if we were to use total ads and ignore the fact that some of it is comparative. Is there then a dip in the data? Let's find more differences in terms of what we'd get wrong if we made a wrong assumption that all is positive. And let us do the econometric analysis, then show how much better we perform when we do it properly.