Information content of advertising: empirical evidence from the OTC analgesic industry

Anderson, Simon and Ciliberto, Federico and Liaukonyte, Jura

University of Virginia, University of Virginia, Cornell University

28 April 2012

Information Content of Advertising: Empirical Evidence from the OTC Analgesic Industry[1]

Simon P. Anderson[2]
University of Virginia
Center for Economic Policy Research

Federico Ciliberto[3]
University of Virginia
Center for Economic Policy Research

Jura Liaukonyte[4]
Cornell University

December 2012

Abstract

We empirically study the *information-persuasion trade-off* in advertising using data on the *information content* of advertisements, which we measure with the number of information cues in ads. We propose a simple theoretical framework to motivate an ordered probit model of information content. We find that stronger vertical differentiation is positively associated with the delivery of more product information in a brand's advertisements: brands with higher levels of quality include more information cues. Next, comparative advertisements contain significantly more product information than self-promotional advertisements. Finally, brands with higher market shares and brands competing against strong generic substitutes have less information content.

Keywords: Information content, Advertising, Information-persuasion tradeoff, Content analysis, Ordered probit.

JEL Codes: L13, M37, L65.

[1] 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"). We thank the Melbourne Business School and the Portuguese Competition Authority for their hospitality. All authors thank the Bankard Fund for Political Economy at the University of Virginia.
[2] Department of Economics, University of Virginia, sa9w@virginia.edu, ph: (434) 924-3861.
[3] Department of Economics, University of Virginia, ciliberto@virginia.edu, ph: (434) 924-6755.
[4] Charles H. Dyson School of Applied Economics and Management, Cornell University, Jurate@cornell.edu, ph: (607) 255-6328.

1. Introduction

We develop an empirical study of how much information brands choose to disclose in advertisements. Recent theoretical contributions have analyzed this choice (Anderson and Renault, 2009; Guo and Zhao, 2009; Sun, 2011). In parallel, recent empirical contributions have coded advertising content (Bertrand et al., 2010; Liaukonyte, 2012; Anderson et al., 2012 (ACLR); and Ching et al., 2011) to better understand the impact of advertising. We contribute to this growing empirical literature by explaining the number of information cues in advertisements through underlying product characteristics and other market fundamentals, and we theoretically found the use of the ordered probit model in content analysis.

We maintain that there exists an optimal amount of information to be included in an advertisement. Advertisements that provide too little objective information about the brand arguably waste the opportunity to sufficiently convince prospective consumers to buy it (Jacoby, 1977). Conversely, those that provide too much information may crowd the ad message and lead to information overload for the consumer (Chervany and Dickson, 1974; Pieters, Wedel and Zhang, 2007). Factors such as motivation and the ability to process information mediate individual responses to advertising (Cacioppo and Petty, 1985; MacInnis and Jaworski, 1989). Therefore, including too many information cues can create attention wear-out (Pieters, Warlop and Wedel, 2002), which suggests that there is an optimal degree of information content.

The underlying notion behind our empirical analysis is that there is a trade-off between the informative and persuasive roles of advertising, subject to random factors intrinsic to specific ads. In our approach, the "persuasive" content of an ad is interpreted as the content that is not objective information. We treat the persuasive component of an advertisement as *unobserved*. The theoretical framework that we develop assumes that the amount of information in the ad is directly proportional

to the number of information cues mentioned in the ad (and inversely related to the persuasive component of an ad).

The optimal amount of information content may vary systematically across brands and may be partially explained by observable factors such as brand type, brand size, suitability of various combinations of information, and recent news about the product. To study this variation, we use an ordered probit model of the number of information cues in advertisements as a function of the market share of a brand, its core vertical characteristics, and the market share of the generic substitute. To motivate the use of an ordered probit, we develop a simple and intuitive framework in which firms decide how much objective information to include in an advertisement.

As predicted by our theoretical model, we find that stronger vertical differentiation, measured by the characteristic advantage of a product relative to its competitors, is positively associated with more information. Likewise, a higher market share of the generic version of a brand is associated with less informative ads. We also find that comparative advertisements contain significantly more information than self-promotional advertisements, and that larger brands provide less information. Finally, we show that significant estimation bias results from not controlling for the endogeneity of the decision to use comparative advertising and from the endogeneity of market shares.

Our analysis contributes to the sparse empirical literature on the persuasive vs. informative content of advertising, which mostly makes use of advertising expenditure/exposure data to analyze *a priori* assumed informative or persuasive roles of advertising (see Bagwell, 2007, for a comprehensive survey). Several papers use insightful identification strategies to separately quantify the informative and persuasive components of advertising (Leffler, 1981; Ackerberg, 2001, 2003; Ching and Ishihara, 2012). Ackerberg (2001, 2003) maintains that advertising is only informative for first-time buyers, while it is both informative and persuasive for repeat buyers. Ching and

Ishihara (2012) assume that the informative component of advertising is specific to the exogenous characteristics of a product and the persuasive component is brand specific. Basuroy, Rao, and Ravid (2011) identify the persuasive and informative effect of advertising utilizing the assumption that if the persuasive role is important, then one would expect the number and quality of reviews included in a movie promotion to play a major role in the success of the advertisement.

Although the identification strategies of these existing works are clever and rely on the interaction between choice behavior and advertising, they still suffer from the limitations that they only use data on advertising expenditure or number of advertisement exposures instead of the actual content of advertisements. Our approach is fundamentally different because we explicitly code the information content. This enables us to distinguish between the informative and the persuasive role of advertising directly from the advertising content data.

Several recent contributions exploit advertising content data but study different questions. Liaukonyte (2012) codes the content of advertisements to distinguish the impact of comparative and self-promotional advertising on demand. Ching et al. (2011) study that the impact of the content of media coverage on anti-cholesterol drugs on consumer demand varies by publicity type. Finally, Bertrand et al. (2010) develop a field experiment in South Africa and show that advertising content significantly affects demand for loans. In contrast, we focus on firm behavior rather than on the effect of advertising on demand. Another paper closely related to ours is ACLR (2012), which uses the same data source to investigate how firms strategically use comparative advertising and measures the magnitude of the damage inflicted with comparative advertising. Our paper differs from ACLR (2012) because we focus on the *amount* of information content of advertising and we explain it with an ordered probit model.

2. Advertising Content and the Information-Persuasion Trade-off Model

Information in advertisements can be (1) about the brand, brand attributes, benefits, users, or usage situation; (2) cognitive, emotional or subconscious; or (3) contextual information, including consumers' past experiences (MacInnis, Moorman, and Jaworski, 2001; Vakratsas and Ambler, 1999). We focus on the first aspect of brand information (traceable, objective information cues) and treat cognitive, emotional, and contextual information as the untraceable, subjective element of an advertisement. We define this subjective unobserved element as the *persuasive component*.[5]

Contrary to the classic content analysis, which determines the fraction of advertisements that fall into each category given a number of cues and uses univariate analysis to compare scenarios (see e.g., Resnik and Stern, 1977), we use an ordered probit model to study the *determinants* of the distribution of cues. We now present a simple way to motivate the use of an ordered probit in this context.

We maintain that firms face a trade-off: given the limited amount of time available in an advertisement, the firm must decide how much of that time to allocate to providing information and how much of it to allocate to persuading consumers to buy the brand through the use of channels other than objective information. We assume that the persuasive power of an advertisement depends on the number of seconds devoted to the persuasive component, and a random term, which reflects the idiosyncratic features of a particular advertisement. Let the persuasive power of the ad be

$$P(s, \varepsilon) = (\bar{p} - \varepsilon)s$$

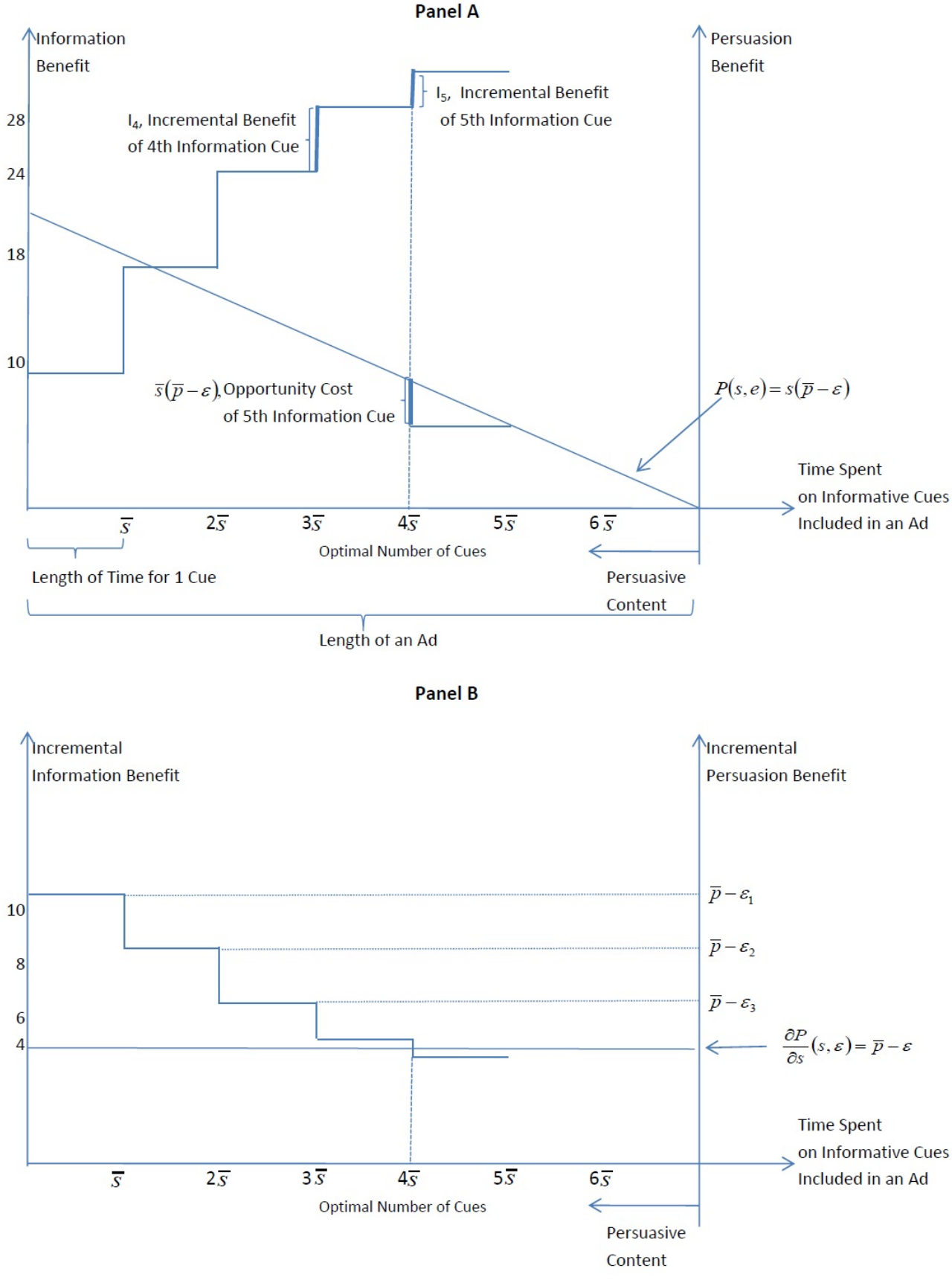
[5] The persuasive component can also be traceable in well-defined studies (e.g., Maheswaran and Meyers-Levy, 1990), but it is outside of the scope of this paper.

where s is the number of seconds spent on persuasion, $\bar{p}$ is a constant and ε is a random term.[6] Panel A of Figure 1 illustrates the function *P(s,ε)*. The x-axis indicates the number of seconds used for persuasion. As more time is allocated to persuasion, the function *P(s,ε)* increases from right to left. The y-axis shows the total benefit from persuasion. The more time spent on persuasion, the larger the total benefit.

The marginal persuasion for an advertisement with s seconds of persuasion and a given draw of the random term ε is therefore $\bar{p} - \varepsilon$. The persuasion function is linear, and therefore the marginal persuasion does not change with the share of the advertisement that is allocated to it. Panel B of Figure 1 illustrates the function $\bar{p} - \varepsilon$. The x-axis again indicates the number of seconds used for persuasion. Because $\bar{p} - \varepsilon$ does not change with s, the function $\bar{p} - \varepsilon$ is parallel to the x-axis. The y-axis depicts the marginal benefit of persuasion. In the empirical implementation of the econometric model that we are developing here, $\bar{p}$ will vary according to the observable features of the advertised brand (e.g., market share, observable quality, generic competition, etc.) and of the advertisement itself, such as whether it is a comparative or self-promotional advertisement.

[6] *P(s,ε)* does not have to be linear in s. It should be an increasing and concave function of s. The linear specification simplifies the exposition significantly.

Figure 1. Trade-off Between Informative and Persuasive Content



Let each information cue take $\bar{s}$ seconds to convey, so that if there are S seconds in the ad (i.e., n information cues are conveyed), there are $s=S-n\bar{s}$ seconds of persuasion. Let I_i be the benefit of the i^{th} information cue i, with $i=1,...,n$. We assume that firms rank the cues from the highest to the lowest information benefit for each given advertisement. We do not observe such information cue ranking, however we assume that the inclusion of the observed cues is based on a firm's intrinsic ranking. Note that our theory is consistent with rankings varying across different types of advertisements for the same brand, since advertisements can be tailored to appeal to target demographics who value product characteristics differently. Clearly, the brand will choose to include the cues delivering the highest information benefit, i.e., those cues for which the values of I_i are the highest.

Because a given advertisement only has a limited amount of available time, S, the brand must decide how much of that time to allocate to providing information vs. for persuasion. This trade-off and the total benefit of information are depicted in Panel A of Figure 1. Panel B of Figure 1 shows the marginal benefit of information, which is decreasing in the amount of information already provided. The marginal benefit of information is also a step function. The firm chooses s (or, alternatively, n) to maximize the sum of the total benefit of persuasion and the total benefit of information. Formally, the firm solves:

$$\begin{cases} \max_{s} D\left(\sum_{i=1}^{n} I_i + P(s,\varepsilon)\right) \\ \quad s.t. \;\; n = \dfrac{S-s}{\bar{s}}. \end{cases}$$

Here $D(.)$ is an increasing function representing the firm's demand as a function of an advertisement's information and persuasion content. The solution to this optimization problem can be described by comparing the incremental benefit from adding a cue to the advertisement to the opportunity cost of reducing the time spent on persuasion. If the advertisement contains $n-1$ cues, then the extra benefit from the n^{th} cue is I_n. We can see this graphically in Panel A of Figure 1.

There we see that I_4, the marginal benefit of the fourth information cue, is larger than I_5, the marginal benefit of the fifth information cue. The slope of the persuasion function, $\bar{p} - \varepsilon$, is such that:

$$I_5 < \bar{s}(\bar{p} - \varepsilon) < I_4.$$

This implies that the optimal number of cues depicted in Panel A of Figure 1 is four.

Panel B of Figure 1 shows this solution in a way that allows us to introduce our ordered probit in a straightforward way. Define for all I_i the value ε_i such that

$$I_i = (\bar{p} - \varepsilon_i)\, \bar{s},$$

so that ε_i is the threshold value of the random error such that the firm chooses to include at least i information cues in the advertisement if $\varepsilon > \varepsilon_i$. In Figure 1, Panel A depicts $\varepsilon_4 < \varepsilon < \varepsilon_5$. Hence, in this particular illustration, the brand chooses to include four information cues in the advertisement.

Accounting for the fact that there can be no fewer than zero cues and no more than $C = S/\bar{s}$, we have the following mapping from the value of ε to the number of cues, here denoted by y:

$$\begin{cases} y = 0, \text{ for } \varepsilon \leq \varepsilon_1 \\ \ldots \\ y = i, \text{ for } \varepsilon_i < \varepsilon \leq \varepsilon_{i+1} \\ \ldots \\ y = n, \text{ for } \varepsilon > \varepsilon_n. \end{cases}$$

The basic intuition of this statistical model is the following: when the (negative) random shock is very small ($\varepsilon < \varepsilon_1$) — which implies that an advertisement has the ability to have very strong persuasive power (relative to the benefit of information in that particular advertisement) — then the firm has no incentive to include information cues. As the persuasive power of an advertisement decreases, the firm chooses to include more information cues.

From the specification above we can construct a probability distribution of observing the corresponding number of cues, where $F(\,.\,)$ denotes the cumulative distribution of ε:

$$\begin{cases} y = 0, \text{ with probability } F(\varepsilon_1) \\ \ldots \\ y = i, \text{ with probability } F(\varepsilon_{i+1}) - F(\varepsilon_i) \\ \ldots \\ y = n, \text{ with probability } 1 - F(\varepsilon_n). \end{cases}$$

If *F(.)* is normally distributed, the formulation corresponds to an ordered probit model.[7] In particular, we show how the probability function in the model above can be written as the textbook version of the ordered probit, where $\varepsilon_i = \alpha_i - X\beta$ and therefore $F(\varepsilon_i) = \Phi(\alpha_i - X\beta)$. In this ordered probit specification, the unobserved components, i.e. *ε,* are drawn from a normal distribution, and the cutoff values (α_i) are such that the realization of a latent variable (explained component plus noise) lies within a range that corresponds to each specific number of cues.[8] β is a $K \times 1$ vector of parameters, and X is a $K \times 1$ vector of observable features of the brand, which does not include a constant. For $\beta > 0$, an increase in X will lower the threshold ε_i which will in turn make it more desirable to add additional information cues. Therefore, for $\beta > 0$, larger *Xs* are associated with more information content in an advertisement.

We can rewrite $F(\varepsilon_i) = \Phi(\alpha_i - X\beta)$. First, recall that $\varepsilon_i = \bar{p} - \frac{I_i}{\bar{s}}$. Here $\bar{p}$ is a variable, the value of which determines the benefit of persuasion content. Hence, we set $\alpha_i - X\beta = \bar{p} - \frac{I_i}{\bar{s}}$. It is useful to extract a constant from $-\frac{I_i}{\bar{s}}$ and to rewrite the equality as $\alpha_i - X\beta = \bar{p} - \bar{I} - \frac{\tilde{I}_i}{\bar{s}}$. Then $X\beta = \bar{I} - \bar{p}$ so that, consistent with the discussion in the preceding paragraph, the X variables increase the benefit of information or, equivalently, decrease the benefit of persuasion. Second, we can define the cutoff value $\alpha_i = -\frac{\tilde{I}_i}{\bar{s}}$, which determines whether the firm is including *i* or *i+1*

[7] For a comprehensive review of ordered models, see Greene (1997) and Woolridge (2001).

[8] For concreteness, we describe the random term as entering the persuasion power, but it could just as well enter the marginal information benefit described below. Or, indeed, the random term could enter on both sides and the analysis would then capture the net effect.

information cues. The econometric model does not identify the constant term $\bar{I}$ separately from the cutoffs α_i, which we need to keep in mind when we interpret the cutoffs.

In our framework, the cutoffs have a clear structural interpretation. In particular:

$$\alpha_{i+1} - \alpha_i = \frac{I_i - I_{i+1}}{\bar{s}}.$$

Thus, conditional on $\bar{s}$ (which is unobserved), differences in the cutoffs provide information on differences in the information benefits of an additional cue.

We estimate this ordered probit model and use the results to examine the relationships between the fundamental variables (e.g., the market share of a firm) and the number of information cues that a firm includes in an advertisement. Before discussing the relationships of interest, we state and prove a helpful result.

Lemma 1. An increase in I_i stochastically increases the number of information cues.

Proof. From the analysis above, i cues will be advertised if $\varepsilon \in (\varepsilon_i, \varepsilon_{i+1})$, where, as above, $\varepsilon_i = \bar{p} - \frac{I_i}{\bar{s}}$. The corresponding probability of observing i cues is $P_i = F(\varepsilon_{i+1}) - F(\varepsilon_i)$, or

$P_i = F(p - \frac{I_{i+1}}{s}) - F\left(p - \frac{I_i}{s}\right)$. Suppose now that I_i increases while retaining its position as the ith largest information benefit. Then P_i increases at the expense of P_{i-1}, while all other probabilities remain unchanged. Hence, the number of cues increases stochastically. Now suppose that the increase in I_i raises it to the j^{th} highest cue, with $j < i$. Then each intervening cue is promoted so that the probability of observing at least that number of cues rises. The probability of observing $i+1$ cues or more stays the same, as does the probability of observing each number less than j. Again, the number of cues stochastically increases. QED

3. Testable Model Implications

First, we consider intrinsic characteristics of a product, which include: strength of pain relief, relative efficacy, and safety. In the industry analyzed in this paper, these variables are naturally exogenous to the information decision since they depend on the medical properties of the active ingredients in analgesic pain relievers, which in turn are inflexibly regulated by the U.S. Food and Drug Administration (FDA). Therefore, the direction of causality can be clearly identified, and we can investigate how different locations in the product characteristics space are associated with information disclosure.

The driving idea here is that the information benefit to a brand from communicating a certain characteristic increases with the brand's strength in that characteristic. Communicating a weak characteristic does not give as much incremental benefit as communicating a strong one. Additional strength in any particular cue will raise its relative information benefit and also make it more likely to be included in the advertisement. Thus, using the result from **Lemma 1** we would expect that brands with higher levels of quality (on each of the quality dimensions for which we have data) will include more information cues in their advertisements.

The second relationship we study is that of information disclosure and the size of the generic counterpart market. An important characteristic of our analyzed market and of many other consumer product categories is the presence of generic substitutes. To the extent that branded product advertisements disseminate information about the benefits of the chemical molecule, generic counterparts are likely to benefit from these promotional efforts. Our second testable implication extends this information externality argument that has been made in the literature that studies the competition between brand-name drugs and generic drugs (e.g., Caves et al., 1991).

The simple framework presented above does not provide any prediction about the presence of generic substitutes on information content. However, we can impose some restrictions to our theoretical framework that would lead to tighter predictions about this relationship. Consider the following simple specification for the objective function of a non-generic firm:

$$D\left(\sum_{i=1}^{n} I_i + P(s,\varepsilon)\right) - L\left(s_g, \sum_{i=1}^{n} I_i\right),$$

where $D(\sum_{i=1}^{n} I_i + P(s,\varepsilon))$ denotes the increase in demand for both the branded product and its generic counterpart, and $L\left(s_g, \sum_{i=1}^{n} I_i\right)$ denotes the leakage to the generic. The leakage is assumed to be increasing in the generic market share, s_g, and in the information content, $\sum_{i=1}^{n} I_i$. Note that *n* information cues will be preferred to *n+1* if :

$$D(\sum_{i=1}^{n} I_i + P(s,\varepsilon)) - L\left(s_g, \sum_{i=1}^{n} I_i\right) > D(\sum_{i=1}^{n+1} I_i + P(s,\varepsilon)) - L\left(s_g, \sum_{i=1}^{n+1} I_i\right)$$

where we have let the argument *n* denote the use of the first *n* information cues. Rewriting this condition as $L\left(s_g, \sum_{i=1}^{n+1} I_i\right) - L\left(s_g, \sum_{i=1}^{n} I_i\right) > D(\sum_{i=1}^{n+1} I_i + P(s,\varepsilon)) - D(\sum_{i=1}^{n} I_i + P(s,\varepsilon))$ and noting that the right hand side of the inequality does not depend upon s_g, then including the lower number of cues, *n*, is more likely to be preferred as the left hand side of the equation increases. That holds if the incremental leakage is increasing in s_g. This is a natural condition given that the leakage itself is increasing in s_g (for example, if leakage were proportional to generic market size times an increasing and concave function of information cues). This is consistent with the rationale that brands that emphasize product quality also provide free advertising for generic counterparts of their products[9], and therefore this would decrease the incentive to include information in the ads. On the other hand, consumers of brands might not be inclined to consume generics, and, in that case, brands might include information in their ads, independently of the size of the generic counterpart.

[9] Unlike other generic substitutes of consumer products, the quality of the active ingredients in the generic versions of OTC analgesics are (and, by regulation, have to be) identical to the quality of the active ingredients in the branded products, since the quality of these OTC drugs is overseen by the FDA..

While this is an intuitive and interesting exercise, our empirical analysis will not impose any (sign) restriction on the relationship between the share of the generic counterpart and the amount of information included in an ad.

Although the two implications tested in this paper may not be entirely surprising or novel from the theoretical viewpoint, testing them with empirical data is challenging since it is generally difficult to measure the product quality and information content of advertising objectively.

4. Other Economic Relationships of Interest

Our theoretical framework does not provide any formal insights into the relationship between a brand's market share and information content or incidence of comparative advertising and information content. One possible interpretation is for market shares and comparative advertising to be variables that shift the constant term $\bar{I}$ in the theoretical model. More generally, these relationships hold empirical interest and we explore them by including them as control variables in our econometric specification.

Larger brands might benefit less than smaller brands from providing information, since larger brands are already well known and have higher advertising goodwill and brand equity (Simon and Sullivan, 1993; Dekimpe and Hanssens, 1995). In other words, the incremental benefit for larger brands might be smaller because consumers are already more aware of the features of commonly used products.

There is no existing theory that tells us whether comparative advertisements should include more or less information than self-promotional advertisements.[10] We expect that more information would be conveyed by comparing two brands than by just promoting a single brand, and comparing

[10] Each comparative advertisement must include at least one cue, because the comparison is made for at least one characteristic. We compare all self-promotional advertisements to all comparative advertisements, conditional on an advertisement having at least one cue.

relative performance provides a more precise and concrete reference point. Similarly, an advertisement with comparative content is likely to have a weaker persuasion effect, even if the amount of time devoted to persuasion is the same. This might happen because mentioning the other brand dilutes the persuasion since it reiterates the existence of the rival brand. Previous research (Chou, Franke, and Wilcox, 1987; Harmon, Razzouk, and Stern, 1983) has found that comparative advertisements have more information, and we expect to find similar patterns. In particular, we expect a comparative claim to both increase the marginal information and decrease the marginal persuasion benefit of an advertisement. Both effects would cause the number of cues to rise stochastically (see Figure 1).

5. Data and Content Analysis

5.1. Industry Description and Data

We follow Liaukonyte (2012) and use sales and advertising data from the OTC analgesics industry in the United States. The OTC analgesics market covers pain relief medications with four major active chemical ingredients: aspirin, acetaminophen, ibuprofen, and naproxen sodium. The nationally advertised brands include Tylenol (acetaminophen), Advil (ibuprofen), Motrin (ibuprofen), Aleve (naproxen sodium), Bayer (aspirin or combination), and Excedrin (acetaminophen or combination).

We focus on the OTC analgesics industry for several reasons. First, television advertising constitutes a large fraction of total advertising in the industry, implying that it is the most important marketing strategy used by the industry to communicate to its consumers. Second, the products sold by firms differ significantly in their characteristics, so that there is a range of meaningful information to potentially communicate (Lancaster, 1971; Christou and Vettas, 2008). Third, the

information is concentrated in experience- and credence-based characteristics, which tend to be advertised, rather than search characteristics that consumers can learn in the store before purchase (Erdem, Keane and Sun, 2008). Fourth, product differentiation—both real and spurious—is important because we emphasize the trade-off between persuasion and information, with the optimal mix depending upon a product's characteristics (Soberman, 2002). Finally, the type of cues mentioned (e.g., "strong," "fast-acting") are clearly identifiable, which enables us to avoid making any subjective judgments while coding the information cues.

The brand and generic sales data come from AC Nielsen and consist of monthly observations of the prices, dollar sales, and revenue market shares of OTC analgesics in the national U.S. market for 58 months, from March 2001 to December 2005. We observe data on essential product attributes such as brand name and active ingredient of a drug, as well as strength and pill type. This product sales and characteristics data is supplemented with the efficacy data on the strength of pain relief, relative speed, and safety for each brand. We collected this information from peer-reviewed medical journals. Clinically, all four main active ingredients have varying degrees of side effects. Each active ingredient has a comparative advantage in at least one attribute dimension. For example, Aspirin (brand name – Bayer) is weak in pain relief but has low, almost nonexistent cardiovascular risk; Naproxen sodium (Aleve) is the most potent drug but is associated with very high gastrointestinal risk; Acetaminophen (Tylenol and Excedrin) has low gastrointestinal risk but is weak in pain relief and has medium cardiovascular risk and Ibuprofen (Advil and Motrin) and naproxen sodium–based brands (Aleve) have the highest cardiovascular risk but are also the fastest in pain relief.

The advertising data come from TNS Media Intelligence and cover the entire U.S. OTC analgesics product category. The dataset contains video files of all unique advertisement creatives and monthly expenditures associated with airing them nationally from 2001 to 2005. The

advertising numbers also include expenditures on other media, but almost all the advertising budgets (approximately 90%) were spent on television advertising, including network and cable networks. In our analysis, we examine only the television advertising data. We watched 4503 individual commercials that were broadcast during the 2001–2005 period, 346 of which had missing video files. Each individual advertisement was usually shown multiple times.

We organize these three data sources (product characteristics and sales data, advertising expenditures and content data, and medical efficacy data) into one final dataset. The observation in this final data set is a unique advertisement creative – month, and variables that are associated with each such observation are the following: advertising expenditures on a specific advertisement creative in a specific month, advertising content for that advertisement creative (coded), brand sales during that month, generic counterpart sales during that month and product quality (based on medical data) associated with the active ingredient of advertised brand.

5. 2. Measuring Information Content

The widely used Resnik and Stern (1977) method for measuring advertising information categorizes the information provided in advertisements into 14 distinct "information cues," including price, quality, performance, components, availability, special offers, taste, nutrition, packaging, warranties, safety, independent research, company research, and new ideas. More than 60 studies have applied the Resnik and Stern approach to measure the information content of advertising in different media (Chou, Franke, and Wilcox, 1987; Harmon, Razzouk, and Stern, 1983; Stern and Resnik, 1991), countries (Hong, Muderrisoglu, and Zinkhan, 1987; Madden, Caballero, and Matsukubo, 1986), and product categories (Stern, Krugman, and Resnik 1981). The results have varied markedly, even within the same medium, because of the lack of a multivariate statistical analysis, redundant or too broad definitions of information cues, and small sample size (Abernethy

and Franke, 1996). The main advantage of the Resnik and Stern classification system is the general nature of the information cue categories, which allow for comparison of products from multiple industries. However, this advantage is also a disadvantage. Categorizing advertising information content into coarse categories inevitably omits some information that consumers might find important. For example, in the OTC analgesics industry, two distinct information cues (e.g., fast and strong) would be coded as one "performance" cue in the Resnik and Stern classification system.

Our attribute coding approach documents every attribute mentioned. For each advertisement, we recorded whether the commercial had any comparative claims and, if so, the specific claim (e.g., faster, stronger). We also noted all information cues mentioned, including the purpose of the drug (e.g., arthritis, headache), drug efficiency (e.g., strength, speed), safety, and other characteristics. The type of information cues that were mentioned (e.g., "strong") are clearly identifiable, which enables us to avoid making any subjective judgments while coding the information content. The disadvantage of this approach is that it is industry-specific. However, focusing on a specific industry is now the standard approach in empirical industrial organization (Bresnahan, 1989).

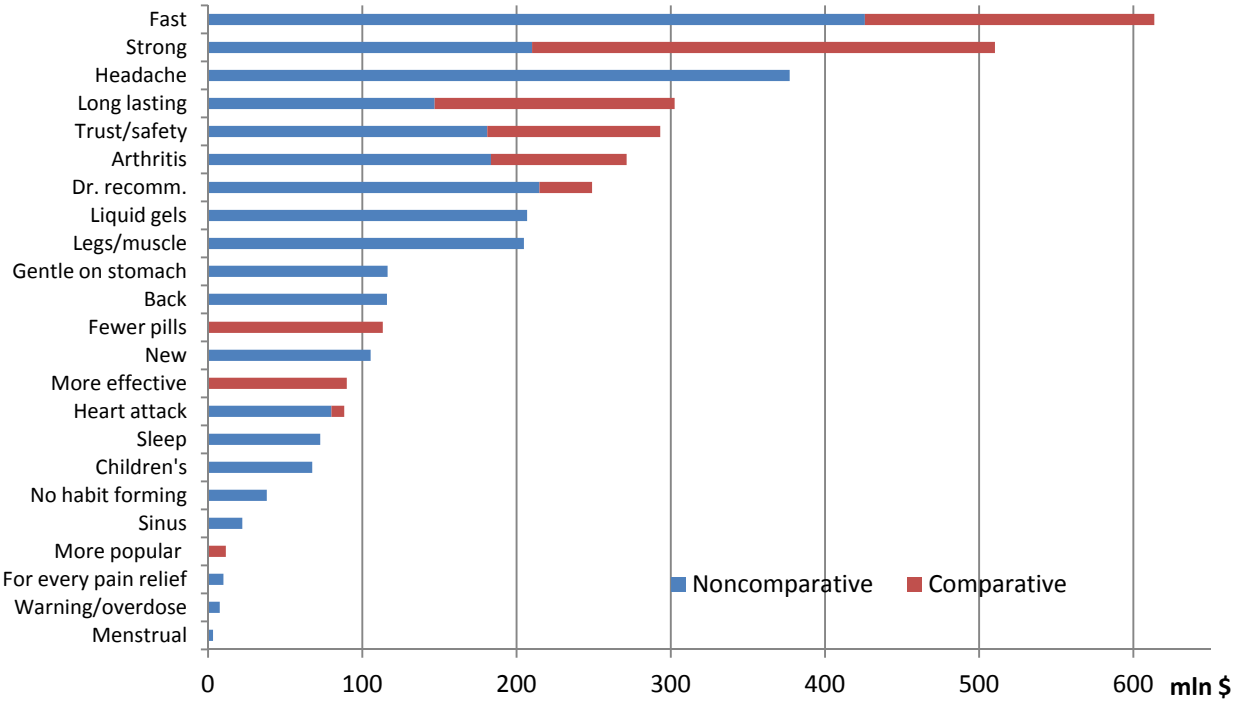
Table 1. Descriptive Statistics of Information Disclosure, Vertical Attributes, Ad Spending and Sales

	Information Disclosure		Vertical Characteristics[1]				Advertising and Sales				
Brand	**Number of Cues**	**Compa-rative?**	**NNT**	**GI Risk**	**CV Risk**	**Relative Speed**	**Avg Monthly Spending per Ad**	**Average Monthly Sales**	**Total Ad Spending**	**Total Sales**	**Ad to Sales Ratio**
Advil	3.6	0.74	2.40	2.00	1.44	1	$0.14	$23.92	$293.10	$1,374	21.30%
	(1.004)	(0.441)					(0.241)	(1.693)			
Aleve	3.77	0.9	2.30	9.10	1.44	3	$0.12	$11.41	$174.80	$659	26.50%
	(1.156)	(0.298)					(0.293)	(1.123)			
Bayer	3.19	0.31	4.00	3.10	1.07	4	$0.10	$7.95	$131.20	$458	28.80%
	(1.320)	(0.461)					(0.222)	(0.964)			
Excedrin	2.4	0.15	3.86	1.67	1.26	2	$0.26	$12.39	$182.40	$689	26.50%
	(0.695)	(0.359)					(0.456)	(1.172)			
Motrin	2.61	0.37	2.40	2.00	1.44	1	$0.10	$8.03	$102.00	$466	21.90%
	(0.937)	(0.484)					(0.240)	(0.762)			
Tylenol	2.54	0.28	3.80	1.00	1.35	2	$0.13	$40.59	$414.90	$2,328	17.80%
	(0.957)	(0.449)					(0.346)	(3.195)			

Notes: [1] Larger number implies weaker performance. These are relative risk and performance measures taken from peer-reviewed medical literature. See Appendix for more details.

We quantify or rank all of the "true" characteristics that were used in advertising associated with each active ingredient as follows. First, we interpret "fast" as the time taken to achieve perceptible or meaningful pain relief (medical literature calls this "onset to perceptible pain relief"). Second, we interpret claims about strength (e.g., "strong," "stronger," "tougher on pain") as the maximum level of pain relief achieved; we use the number-needed-to-treat (NNT) measure to approximate analgesic efficiency claims. Finally, we collect reported data on gastrointestinal and cardiovascular risks (GI Risk and CV Risk, respectively) for each of the active ingredients. Appendix A describes these measures in greater detail and presents a list of peer-reviewed medical articles that quantify these performance attributes. The first two columns of Table 1 also present each brand's (expenditure-weighted) average number of information cues mentioned in ads as well as the incidence of comparative ads. We also add advertising spending and sales descriptive statistics to Table 1.

Figure 2. Advertised Attributes and Expenditures



During our analyzed period, 30 different product attributes were mentioned. Figure 2 represents the top 23 attributes and shows the advertising expenditures (in millions of dollars) spent on advertising those attributes during the sample period.[11] We separate advertising expenditures by type of advertisement (comparative versus self-promotional). The attributes "fast," "strong," "long lasting," and "trust and/or safety" are among the top five most heavily advertised attributes. These attributes are directly related to the inherent (exogenous) chemical characteristics of each active ingredient in each analyzed brand.

We examine the attribute usage correlations to investigate whether the coded cues represent distinct information. For example, although we code "strong" and "fast" as separate information categories, we also ensure that the coded information categories are indeed distinct information cues. Table 2 portrays the correlation matrix of cue usage and shows that the cue descriptors that we use are distinctive. For example, both "fast" and "strong" are often used together, but they are also used separately in more than half of the occurrences (in dollar terms).[12] Thus, two cues may be used together frequently, but each still provides consumers with important information. "Strong" denotes how powerful the medicine is, and "fast" denotes the speed of the onset of pain relief.

Our careful coding and investigation of cue correlations (see Table 2 and Figure 2) suggest that the number of coded attributes within an advertisement is a reasonable approximation of the information content. We use these information cue numbers as our dependent variable in our empirical estimation described below.

---

[11] The remaining 7 attributes (not reported) had negligible advertising expenditures. The sum of the expenditures in Figure 2 exceeds total ad spending because many advertisements promote multiple characteristics and, for the purpose of this Figure only, we attributed total ad spending to each characteristic mentioned.

[12] There are two instances of high correlation that merit comment. First, whenever "liquid gels" are mentioned, "fast" is almost always mentioned. Second, "long lasting" and "fewer pills" are used together 33% of the time. Conversely, when "fewer pills" are mentioned, long lasting was mentioned 88% of the time. We experimented with providing an umbrella classification that encompasses both, but the differences in the results were minor.

Table 2. Matrix of Frequency of Attributes that are Mentioned Together

	Fast	Strong	Head-ache	Long lasting	Safe	Arthritis	Dr. recomm	Liquid gels	Legs/ muscle	Gentle on stomach	Back	Fewer pills	**Total**
Fast		$251.61 49.32%	$296.52 78.63%	$70.25 23.22%	$32.76 11.17%	$23.80 8.77%	$16.69 6.70%	$204.55 98.90%	$101.42 49.52%	$66.32 56.98%	$25.65 22.11%	0 0	**$613.52**
Strong	$251.61 41.01%		$126.43 33.53%	$103.37 34.16%	$93.88 32.02%	$123.06 45.34%	$97.80 39.28%	$140.66 68.01%	$133.31 65.10%	$45.46 39.05%	$51.94 44.78%	$53.01 46.81%	**$510.20**
Headache	$296.52 48.33%	$126.43 24.78%		$6.22 2.06%	$28.85 9.84%	$11.95 4.40%	$25.42 10.21%	$84.88 41.04%	$17.53 8.56%	$24.40 20.96%	$15.88 13.69%	$6.22 5.49%	**$377.09**
Long lasting	$70.25 11.45%	$103.37 20.26%	$6.22 1.65%		$68.50 23.36%	$153.97 56.73%	$83.96 33.72%	$14.43 6.98%	$64.17 31.34%	$23.18 19.91%	$44.99 38.79%	$100.06 88.35%	**$302.56**
Safe	$32.76 5.34%	$93.88 18.40%	$28.85 7.65%	$68.50 22.64%		$115.84 42.68%	$77.64 31.18%	$5.51 2.67%	$29.96 14.63%	$55.50 47.68%	$39.29 33.87%	$30.46 26.90%	**$293.20**
Arthritis	$23.80 3.88%	$123.06 24.12%	$11.95 3.17%	$153.97 50.89%	$115.84 39.51%		$125.59 50.43%	$21.58 10.43%	$18.38 8.97%	$55.21 47.43%	$19.22 16.57%	$80.12 70.75%	**$271.42**
Dr. recomm.	$16.69 2.72%	$97.80 19.17%	$25.42 6.74%	$83.96 27.75%	$77.64 26.48%	$125.59 46.27%		0 0	$23.88 11.66%	$38.35 32.95%	$4.71 4.06%	$50.65 44.73%	**$249.02**
Liquid gels	$204.55 33.34%	$140.66 27.57%	$84.88 22.51%	$14.43 4.77%	$5.51 1.88%	$21.58 7.95%	0 0		$23.06 11.26%	$44.74 38.43%	0 0	0 0	**$206.82**
Legs/ muscle	$101.42 16.53%	$133.31 26.13%	$17.53 4.65%	$64.17 21.21%	$29.96 10.22%	$18.38 6.77%	$23.88 9.59%	$23.06 11.15%		$23.06 19.81%	$56.44 48.66%	$7.54 6.65%	**$204.79**
Gentle on stomach	$66.32 10.81%	$45.46 8.91%	$24.40 6.47%	$23.18 7.66%	$55.50 18.93%	$55.21 20.34%	$38.35 15.40%	$44.74 21.63%	$23.06 11.26%		$0.88 0.76%	0 0	**$116.40**
Back	$25.65 4.18%	$51.94 10.18%	$15.88 4.21%	$44.99 14.87%	$39.29 13.40%	$19.22 7.08%	$4.71 1.89%	0 0	$56.44 27.56%	$0.88 0.76%		$14.83 13.10%	**$115.98**
Fewer pills	0 0	$53.01 10.39%	$6.22 1.65%	$100.06 33.07%	$30.46 10.39%	$80.12 29.52%	$50.65 20.34%	0 0	$7.54 3.68%	0 0	$14.83 12.79%		**$113.25**

6. Identification Strategy

A brand's decision about how much information to include in an advertisement is likely to be made simultaneously with the decision about the type of advertisement (comparative or self-promotional). Therefore, these two decisions are interdependent, in much the same way as equilibrium price and quantity are determined at the same time in a simple demand–supply model. In other words, there is some unobservable exogenous variable that explains both the information content and whether that content is a self-promotional or comparative advertisement. For example, a higher-quality brand

might provide less information and have more comparative advertisements than a lower-quality brand.

The other two potentially endogenous explanatory variables that we consider are the measures of the market share of a brand and the market share of a brand's generic counterpart. To see why endogeneity might be an issue, note that market size and information content are outcomes of brands' strategic interactions, and a fully structural equilibrium model would specify three equations, one for each of the three variables (information content, market size, and the generic counterpart's market size). Here, we estimate only one of the three equations—the one that explains information content as a function of the other two endogenous variables—but we control for the endogeneity of the other two variables. Equivalently, we could consider an unobservable variable (e.g., quality) that is correlated with the brand's market size and information content. By omitting that variable from our regressions, we would introduce a bias into the estimation of the parameters of the model.

We use an instrumental variable approach to determine whether our concerns about the endogenous variables are empirically relevant. Instruments that correlate with sales and advertising but not with unobserved quality provide information on how important the endogeneity problem is likely to be. Our main identification strategy is to use variation in generic prices as the exogenous source of variation in the industry.

The crucial observation for our identification strategy is that all of the patents for the active ingredients of the OTC analgesics have expired. The patent for naproxen was the last one to expire – in 1993. After patents expire, generic counterparts are produced and sold at prices that are substantially lower than the brand name product (Aleve). As time passes, new entry of generic competitors brings the price of the generic counterpart down to marginal cost, as shown by Grabowski and Vernon (1992) and discussed by Ching (2004) and Scott Merton (2004). If the

marginal cost is constant and the generic prices are set at the marginal cost, then the generic prices are independent of the prices set by the national brands and can be appropriately used as instrumental variables.[13] [14]

More specifically, we include functions of a brand's active ingredient generic counterpart prices and generic prices of a brand's competitors (constructed means, squared and cubed terms). We also construct interactions between generic prices (own and competitors') and a dummy indicating that a firm produces multiple brands. Finally, we interact the characteristics with the 2005 year dummy to capture advertising content changes—2005 was considered one of the most turbulent years in the analgesics industry.[15] We discuss in detail our identification strategy and the instrumental variables in the online supplement.

To deal with the endogenous variables in our nonlinear ordered probit model, we follow Rivers and Vuong's (1988) proposed approach. First, we rewrite the information content as $\alpha_i - X\beta - w\gamma = \bar{p} - \bar{I} - \frac{\tilde{I}_i}{\bar{s}}$, where w is a vector of the three endogenous variables. The main identification assumption is that instruments are not correlated with the error term (i.e., $(Z'\epsilon) = 0$). Here, Z includes all of the exogenous variables, such as brand characteristics (X) and functions (here, the average) of the characteristics of the brand's competitors. We use a two-step procedure described in Rivers and Vuong (1988). First, we run an ordinary least squares regression, $w = Z\delta + \upsilon$, where υ is not observed and is the omitted variable that generates the endogeneity problem. This

[13] We check to make sure that there is no downward trend in the generic price data, the existence of which would suggest that generic prices have not yet flattened out as predicted by Grabowski and Vernon (1992). Fortunately, the scatter plots of average generic prices in each period exhibit fluctuations around the same level. The expectation that the generic prices will show a downward trend in the data is a characteristic of younger drug markets (where the switch to off-patent occurred relatively recently). In our situation, the generic and branded markets are mature and they have coexisted for decades, hence, it is not surprising that our generic prices do not exhibit such a downward trend. We thank an anonymous reviewer for suggesting this test of instrument validity. See also the online supplement.

[14] 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.

[15] The growth of information content during 2005 is most likely due to the FDA's announcement of the results of a clinical study, at the end of 2004, which indicated that patients taking naproxen sodium (Aleve) may be at an increased risk of heart attack or stroke (the withdrawal of Vioxx was also associated with this clinical study). By the end of January 2005, sales of Aleve plummeted by more than 50%, suffering the largest decline in brand history.

first stage regression yields residuals $\hat{v}$, which we include in the ordered probit in the second stage of the estimation.[16] We report first stage R^2 results in the footnote of Table 3. The estimation of this ordered probit provides consistent estimates of all of the parameters. Here, we are not interested in the magnitude of the parameters of the information content relationship (i.e., β) but rather in the marginal effects of a change in the endogenous variables, which we can consistently estimate using Blundell and Powell's (2003) approach.

7. Empirical Findings

Table 3 reports the findings of the various specifications estimated with our data. We investigate whether brands with higher vertical quality transmit more information, by associating the number of information cues (our dependent variable) with the values of the exogenous medical characteristics of the active ingredients. The first column of Table 3 includes number-needed-to treat (NNT), relative speed, and measures of cardiovascular and gastrointestinal risk as such explanatory variables[17].

Our results are consistent with the predictions of our theoretical model. Overall, the results in the first column of Table 3 suggest that there is a positive relationship between the amount of information provided and a stronger vertical differentiation of the brands. Keeping all else constant, we find that brands with inherently stronger pain relief have advertisements with more information content. As explained in the Appendix A, for NNT (as well as for Relative Speed, CV Risk and GI Risk), a higher number implies worse performance and a less effective drug. Thus, the negative

[16] The variable *Comparative* is a dummy variable. Here we run a linear probability model.

[17] We transform the branded and generic sales numbers as well as vertical characteristics levels into standardized values. Essentially we produce a variable with mean zero and standard deviation 1: such standardized variable represents how many standard deviations each observation is from the mean (which is zero).

coefficient on NNT is consistent with strong and efficient drugs having more information in their advertisements. We find a similar pattern for brands that have lower gastrointestinal and cardiovascular risks: their advertisements also tend to be more informative. Finally, consistent with our theoretical framework, brands that offer faster relief also have more information content.

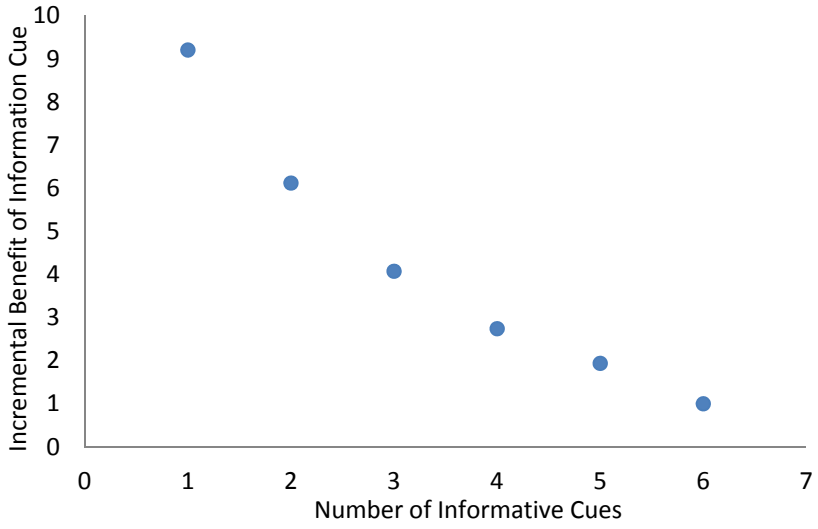
Table 3. Determinants of Information Disclosure

	(1)	(2)	(3)
Comparative?		0.476***	0.968***
		(0.026)	(0.077)
Standardized Sales		0.445***	0.309***
		(0.027)	(0.036)
Standardized Sales Squared		-0.204***	-0.130***
		(0.017)	(0.021)
Standardized Generic Sales		-0.136***	-0.140***
		(0.043)	(0.045)
Standardized NNT	-0.724***	-0.785***	-0.670***
	(0.029)	(0.047)	(0.050)
Relative Speed	0.336***	0.260***	0.198***
	(0.041)	(0.047)	(0.048)
Standardized GI Risk	-0.133***	-0.151***	-0.189***
	(0.035)	(0.056)	(0.059)
Standardized CV Risk	-0.151***	-0.478***	-0.465***
	(0.029)	(0.053)	(0.054)
Residuals-Comparative			-0.557***
			(0.082)
Residuals-Sales			-0.175**
			(0.068)
Residuals-Generic Sales			-0.167***
			(0.041)
Cutoff (0->1 Cues)	-3.085***		
	(0.194)		
Cutoff (1->2 Cues)	-1.041***	-1.199***	-1.042***
	(0.090)	(0.100)	(0.104)
Cutoff (2->3 Cues)	0.291***	0.196**	0.360***
	(0.089)	(0.098)	(0.103)
Cutoff (3->4 Cues)	1.095***	1.054***	1.218***
	(0.090)	(0.099)	(0.103)
Cutoff (4->5 Cues)	2.032***	2.029***	2.196***
	(0.091)	(0.100)	(0.105)
Cutoff (5->6 Cues)	3.034***	3.035***	3.213***
	(0.097)	(0.105)	(0.109)
# of Obs.	9739	9708	9708
Log-Likelihood	-13847.5	-13378.8	-13351.8

note: *** p<0.01, ** p<0.05, * p<0.1; Bootstrapped standard errors reported in Column 4. Endogenous variables: Comparative?, Standardized Sales, Standardized Generic Sales. Instrumental Variables: Functions of Generic Prices. The first stage R^2 for endogenous variables is the following: Comparative? - R^2 =.35, Standardized Sales - R^2= 0.96, Standardized Generic Sales - R^2=0.77)

The cutoffs estimated in these regressions have a clear interpretation. In particular, we showed that $\alpha_{i+1} - \alpha_i = \frac{I_i - I_{i+1}}{\bar{s}}$. In our analysis $\bar{s}$ is unknown, so assume for the sake of interpretation that $\bar{s} = 1$.[18] Then, for each i we can compute the difference I_i - I_{i+1}, which is the incremental benefit of the i^{th} information cue i. Consider the case of i=5. Then, $I_5 - I_6 = 1.002$. We can then find the point (6, 1.002) in Figure 3 to represent the empirical analogue of the theoretical incremental benefit corresponding to six information cues that we had derived in Panel B of Figure 2. We can replicate this exercise for the case of i=4, for which the scatter point will correspond to (5, 1.939), which is equal to 1.002 + 0.937. By repeating this exercise for all i = 1,…,6, we derive the scatter plot in Figure 3. What we see in Figure 3 is remarkably similar to what we derive theoretically in Panel B of Figure 1: there is a fundamental trade-off between information and persuasion content that firms face when preparing an advertisement. This relationship, depicting the diminishing marginal returns to information, holds in every specification that we estimate and looks very similar to the one illustrated in Figure 3.

Figure 3. Estimated Decreasing Marginal Returns of Informative Content



[18] Instead of normalizing $\bar{s} = 1$, one could also report $\frac{\alpha_{i+1} - \alpha_i}{\alpha_6 - \alpha_1}$, but that would only rescale Figure 2. The normalization $\bar{s} = 1$ gives a more intuitive interpretation for Figure 2.

In Column 2 we include the variables that approximate brand share, brand share squared (to capture a possible nonlinear relationship), and the share of the brand's generic counterpart[19].

To begin with, the evidence is also consistent with a relatively large spillover effect from informative advertising, as predicted by the theoretical model. We find that branded firms include less information content when the size of their generic competitors is large because the parameter of standardized generic sales is negative.

Next, we consider the other relationships of interest discussed in Section 4. First, we find that the largest brands transmit less information than the rest of the brands. Second, we find that the comparative advertising dummy is highly statistically significant and positive. This result indicates that comparative advertisements have more informational cues than self-promotional ads and that the likelihood that an advertisement is comparative increases with the number of cues. Note that comparative advertisements will always have at least one cue because a brand must compare itself with another brand along at least one dimension. Therefore, the analysis examines whether comparative and self-promotional advertising information content is different, conditional on having at least one information cue. Here, we treat the choice of using a comparative advertisement as an exogenous variable.

In Column 3 we follow an instrumental variable approach to estimate the effect of comparative advertisements, sales, and generic sales on information content. Column 3 uses the generic prices as instrumental variables. The results in Column 3 are those that we will use to discuss the main implications of our analysis. We find that the estimated coefficient of comparative advertisements is larger than in Column 2, implying that comparative advertisements have more information content than self-promotional advertisements and that such information content is

[19] Since only the relative market share position is important in our empirical model, we approximate brand and generic counterpart market shares with standardized values of respective monthly sales.

significantly higher than under the exogenous treatment. The strong endogeneity of this dummy variable is confirmed by the estimated coefficient of the control function, which is statistically significant at the 1% confidence level. Thus, we cannot reject the hypothesis that the variable "Comparative" is endogenous (Smith and Blundell, 1986). Therefore, our omitted variable concern in this setting is valid, and such unobserved quality is associated with the attractiveness of having comparative advertisements.

Both the coefficient on standardized sales and on the squared standardized sales are smaller in absolute value than in Column 2, suggesting a flatter, but still negative, relationship between brand shares and information content. The coefficient estimate of the control function is statistically significant, confirming the endogeneity of shares.

Finally, the marginal impact of the generic counterpart's market share is the same as in Column 2, suggesting that its endogeneity is not empirically significant. The coefficient is negative, implying the negative relationship between the generic counterpart and information disclosure, as suggested by our extended theoretical framework.

To understand the economic importance of the results regarding a brand and its generic counterpart market sizes, we construct figures that associate a brand's probability of choosing a given amount of information (i.e., a certain number of cues) with the distribution of a brand's share and the share of the market for a generic substitute (at the 10th, 25th, 50th, 75th, and 90th percentiles of market share distribution). For example, Figure 4 shows that the likelihood of an advertisement including only one cue increases sharply with size. In contrast, the probability of observing an advertisement with four or more cues decreases with size. Specifically, a move from the 10th percentile in the size distribution to the 90th percentile increases the probability of providing only one cue by approximately 16%. Similarly, Figure 5 represents the changes in the likelihood that an advertisement includes a certain number of information cues depending on the

size of the generic counterpart market share. For example, the likelihood that a product with a large generic counterpart will include one information cue is approximately 10% higher than the likelihood that a product with a small generic counterpart will include one information cue. The converse is true for three or more information cues: brands with smaller generic counterparts are significantly more likely to include more information in their advertisements.

Figure 4. Marginal effects on Information Disclosure by Brand Share

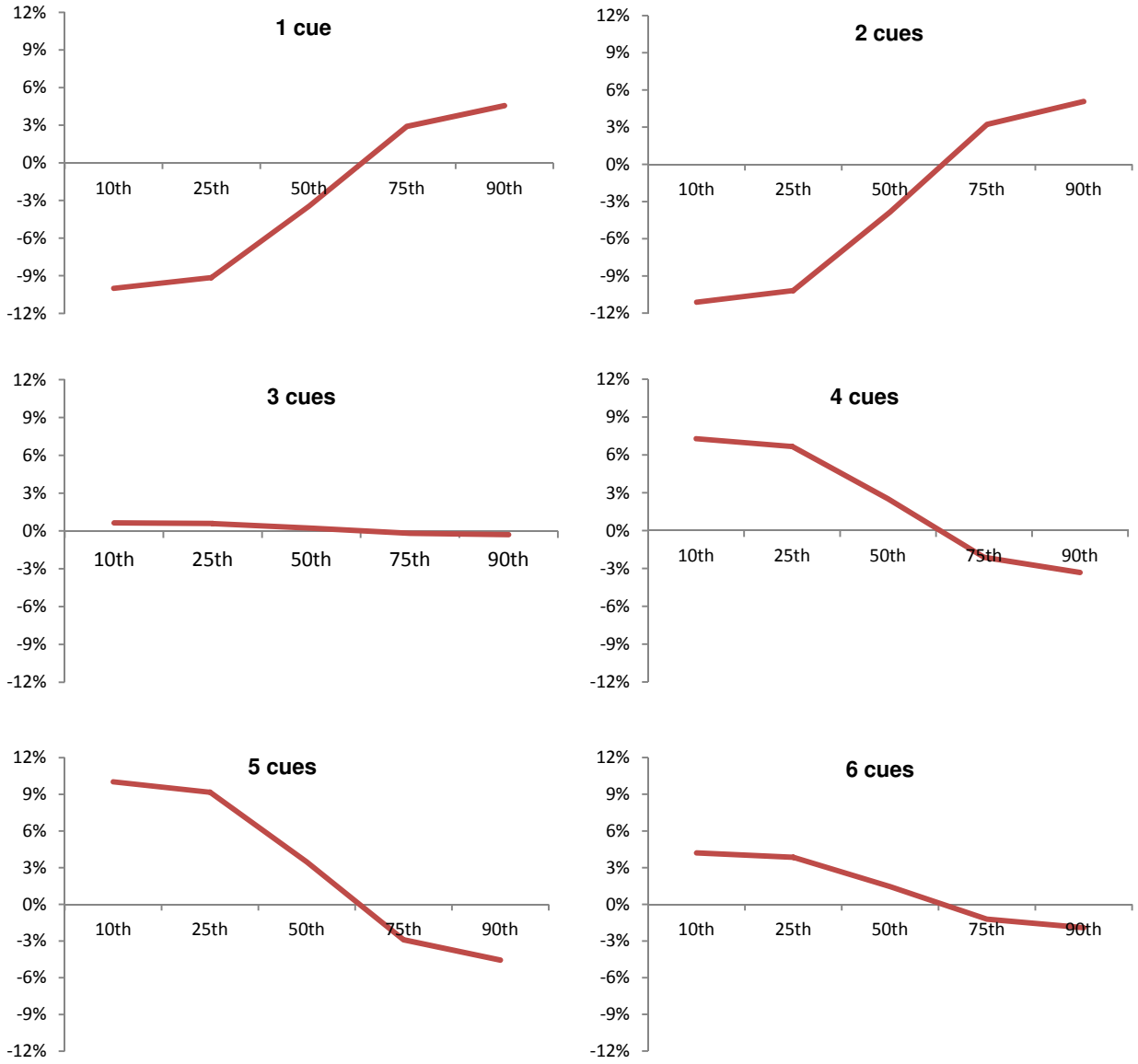
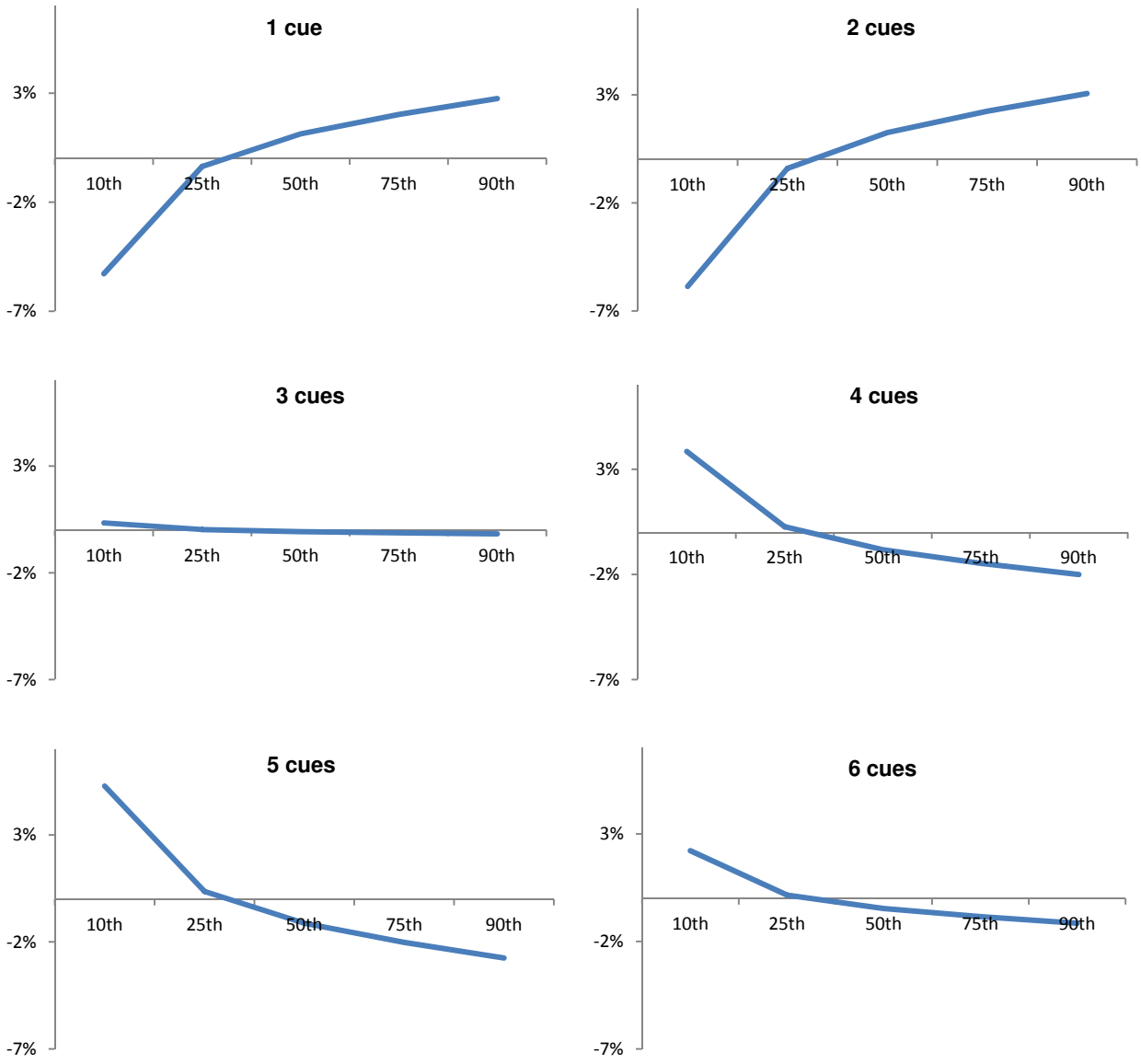


Figure 5. Marginal effects on Information Disclosure by Generic Counterpart Share



8. Concluding Remarks

We find that advertisements for brands with higher vertical quality disclose more information, and advertisements for brands with a high market share provide less information. These two results are not contradictory. The endogeneity of brand size underscores the importance of correcting for it. Otherwise, it might be assumed that larger brands are fundamentally of higher quality than smaller

brands and, therefore, their advertisements should have more information content. Correcting for endogeneity, we also show that more competition from generics gives rise to less information transmission by branded products, which is in line with the view that there are significant spillovers to informative advertising. Finally, we quantify the extent to which comparative advertisements have significantly more information content than self-promotional ones, and find that this effect is much larger than would be predicted without correcting for endogeneity.

From a methodological standpoint, we describe a method that is appropriate for dealing with information content with multiple explanatory variables, and show how the analysis can be corrected for endogeneity and how this alters the results. Our empirical analysis is restrictive in several aspects, and therefore suggests extensions that constitute themes for further research into information content. First, further research could use a sub-classification of cues (e.g., into vertical cues such as "fast" that all consumers would appreciate, or horizontal cues such as "headache" or "menstrual pain" that only some consumers desire) to explore the differential content of various cue types. Second, an information cue can be deployed only if a product has the attribute communicated in the cue, and can be used comparatively only if a product has an advantage over another product. Thus, further research might examine the amount of information advertised as a function of the total number of cues that could feasibly be advertised. Likewise, investigating whether products use comparative advertisements more often against similar or dissimilar products would be an interesting extension. Third, we do not examine product advertising campaigns in which advertisements address a subset of themes over a limited horizon. Fourth, we code only the objective content of advertisements as quantified by their reference of specific characteristics and competitors. We recognize that advertising may persuade through channels other than pure information, thereby leading consumers to act on emotional factors. While our paper recognizes that the trade-off between persuasive and informative content of an ad exists, we have not attempted to

code persuasive effects. The primary purpose of the empirical component of this paper was to measure the objective content of advertising along the lines of traditional content analysis, and incorporating the subjective side would be an important aspect to explore in future research. Finally, we do not address whether the market provision of information is optimal, how valuable the information is to consumers (Ippolito and Pappalardo, 2002; Pappalardo and Ringold, 2000), or how government policy towards advertising content might affect promotion effectiveness (Goldfarb and Tucker, 2011). Instead, the purpose of this paper was to document empirical regularities present in the data, rationalized via our *information-persuasion trade-off* theoretical framework, and to provide measures of the fundamental variables that can be used to answer such questions.

Our content analysis methodology can readily be applied to other product categories and industries. Extensions of the proposed methodology could be applied to answer questions such as: How does advertising information content differ for experience, search, and credence products? Does the relationship between brand size and information hold in more broad contexts? Do new products provide more information? These questions should hold both empirical and theoretical interest.

REFERENCES

Abernethy, A. M., Franke, G.R., 1996. The Information Content of Advertising: A Meta-Analysis. Journal of Advertising 25 (2), 1-17.

Ackerberg, D. A., 2001. Empirically Distinguishing Informative and Prestige Effects of Advertising. RAND Journal of Economics 32 (2), 316-333.

Ackerberg, D. A., 2003. Advertising, learning, and consumer choice in experience good markets: an empirical examination. International Economic Review 44 (3), 1007-1040.

Anderson, S. P., Ciliberto, F., Liaukonyte, J., Renault, R., 2012. Push-Me, Pull-You: Comparative Advertising in OTC Analgesics Industry. Working paper, University of Virginia.

Anderson, S.P., Renault, R., 2009. Comparative Advertising: Disclosing Horizontal Match Information. RAND Journal of Industrial Economics 40 (3), 558-81.

Bagwell, K., 2007. The Economic Analysis of Advertising. In: Armstrong, M. and Porter, R. (eds.), Handbook of Industrial Organization, Vol. 3, North Holland, Amsterdam, pp. 1701–1844.

Basuroy, S., Rao, V., Ravid, S. A., 2011. What is Advertising Content Worth? Evidence from the Motion Pictures Industry. Johnson School Research Paper Series, #45-2011.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., Zinman, J., 2010. What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment. Quarterly Journal of Economics 125 (1), 263-305.

Blundell, R., Powell, J.L., 2003. Endogeneity in Nonparametric and Semiparametric Regression Models. In: Dewatripont, M. , Hansen, L.P., Turnovsky, S.J., (eds.) Advances in Economics and Econometrics, Theory and Applications. Eighth World Congress, Volume II. Cambridge University Press: Cambridge, UK.

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

Cacioppo, J.T., Petty, R.E., 1985. Central and Peripheral Routes to Persuasion: The Role of Message Repetition. A.A. Mitchell, L.F. Alwitt, eds. Psychological Processes and Advertising Effects. Lawrence Erlbaum Associates , Hillsdale, NJ, 91-112.

Caves, R., Whinston, M.D., Hurwitz, M.A., 1991. Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry. Brookings Papers in Economic Activity, Microeconomics 1 (1), 1-48.

Chervany, N., Dickson, G., 1974. An experimental evaluation of information overload in a production environment. Management Science 20 (10), 1335–1344.

Ching, A.T.,Ishihara, M., 2012. Measuring the Informative and Persuasive Roles of Detailing on Prescribing Decisions. Management Science 58 (7), 1374-1387.

Ching, A., 2004. A Dynamic Oligopoly Structural Model for the Prescription Drug Market After Patent Expiration. Econometric Society 2004 Far Eastern Meetings 735, Econometric Society.

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

Ching, A.T., Clark, R., Horstmann, I., Lim, H., 2011. The Effects of Publicity on Demand: The Case of Anti-Cholesterol Drugs. Working paper, Rotman School of Management, University of Toronto.

Chou, L., Franke, G.R., Wilcox, G.B., 1987. The Information Content of Comparative Magazine Ads: A Longitudinal Analysis. Journalism Quarterly 64 (1), 119-24.

Christou, C., Vettas, N., 2008. Informative advertising and product differentiation. *International Journal of Industrial Organization.* **26**(1) 92-112.

Dekimpe, M.G., Hanssens, D.M., 1995. The Persistence of Marketing Effects on Sales. Marketing Science 14 (1), 1-21.

Erdem, K., Keane, M.P., Sun, B., 2008. The Impact of Advertising on Consumer Price Sensitivity in Experience Goods Markets. Quantitative Marketing and Economics 6 (2), 139-176.

Goldfarb, A., Tucker, C.E., 2011. Privacy Regulation and Online Advertising. Management Science 57 (1), 57–71.

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

Greene, W.H. 1997. Econometric Analysis, 3rd ed. Prentice Hall; London.

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

Harmon, R., Razzouk, N., Stern, B.L., 1983. The Information Content of Comparative Magazine Advertisements. Journal of Advertising 12 (4), 10-19.

Hong, J.W., Muderrisoglu, A., Zinkhan, G.M., 1987. Cultural Differences and Advertising Expression: A Comparative Content Analysis of Japanese and U.S. Magazine Advertising. Journal of Advertising 16 (1), 55-68.

Ippolito, P.M., Pappalardo, J.K., 2002. Advertising Nutrition and Health: Evidence from Food Advertising 1977-1997. Washington, DC: Bureau of Economics Staff Report, Federal Trade Commission.

Jacoby, J., 1977. Information load and decision quality: Some contested issues. Journal of Marketing Research 14 (4), 569–573.

Lancaster, K., 1971. Consumer Demand: A New Approach. Columbia University Press, New York&London.

Leffler, K. B., 1981. Persuasion or Information? The Economics of Prescription Drug Advertising. Journal of Law and Economics 24 (1), 45-74.

Liaukonyte, J., 2012. Is Comparative Advertising an Active Ingredient in the Market for Pain Relief? Working Paper, School of Applied Economics and Management, Cornell University.

MacInnis, D.J., Jaworski, B.J., 1989. Information Processing from Advertisements: Toward an Integrative Framework. Journal of Marketing 53 (4), 1-23.

MacInnis, D.J., Moorman, C., Jaworski, B.J., 1991. Enhancing and Measuring Consumers' Motivation, Opportunity, and Ability to Process Brand Information from Ads. Journal of Marketing 55 (4), 32-54.

Madden, C.S., Caballero, M.J., Matsukubo, S., 1986. Analysis of Information Content in U.S. and Japanese Magazine Advertising. Journal of Advertising 15 (3), 38-45.

Maheswaran, D., Meyers-Levy, J., 1990. The Influence of Message Framing and Issue Involvement. Journal of Marketing Research 27 (3), 361-367.

Pappalardo, J.K., Ringold, D.J., 2000. Regulating Commercial Speech in a Dynamic Environment: Forty Years of Margarine and Oil Advertising Before the NLEA. Journal of Public Policy & Marketing 19 (1), 74-92.

Pieters, R., Warlop, L., Wedel, M., 2002. Breaking Through the Clutter: Benefits of Advertisement Originality and Familiarity for Brand Attention and Memory. Management Science 48 (6), 765–781.

Pieters, R., Wedel, M., Zhang, J., 2007. Optimal Feature Advertising Design Under Competitive Clutter. Management Science 53 (11), 1815–1828.

Resnik, A.J., B.L. Stern. 1977. An Analysis of Information Content in Television Advertising. Journal of Marketing 41 (1), 50-53.

Rivers, D., Vuong, Q.H., 1988. Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. Journal of Econometrics 39 (3), 347-366.

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

Simon, C.J., Sullivan, M.J., 1993. The measurement and determinants of Brand Equity: A Financial Approach. Marketing Science 12 (1), 28-52.

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

Soberman, D.A., 2002. Research Note: Additional Learning and Implications on the Role of Informative Advertising. Management Science 50 (12), 1744–1750.

Stern, B.L., Krugman, D.M., Resnik, A.J., 1981. Magazine Advertising: An Analysis of Its Information Content. Journal of Advertising Research 21 (2), 39-44.

Stern, B.L., Resnik, A.J., 1991. Information Content in Television Advertising: A Replication and Extension. Journal of Advertising Research 31 (3), 36-46.

Sun, M., 2011. Disclosing Multiple Product Attributes. Journal of Economics & Management Strategy 20 (1), 195-224.

Vakratsas, D., Ambler, T., 1999. How Advertising Works: What Do We Really Know? Journal of Marketing 63 (1), 26-43.

Woolridge, J. 2001. Econometric Analysis of Cross Section and Panel Data. The MIT Press; Cambridge, MA.

APPENDIX A

Vertical Product Performance and Medical Measures

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

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

These 4 efficacy measures for each brand are reported in Table 1.

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

Relative risk is the risk of an event (e.g., developing a disease) relative to exposure. Relative risk is the ratio of the probability of the event (E) occurring in the exposed group versus the control (nonexposed) group:

$$RR = \frac{\Pr\,(E|treatment)}{\Pr\,(R|control)}$$

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

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

$$NNT = \frac{1}{P_B - P_A}$$

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

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

(2) Hyllested, M., Jones, S., Pedersen, J.L., Kehlet, H., 2002. Comparative Effect of Paracetamol, NSAIDs or Their Combination in Postoperative Pain Management: A Qualitative Review. British Journal of Anaesthesia 88, 199-214.

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

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

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

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

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

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

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

(10) Hersh, E., Moore, P.A., Ross, G.L., 2000. Over-the-Counter Analgesics and Antipyretics: A Critical Assessment. Clinical Therapeutics 22 (5), 500-548.