Competition in the pharmaceutical industry: how do quality differences shape advertising strategies?

Maria-Angeles de Frutos, Carmine Ornaghi and Georges Siotis*

February 23, 2012

Abstract

We present a Hotelling model of price and advertising competition between prescription drugs that differ in quality/side effects. Promotional effort results in the endogenous formation of two consumer groups: brand loyal and non-brand loyal ones. We show that advertising intensities are strategic substitutes, with the better quality drugs being the ones that are most advertised. This positive association stems from the higher rents that firms can extract from consumers whose brand loyalty is endogenously determined by promotional effort. The model's principal results on advertising and pricing strategies are taken to the data. The latter consists of product level data on price and quantities, product level advertising data as well as the qualitative information on drug quality contained in the Orange Book compiled by the Food and Drug Administration (FDA). The empirical results provide strong support to the model's predictions.

Keywords: Product differentiation, market segmentation, advertising, pharmaceutical industry.

JEL Classification Numbers: I11, L11, L13, L65 and M37

*de Frutos: Department of Economics, Universidad Carlos III de Madrid, Spain. Ornaghi: Department of Economics, University of Southampton, UK. Siotis: Universidad Carlos III de Madrid (on temporary leave), European Commission (member of the task Force for Greece), and CEPR. This paper was written before Siotis joined the European Commission and the views expressed are those of the author and do not necessarily reflect those of the European Commission. We would like to thank Nikos Vettas, Patrick Rey, Gerad Llobet and Natalia Fabra for helpful suggestions, and Marisa de Frutos whose pharmacological expertise greatly helped us navigate through the idiosyncracies of the industry. Valuable comments were also received from seminar participants at University of Edinburgh, University of Southampton, and at the CRETE2010 conference. Financial support from Project SEJ 2007-04339-001 is kindly acknolwledged by de Frutos. Siotis gratefully acknowledges support from projects SEJ2007-66268, ECO2010-20504 and I3 2006/04050/011.

1 Introduction

A particular feature of the market for prescription drugs is that patients usually do not establish their own diagnosis nor are they fully aware of the effectiveness or side effects associated with the different drugs. As a consequence, the choice of drug to administer is generally made by a physician. It may however also be the case that a patient expresses a preference for a drug over another, in particular if she has been exposed to some form of advertising.[1] Accordingly, a consumer is best represented by a physician-patient pair whose choice to address a given pathology is determined by the intrinsic characteristics of the available drugs, their prices, and promotional effort. In the US, the latter takes three forms. The bulk consists in "detailing" i.e., salespeople personally visiting doctors to promote a set of drugs, often leaving free samples in the process. The second type emerged in late 1996 when the US Food and Drug Administration (FDA) allowed "plain vanilla" advertising for prescription drugs, for instance via television ads.[2] Since then, spending on direct-to-consumer advertising (DTCA, from now on) has increased more than any other marketing activity (Iizuka (2004)). The last category is made-up of adverts appearing in specialised medical publications.[3]

Differences across consumers' responsiveness to price and advertising are, in pharma, at least as strong as in other industries. The combination of differences in insurance coverage across patients logically leads to heterogeneous responses to price. Doctors prescribing drugs to patients that benefit from a generous employer or State financed health coverage are unlikely to be very price sensitive (without however ignoring it altogether in their decisions), whereas physicians in hospitals and/or physicians attending uninsured patients are often well aware of the budgetary costs of their prescription decisions. In the same line, some doctor/patients are almost oblivious to promotional effort (in all its forms), whereas others tend to be more influenced by face-to-face meeting with sale representatives and TV ads, prescribing/consuming what they are most familiar with.[4] Promotional effort affects the proportion of doctor/patient pairs falling in each category, which we will respectively coin as "non-loyal" and "loyal".[5] In the absence of promotional effort, doctor/patients choices would solely be driven by price and intrinsic drug characteristics. If promotional effort were to tend to infinity, even the most reluctant patient/doctor pairs would end-up being influenced by it.

Pharmaceutical products are chemicals that improve the health of some humans but can cause serious side-effects in others. Consider, for instance, blood pressure control, the largest market in value terms, with worldwide sales exceeding 30 billion. Drugs to treat hypertension act via different parts of the body: central nervous system, heart (beta blockers), kidney (diuretics, saluretics), and vessels (alpha blockers,

[1] For cholesterol reducing drugs, Wosinska (2002) finds that direct to consumer advertising (DTCA) may affect the demand for an individual brand positively provided that brand is on the third party's payer formulary. This is also indirect evidence that, in the US, "price matters", albeit indirectly (via the presence on the formulary).

[2] The FDA introduced changes in August 1996. Prior to that date, rules stipulated that advertising had to provide detailed information on the drug, thus implying that TV ads were prohibitively expensive (because of their time length) in most cases. In the European Union (EU), direct to consumer advertising for prescription drugs remains prohibited.

[3] See Azoulay (2002) for evidence that advertising and scientific information stemming from clinical trials can affect physicians' prescription choices.

[4] Iusuka and Jin (2007) find that directed-to-physician advertising (i.e., detailing and medical journal advertising) has positive, significant, and long-lasting effects on the prescription choice of allergy drugs.

[5] This characterisation is akin to that found in Bala and Badgway (2010) who distinguish between "strong preference" patients whose choices are influenced by DTCA, and "other patients" who are not.

ACE inhibitors, AT1 and calcium antagonists). The efficacy of these drugs in terms of bringing blood pressure in the desired range differs across patients. In addition, they differ in terms of (numerous) side effects whose incidence vary substantially across the population. For some patients, a single molecule is a perfect cure: blood pressure is lowered within the desired range with no side effects. For other patients, the efficacy may be more limited (blood pressure lowered but above the optimal range) and side effects may be pervasive.[6] In short, one of the characteristics of pharmaceutical products is the existence of side-effects and/or contraindications that result in mismatch costs whenever consumers' ideal treatment is not available.

To capture the features of the pharma industry described above, we present a Hotelling model of competition among prescription drugs potentially characterised by different quality/side effects. Producers of these drugs compete both in prices and advertising. The latter gives rise to the endogenous formation of two consumer groups: brand loyal and non-brand loyal ones. We show that promotional effort and prices are strategic complements so that equilibrium prices are higher the more advertising firms do. Moreover, advertising efforts are strategic substitutes as they neutralise one another. This occurs because higher advertising by one firm results in a lower advertising by its rival.[7] By reducing its promotional effort, the firm enlarges the mass of non-loyal consumers on which it can focus while keeping prices at a relatively high level. In other words, it takes advantage of the fact that its rival price high as it has a large base of loyal consumers. In equilibrium, the firm that invest more in advertising is the one with a better quality drug. Thus, in our model, heterogeneity in firms' advertising behavior is driven by quality differentials. We perform comparative statics with respect to changes in the mismatch (transport) cost, in quality asymmetries and in the level of co-payment. We show that they all affect advertising levels and hence equilibrium prices. Larger co-payments or lower side effects both result in lower aggregate advertising expenditures and in lower prices.

Our results indicate that, for a given quality differential, the better quality drugs are also the ones that are most advertised. This positive association stems from the higher rents that firms can extract from consumers that endogenously exhibit brand loyalty as a consequence of promotional effort. It is however not possible to conclude that the link between quality and profit maximising advertising spend provides incentives to the development of superior drugs. On the one hand, while advertising increases profits of all firms, a firm with a lower quality product benefits relatively more from it. Potentially, this can have negative effects on the incentives to target path-breaking R&D. On the other hand, in the presence of large sunk costs, the prospect of large (absolute) profits may be necessary to induce firms to undertake risky research projects.

The model provides a number of testable hypotheses on pricing and advertising strategies that are taken to the data. The latter has been gathered by the market intelligence firm IMS-Health and consists of product level data that allows us to retrieve price and quantities. It encompasses the entire universe of prescription drugs sold in the US during the period 1994-2003. IMS sales data is complemented with product level DTCA data gathered by TNS Media and Intelligence/Competitive Media Reporting and

[6] Heart rhythm disorders, hypotension, impotence, mediastinal and gastro-intestinal disorders, abdominal pain, eye disorders, or subcutaneous tissue disorders are some of the side effects.

[7] The result is not driven by free-riding as we assume that market size is given and hence it is independent of firms' promotional efforts. In contrast, Linnosmaa (2008) assumes that the number of patients visiting a physician is determined by aggregate DTCA expenditures, thus attributing "public good" characteristics to DTCA.

detailing expenditure (promotion to office-based and hospital-based physicians) by IMS-Health. Last, proxies for drugs' quality are obtained from the Orange Book published by the Food and Drug Administration (FDA). In line with the prediction of the model, our results suggest that i) better products are advertised more intensively, ii) advertising has a positive impact on prices, and, iii) prices are higher in those markets where payers enjoy lower co-payment obligations.

Since the seminal work by Grossman and Shapiro (1984), several papers have investigated the role of advertising in markets with product differentiation. From a theoretical perspective our model shares some features with those of Brekke and Khun (2006) and Königbauer (2007). The former paper examines pricing and advertising decisions in a duopoly market where pharmaceutical firms use DTCA and detailing in sales promotion. Contrary to us, they focus on informative advertising and on drugs for which the loyal (monopolistic) segment of the market is not fully covered. As in our paper, Königbauer (2007) analyzes the impact of persuasive advertising on prescription decisions, but it focuses on the competition between a branded firm and a generic competitor.

Several empirical studies have analysed the competitive effect of advertising on prescription drugs' sales. The recent article by Dave and Saffer (2010) provides an exhaustive overview of the results of these studies. Two papers have studied the effect of advertising on price elasticity: Rizzo (1999) (for antihypertensive drugs), and Meyerhoefer and Zuvekas (2008) (for antidepressants). Both papers establish that advertising has a positive direct effect on sales (i.e., shifts demand outward). However, while the former reports that advertising reduces the price elasticity of demand (i.e., demand also rotates clockwise), the latter finds that advertising makes the demand more elastic. To the best of our knowledge, Dave and Saffer's paper (2010) and our's are the only papers that have directly estimated the effect of advertising on prices. Controlling for promotion aimed at physicians, they find that DTCA has a positive effect on price, with an estimated elasticity of 0.04. Their estimates of the impact of advertising on prices are qualitatively similar to those reported in our paper.

The contribution of this paper to the existing literature is threefold. First, we extend the analysis of price and advertising competition to an asymmetric environment, where drugs differ in their effectiveness and/or side effects. The model yields interesting prediction regarding the difference in advertising between "high" and "low" quality drugs as well as total advertising expenditures in a given market. Second, we identify all the necessary condition for an equilibrium to exist in a two stage game where drugs' producer first choose their advertising level and then set their prices. In particular, we provide a complete characterization of the firms' best-reply function in the presence of market segmentation (see Appendix A). The latter give rises to profits functions which fail to be quasiconcave. Third, the empirical results contributes to the literature by providing novel evidence on the relationship between advertising and quality differentials as well as on the effect of advertising and co-payments on prices.

The next section presents the general model, while section 3 focuses on the determination of market shares. Sections 4 and 5 respectively deal with price and advertising competition. Section 6 describes the data and presents our main empirical results, while section 7 concludes.

2 Model

Assume that the market for a therapeutic drug consists of a continuum of consumers (physician-patient pairs) with mass one uniformly distributed on the unit interval [0,1]. In that market, there are two branded firms located at the two extremes of this interval. These two brands compete in both advertising and prices. Each firm sells its own drug produced at the same constant marginal cost c. The cost of advertising expenditures is given by the increasing and convex function $C(a_i)$ with $C(0) = 0$. Convexity simply reflects that, at the margin, getting one extra sale requires more advertising.

Patients differ in terms of the possible side-effects they experience when exposed to a particular drug. The presence of promotional spend generates an additional source of heterogeneity. More precisely, advertising creates market segmentation with patient/physician pairs falling into two categories: brand-loyal and non-brand loyal.[8] The latter group is formed by consumers whose choice is solely driven by price and the intrinsic characteristic of a particular drug, i.e., their choice is not influenced by advertising. By contrast, in the brand loyal segment, choices are also determined by promotional effort (be it in the form of detailing directed at doctors, or DTCA, mainly directed at patients, but that could also affect physicians). We assume that promotional effort (both detailing and DTCA) is aimed at enhancing brand loyalty.

We first describe the behaviour of the **non-brand loyal segment**. We model competition for this segment by assuming that there is product differentiation à la Hotelling so that consumers (physicians) perceive the two drugs as horizontally differentiated and prescribe the most suitable drug in view of the pathology. The utility received by a non-loyal patient/physician pair located at x is given by:

$$U^{nl} = \max\{(u - t_0 d_{x-0}) - kp_0, (u - t_1 d_{x-1}) - kp_1, 0\}$$

where u is the utility received from treating the disease, d_{x-i} measures the distance of consumer x from drug i, k stands for the co-payment percentage, while the term t_i captures the utility loss ("mismatch costs") per unit distance from the most preferred drugs. The product $t_i d_{x-i}$ can thus be interpreted as the side effects associated with taking drug i. The introduction of different transport costs will allow us to analyse asymmetric equilibria that result from differences in the quality of the drugs (proxied by the magnitude of side effects). At this stage, and without loss of generality, we assume that the firm located at the beginning of the unit interval has a transport cost not greater than the one of its competitor, i.e., $t_0 \leq t_1$. We will denote total transport costs $t_0 + t_1$ by T. Note that a consumer located at point 0 (point 1) does not suffer from any side effects if she takes drug 0 (drug 1) since $d = 0$. In this setting, both drugs can treat the entire set of patients: from a therapeutic perspective, the difference between them lies in the side-effects.[9] Finally, p_i is the price of the drug. Prices are bounded from above by consumers' maximal reservation utility u/k. Further, we define $v = u - ck$; thus, v stands for the (social) surplus associated with treating the patients that do not suffer any side effects.

[8] Since the seminal paper by Frank and Salkever (1992), it is common to model the demand side of the market for prescription drugs as consisting of two segments (see, for instance, Regan (2008)).

[9] For instance, many blood pressure control drugs fit this description: they all reduce blood pressure but with varying intensity across the patient population: given a posology, a given drug may lead to a negligible reduction in hypertension, while for others it may lead to hypotension. In addition, these drugs differ in terms of the intensity of the side effects (tachycardia, sleepiness, eudemas, etc..). A similar comment applies to drugs used for the treatment of asthma, where LABA/ICS combinations and single LABAs are prescribed.

In the **brand loyal segment**, advertising is aimed at inducing subjective product differentiation and hence enhance loyalty to a particular brand. This is achieved by undertaking persuasive advertising thus inducing consumers (physicians/patients) to attach more importance to those differences that already exist between products.[10] Alternatively, detailing or advertising in medical journals reduces the physicians' costs of finding treatment information, making physicians more prone to prescribe drugs that are heavily promoted.

The utility received by a loyal patient/physician pair located at x and loyal to i is given by:

$$U^l = \max\{(u - t_i d_{x-i}) - kp_i, 0\}$$

The physician/patient pair subjected to detailing by i only contemplates the choice between not prescribing the treatment or prescribing i. In other words, the pair does not consider j as an alternative. As will be seen below, in equilibrium, all patients are prescribed a drug as long as their utility (net of pecuniary costs and side effects) is positive. Thus, the assumption that loyal practitioners/patients pairs focus on a single drug (and ignore the potential substitute) does not lead to patients needing treatment being left without it, nor does it imply "improper" prescriptions.

Our modelling assumption encompasses alternative interpretations for the emergence of brand loyalty, such as differences in doctors' habit prescriptions and in the time constraints they face for finding treatment information,[11] or the existence of patients that develop "strong" preferences for a particular drug after watching TV ads.[12] Our model also encompasses the case of "captive" consumers in Bagwell's (2007) taxonomy: patient/doctor pairs that are only aware of one drug because they have been have exposed to promotional effort by one firm.

We next turn to the distribution of patients across the two segments. Following Chioveanu (2008) we assume that the proportion of loyal consumers L is an increasing and concave function of the aggregate advertising expenditures of the firms,[13] i.e., $L = L(a_0 + a_1)$ with $\lim_{\sum a_i \to \infty} L(\sum a_i) = 1$, where a_i denotes the advertising expenditure chosen by firm i. In order to work with closed-form solutions, we further assume that $L = (a_0 + a_1)/(1 + a_0 + a_1)$. Brand-loyal consumers are split among the two firms according to the proportional market sharing function which goes back to Tullock's (1967) contest success function, $\mu_i = a_i/(a_0 + a_1)$. Consequently, the proportion of loyal consumers of firm i, θ_i, is given by:[14]

$$\theta_i = L \times \mu_i = \frac{a_0 + a_1}{1 + a_0 + a_1} \times \frac{a_i}{a_0 + a_1} = \frac{a_i}{1 + a_0 + a_1}.$$

[10] This is in line with von der Fehr and Stevin (1998), where advertising increases perceived product differences. See also Scott Morton (2000) for further arguments supporting this interpretation.

[11] Doctors that are used to prescribe drug i need to incur a "switching" cost before start prescribing drug j. For some doctors, these switching costs are sufficiently high to imply that drug j is never prescribed. See Linnosmaa (2008) for a model of prescription decisions based on the opportunity costs of physicians' time.

[12] Hollon (2005) reports that: "Survey data suggest that approximately 40% of visits in which a DTCA discussion occurs result in a prescription for the advertised drug. In more than half the cases, a physician prescribes the drug partly to accommodate a patient's request". See also Bala and Bhardway (2010) for a model in which patients differ in the strength of their brand-preference: some patients have "strong preferences" so that they get prescribed their preferred brand (similar to our modelisation of loyal consumers) while other have weak preferences and hence talk about both products with their physician (as our non-loyal consumers).

[13] For empirical evidence of diminishing returns to scale in advertising, see Bagwell (2007).

[14] In Baye and Morgan (2008), brand advertising also creates loyal consumers through an advertising response function. As in our paper, they assume a particular advertising response function to be able to get closed-form solutions.

A random utility model can lead to these reduced-forms whereby patient/doctor's utility is the sum of two independent components: the systematic component, representing what is known about the drug i (hence affected by a_i), and the disturbance term, which is randomly distributed. If we assume that all errors associated with the different choices are independent, identically distributed with type I extreme value distribution (in line with the Conditional Logit model), then the probability of choosing alternative i (θ_i in our model) can be written as:

$$\frac{\exp(x_i\beta_i)}{1+\exp(\exp(x_i\beta_i))+\exp(x_j\beta_j)} = \frac{a_i}{1+a_0+a_1} \text{ for } \exp(x_i\beta_i) \equiv a_i$$

Consequently, our modelling strategy can be thought of as a multinomial logit whereby only a fraction of patient/doctor pairs are exposed to advertising, and of the latter, only some are influenced by it.

In the empirical analysis of Section 6, we also distinguish between two distribution channels, hospitals vs. pharmacies. Within each channel, we assume that firms cannot price-discriminate between the two type of consumers. Thus, depending on the intensity of advertising that doctors are subjected to in each channel, equilibrium configurations may differ. For instance, this will occur if, for a given drug, promotional effort is principally (but not exclusively) directed at one particular distribution channel.

The timing of the game is as follows. In the first stage both firms simultaneously choose their advertising expenditure levels and in the second stage they compete in prices.[15] Thus, price and advertising strategies determine firms' market shares.

3 Market shares

Consumers/physicians decide which brand to consume by taking into account firms advertising effort and the prices set by them. The total number of consumers in each segment is determined by first-stage advertising strategies. Pricing strategies affect consumers decisions: whether to buy or not for a loyal consumer, and which firm to buy from for a non-loyal one.

A **loyal consumer** will compare the utility derived from consuming its preferred drug with the utility derived from the outside option of not getting any medical treatment. Thus, a loyal consumer located at x will buy drug i iff:

$$u - t_i d_{x-i} - kp_i \geq 0.$$

Firm i's market share of loyal consumers equals $\theta_i M_i^l$ where:

$$M_i^l = \begin{cases} 1 & \text{if} \quad kp_i \leq u - t_i, \\ \tilde{x}_i \in [0,1) & \text{if} \quad u - t_i < kp_i \leq u, \end{cases}$$

with $\tilde{x}_i = \frac{u-kp_i}{t_i}$. If firm i were to sell only to its loyal consumers, then its optimal price (kp_i) would be equal to $u - t_i$ if $v \geq 2t_i$, and $(u + ck)/2$ otherwise. Consequently, if $v \geq 2t_i$ holds, then in markets populated by loyal consumers, firm i's optimal strategy is to fully cover the market at the largest price that allows firm i to do so, namely at $kp_i = u - t_i$. In what follows, we assume that $v > 2t_1$ does

[15] Sequentiality comes from the fact that pharma companies, on the basis of phase III trials and the market intelligence at their disposal, decide whether "go for it" with heavy advertising. This reflect the fact that they have fairly precise ex-ante knowledge regarding the drug's potential. Posterior price adjustments come as fine tuning.

hold, so that we can restrict our attention to cases where firms find profitable to attend their entire loyal segment.[16] This appears as a sensible assumption (in particular, in a developed country context) given that we are modelling competition among prescription drugs. The latter are typically directed at addressing seriousness illnesses; thus, the benefit of treatment are presumably high. In the same line, the severity of side effects is often correlated with therapeutic benefits (for instance, it is a unlikely that a market for stomach burns drugs would have developed if they generated side effects similar to antipsychotics).

A **non-loyal consumer** will compare both product characteristics and chose the drug with the lower relative price. More precisely, a non-loyal consumer located at $\hat{x}$ on the unit interval will be indifferent between the two products if:

$$u - t_0\hat{x} - kp_0 = u - t_1(1-\hat{x}) - kp_1$$

so that:

$$\hat{x} = \frac{k(p_1 - p_0) + t_1}{t_0 + t_1},$$

where $\hat{x} \in (0,1)$ if $k(p_1 - p_0) \in (-t_1, t_0)$. The indifferent consumer will prefer buying from from either branded firm to non-consuming only if $U(\hat{x}, p_i) \geq 0$, which imposes the additional "participation constraint" $k(p_0 t_1 + t_0 p_1) \leq uT - t_0 t_1$. The participation constraint will be satisfied only if condition below holds:

$$kp_0 \leq \frac{uT - t_0 t_1}{t_1} - \frac{k t_o p_1}{t_1} \equiv \Xi(p_1). \tag{1}$$

Thus, if (1) holds then the market share of brand 0 equals $\hat{x}$ and the one of brand 1 equals $1 - \hat{x}$.

For sufficient low prices $kp_0 \leq kp_1 - t_0$, firm 0 market share of non-loyal consumers will equal one as for such low prices U^{nl} is non-negative and any non-loyal consumers always prefers firm 0. For higher prices, some non-loyal consumers may prefer not to buy. In particular, since $U(\hat{x}, kp_0 = kp_1 + t_1) = 0$ at $kp_1 = u - t_1$, it follows that any $kp_0 < kp_1 + t_1$ will satisfy (1) if $kp_1 \leq u - t_1$ holds. By contrast, for $kp_1 > u - t_1$, the non-loyal consumer $\hat{x}$ will buy product 0 only if offered a price $kp_0 < \Xi(p_1) < kp_1 + t_1$. For $kp_0 > \Xi(p_1)$, the fact that (1) is binding implies that the proportion of non-loyal consumer buying drug 0 is defined by the location of the consumer that is indifferent between buying drug 0 or buying nothing (exactly as for the loyal market). Therefore, the non-loyal market will not be fully covered. Finally, if $kp_0 \geq kp_1 + t_1$ then no non-loyal consumer will buy brand 0.

Accordingly, using $\hat{x}_i$ to measure the distance of the indifferent consumer from drug i, with $\hat{x}_0 = \hat{x}$ and $\hat{x}_1 = 1 - \hat{x}$, firm i's market share of non-loyal consumers is given by $(1 - \theta_0 - \theta_1)M_i^{nl}$ where:

$$M_i^{nl} = \begin{cases} 1 & \text{if} \quad kp_i \leq kp_j - t_i, \\ \hat{x}_i \in [0,1) & \text{if} \quad kp_j - t_i < kp_i \leq \min\{\Xi(p_j), kp_j + t_j\}, \\ \frac{u - kp_i}{t_i} < \hat{x}_i & \text{if} \quad kp_i \in (\Xi(p_j), u) \text{ and } kp_j > u - t_j. \end{cases}$$

To determine the overall market share we focus next on the relationship between M_i^{nl} and M_i^l, noting that $M_i^{nl} \subseteq M_i^l$. To this end, it is useful to compare the smallest kp_i at which firm i loses all its non-loyal

[16] Brekke and Khun (2006) focus on price competition in markets with identical side effects for both drugs and with elastic monopolistic demand, i.e., with $M_i^L < 1$, by assuming that $u \in (t, 2t)$. By contrast, we will here focus on markets for which $v > 2t_1$, for which all needy patients receive a treatment.

consumers ($kp_j + t_j$) with the largest kp_i at which it maintains its entire loyal base ($u - t_i$). Note that $\min\{kp_j + t_j, u - t_i\} = u - t_i$ if $kp_j \geq u - t_i - t_j$. Thus, for firm j prices below $u - t_i - t_j$, firm i setting a price slightly above $kp_j + t_j$ implies that it will sell nothing to non-loyal consumers, while selling to all its loyal. For prices of firm j belonging to $(u - t_i - t_j, u - t_j]$, if firm i sets a price above $u - t_i$, it will lose part of its loyal consumers but can still attend some non-loyal ones. Finally, for firm j prices above $u - t_j$, firm i that sets a price belonging to $(\Xi(p_j), kp_j + t_j)$ gets an identical market share in the two segments with $M_i^{nl} = M_i^l = \tilde{x}_i < \hat{x}_i$. Thus, to determine firm i's market share, one must distinguish among the three aforementioned cases. The third case ($kp_j > u - t_j$) can be ignored, as no firm will find it optimal to exclude any of its loyal given the assumption that $v \geq 2t_i$. In other words, for firm j, $kp_j = u - t_j$ strictly dominates all higher prices, so that $p_j^{\max} = (u - t_j)/k$.[17]

Based on the discussion above, if $kp_j \leq u - t_i - t_j$, then firm i loses all non-loyal consumers before the first loyal one drops from its consumer base. Firm i's market share $S_i = \theta_i M_i^l + (1 - \theta_i - \theta_j) M_i^{nl}$ is given by:

$$S_i = \begin{cases} \theta_i + (1 - \theta_i - \theta_j) & \text{if} \quad kp_i \leq kp_j - t_i, \\ \theta_i + (1 - \theta_i - \theta_j)\hat{x}_i & \text{if} \quad kp_j - t_i < kp_i < kp_j + t_j \\ \theta_i & \text{if} \quad kp_j + t_j \leq kp_i \leq u - t_i \end{cases}$$

If $kp_j \in (u - t_i - t_j, u - t_j)$, firm i loses some loyal consumers before the non-loyal consumer $\hat{x}$ switches suppliers so that:

$$S_i = \begin{cases} \theta_i + (1 - \theta_i - \theta_j) & \text{if} \quad kp_i \leq kp_j - t_i, \\ \theta_i + (1 - \theta_i - \theta_j)\hat{x}_i & \text{if} \quad kp_j - t_i < kp_i \leq u - t_i. \end{cases}$$

4 Price Competition

The existence of two market segments generates profit functions that may fail to be quasi-concave. This is due to the fact that firms enjoy monopoly power on their "captive" (or loyal) consumers and may have an incentive to set a high price even if such a high price implies losing all non-loyal consumers.[18]

More precisely, for any $kp_j \leq u - t_i - t_j$, firm i's profit function fails to be quasi-concave as they have two local maxima, one where $kp_i \in (kp_j - t_i, kp_j + t_j)$ and another one at $kp_i = u - t_i$, which is the best price to charge when only loyal consumers buy. By contrast, if $kp_j > u - t_i - t_j$, then the profit function has a unique maximum which lies in the region of prices for which $S_i = \theta_i + (1 - \theta_i - \theta_j)\hat{x}$. This maximum is either the unconstrained maximum, i.e., the point which maximizes $(p_i - c)S_i$, or the demand kink, i.e., the largest price that preserves the loyal base, namely $kp_i = u - t_i$. A numerical example helps to illustrate this point. Consider a symmetric mismatch cost environment $t_0 = t_1 = 1$. Let $a_0 = 0.75$, $a_1 = 0.7$ with $v = 4$ ($u = 4$, $c = 0$) and $k = 0.65$. At $kp_j = 1.49 < u - 2t_i = 2$, firm i profit function has

[17] This implies that the strategy space can be restricted to pairs $(p_0, p_1) \in [0, (u - t_0)/k] \times [0, (u - t_1)/k]$.

[18] The non-quasi-concavity of the profit function is a common feature of any model in which advertising creates market segmentation, in either localized or non-localized competition. Furthermore, it is independent of whether advertising is considered to be informative or persuasive. Grossman and Shapiro (1984) do not examine this possibility by implicitly restricting the analysis to parameter ranges where a deviation to a high price is not profitable. See Christou and Vettas (2004) for a further discussion on this issue.

two maxima whereas at $kp_j = 2.4 > 2$ it has a unique maximum (see Figures below).

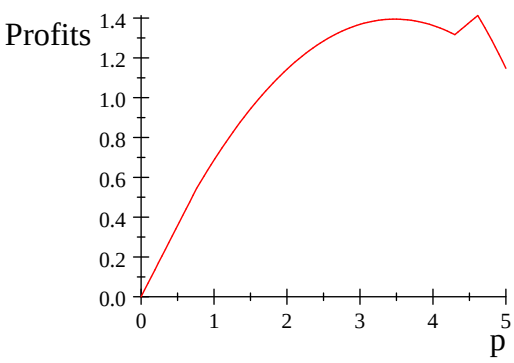

Figure 1.a. Firm i profits for $kp_j = 1.49$

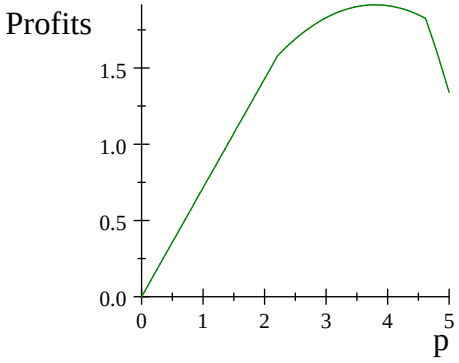

Figure 1.b. Firm i profits for $kp_j = 2.4$

Since the profit functions may fail to be quasi-concave, it is necessary to fully characterize firms' best replies, an issue that is explored in Appendix A. In what follows, we concentrate on the interior price equilibrium in pure-strategy.

Proposition 1 *i) There is a pure strategy price equilibrium in every second-stage subgame if and only if* $v - T \leq 3t_0$. *Moreover, if* $v - T \leq 3t_0$ *holds, then the pure-strategies price equilibrium is unique and it entails:*

$$S_i = \theta_i + (1 - \theta_0 - \theta_1)\left(\frac{k(p_j - p_i) + t_j}{T}\right) \tag{2}$$

ii) The pure-strategy equilibrium is interior if $v - T > \frac{4t_1+2t_0}{3}$ *holds.*

Proof. See Appendix B. ■

There are several reasons that justify the choice of focusing on the case where $v - T \in \left(\frac{4t_1+2t_0}{3}, 3t_0\right)$ for which we have a unique interior equilibrium in pure-strategy.[19] First, the condition $t_1 < 7t_0/4$ (which ensures that $\frac{4t_1+2t_0}{3} < 3t_0$) implies that the differences in mismatch costs between the two drugs are not too large, meaning that they are therapeutic substitutes in a "genuine" sense. This case encompasses the symmetric case ($t_0 = t_1$) and, more generally, competition between a drug and its "me-too" substitutes. Second, in this parameter constellation, prices are strategic complements, as each firm responds to a price increase by increasing its own price. Last (but not least), the tractability associated with this parameter constellation allows us to obtain testable predictions on equilibrium advertising strategies.

It is interesting to note that setting firm i demand equal to S_i in (2) implies that the price elasticity of the demand faced by firm i is given by:

$$\varepsilon_i = k\frac{p_i}{t_j + a_i T + k\left(p_j - p_i\right)},$$

which increases with the level of co-payment and decreases with advertising effort. The sign of these effects are consistent with existing empirical findings. An increase in a_i generates a clockwise rotation of

[19] The condition $v - T > \frac{4t_1+2t_0}{3}$ implies that both firms serve both loyal and non-loyal consumers. Note that this particular equilibrium can hold for a larger constellation of parameters given that the restriction $v - T > \frac{4t_1+2t_0}{3}$ is a sufficient but not a necessary condition to have an interior equilibrium in pure-strategies.

the inverse demand curve rather than a parallel translation, meaning that advertising activities involve the provision of loyalty enhancing "real" information in the Johnson and Myatt (2006) taxonomy of advertising.

The profits associated with the demand S_i in (2) are given by:

$$\Pi_i = (p_i - c)\frac{1}{1 + a_0 + a_1}\left(a_i + \left(\frac{k(p_j - p_i) + t_j}{t_0 + t_1}\right)\right).$$

The first-order condition for profit maximization with respect to p_i yields the following best-response $p_i(p_j)$:

$$p_i(p_j) = \frac{p_j + c}{2} + \frac{(t_j + a_i T)}{2k}, \tag{3}$$

Using (3), equilibrium prices are given by:

$$p_i = c + \frac{(2a_i + a_j + 1)\,T + t_j}{3k} \tag{4}$$

and the associated equilibrium market shares and profits are given by:

$$\begin{aligned} S_i &= \frac{T\,(2a_i + a_j + 1) + t_j}{3T(1 + a_0 + a_1)}, \\ \Pi_i &= \frac{(t_j + (1 + 2a_i + a_j)\,T)^2}{9kT(1 + a_0 + a_1)} - C(a_i). \end{aligned}$$

Note that the price of brand i increases with both own and rival's advertising , and decreases as the co-payment decreases. A larger a_i implies a larger proportion of captive consumers in firm $i's$ consumer base, thus leading to a higher price. Similarly, as a_j increases, the proportion of loyal consumers increases in the rival's consumer base, and therefore it is less costly to increase prices. Regarding co-payments, as k increases, the non-loyal segment becomes more price-sensitive which pushes down firm $i's$ equilibrium price. The next lemma summarizes (straightforward) comparative static results on equilibrium prices.

Lemma 1 *i) Advertisement by firm i increases both firms' equilibrium prices.*

ii) Advertisement by firm i increases its market share (and, trivially, decreases its rival's in the same proportion).

iii) Advertisement strategies are complementary (substitutive) for the equilibrium market shares if a lower t_i implies a higher (lower) a_i.

Thus, the model generates the well-known result that advertising softens price competition. Firm j benefits from the higher price set by its rival as it triggers its own price increase. As a_i increases, S_i becomes more price inelastic which results in a higher equilibrium price.

5 Competition in Advertising

At the first stage of the game, firms choose their advertising expenditures taking into account the impact of their investments on the equilibrium prices that ensue.

Firms' payoffs are the profits emerging from their equilibrium pricing strategies minus the costs of advertising:

$$\Pi_i = \frac{\left(t_j + \left(1 + 2a_i + a_j\right)T\right)^2}{9kT\left(1 + a_0 + a_1\right)} - C(a_i) \tag{5}$$

while marginal returns are given by:

$$\frac{\partial \Pi_i}{\partial a_i} = \frac{\partial p_i}{\partial a_i} S_i + p_i \frac{\partial S_i}{\partial a_i} - C'(a_i).$$

In what follows, we first derive some general properties of the profit functions. We then solve the symmetric case that is used as a benchmark. Finally, we analyze the asymmetric case and provide comparative statics results. To do so, it is convenient to express t_0 by t and t_1 by βt, with $\beta > 1$. By doing so, we can analyze the impact on advertising and prices of changes in the rival's relative quality (as measured by changes in β). This formulation also permits an analysis of a change in the degree of product differentiation (as measured by t) and in the co-payment parameter (k). Note that a ceteris paribus fall in td represents an across the board quality improvement in the form of reduced side effects for all patients taking the drug.

We first note that payoff functions are submodular. Submodularity has a negative complementarity interpretation: the marginal returns to increasing advertising expenditures decrease with the rival's advertising expenditure (see Vives (2006) for further details). Hence, the advertising strategies are strategic substitutes given the negative sign of the second-order cross derivative:

$$\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j} = -\frac{2\left(t_j + a_i T\right)\left(t_i + a_j T\right)}{9kT\left(1 + a_0 + a_1\right)^3} < 0$$

The fact that firms take into account their first and second stage decisions allows us to gain some insights into the strategic substitutability of advertising strategies. A larger first stage investment implies softer price competition. This is because firm $i's$ best reply in the pricing stage increases with a_i. If a firm increases its advertising effort, its rival is better off by reducing its own because, in addition to generating savings on advertising costs, it also increases the mass of the non-loyal segment on which it can focus while keeping prices at a relatively high level.

The submodularity property of the payoff function is important as it ensures existence of equilibrium without requiring concavity assumptions on profits. Moreover, the equilibrium will be unique provided that the advertising costs are sufficiently convex. A sufficient condition for this to occur is that $C''(a_i) > \frac{2T}{9k}$ holds. We will refer to this condition as *Condition U*.

Condition U: The cost of advertising satisfies $C''(a_i) > \frac{2(t_0+t_1)}{9k}$.

The Proposition that follows summarizes the discussion.

Proposition 2 *The advertising game is submodular so that it has a non-empty set of equilibria. Moreover, if Condition U holds, the equilibrium is unique, and it is given by the solution to firms' first order conditions.*

Proof. See Appendix B. ■

5.1 Benchmark case: Identical side effects

We start by assuming that firms are ex-ante symmetric, i.e., that $t_0 = t_1 = t$. The motivation for carrying out this exercise is two-fold. On the one hand, it allows us to construct a reference benchmark to analyze the impact of asymmetries. On the other hand, by focusing on the symmetric equilibria, we can compare our results with those in Brekke and Kuhn (2006) (BK, for short) as they focus on symmetric equilibria.

In order to work with an explicit solution, we assume an advertising cost given by $C(a_i) = a_i^2$. Note that for this particular functional form, condition U holds for $k \geq 2t/9$. Substituting $T = t_0 + t_1 = 2t$ and $C(a_i) = a_i^2$ in (5) , and exploiting the first order condition $\frac{\partial \Pi_i}{\partial a_i}$, we find that there is a unique symmetric equilibrium given by:

$$\begin{aligned} a_0 &= a_1 = \frac{5}{12k}t. \\ p_0 &= p_1 = c + \frac{(6k+5t)\,t}{6k^2} \end{aligned}$$

The two equations above show that both advertising and prices increase with t and decrease with k. The effect of higher co-payments is identical in BK's and our model. A higher co-payment increases demand elasticity and therefore curbs equilibrium prices and advertising intensities. Similarly, in both models, equilibrium prices increase as the degree of differentiation increases (higher t). The difference lies in the effect of an increase in the degree of differentiation on advertising efforts. In our model, there is a positive relationship between them. By contrast, in BK, the result only holds if t is large enough. The reason behind this different result is related to the price inelasticity (elasticity) of the loyal segment. In BK, an increase in t reduces the demand of the monopolistic segment. This demand reducing effect dominates the price-increasing effect for low values of t. Since here there is no demand reduction effect, a larger t unambiguously increases advertising.

5.2 Different side effects in the presence of co-payments

Since one of the characteristics of pharmaceutical products is the existence of side-effects and/or contraindications that are specific to a given patient/drug pair, we next solve the advertising followed by prices game when transport costs are drug-specific. More precisely, we consider the situation where $t_0 = t$ and $t_1 = \beta t$. The first consequence of unequal transport costs is that the equilibrium will be asymmetric, as $a_i \neq a_j$.

The advertising strategies of the two firms are obtained by maximising (5) with respect to a_i. Using $T = t_0 + t_1 = (1+\beta)t$, the solution to the problem is defined by the following first-order conditions:

$$\begin{aligned} \frac{\partial \Pi_0}{\partial a_0} &= \frac{t\left(\beta + (1+\beta)\left(1 + 2a_0 + a_1\right)\right)\left(1 + (1+\beta)\left(2 + 2a_0 + 3a_1\right)\right)}{9k\,(1+\beta)\left(1 + a_0 + a_1\right)^2} - C'(a_0) = 0 \\ \frac{\partial \Pi_1}{\partial a_1} &= \frac{t\left(1 + (1+\beta)\left(1 + a_0 + 2a_1\right)\right)\left(\beta + (1+\beta)\left(2 + 3a_0 + 2a_1\right)\right)}{9k\,(1+\beta)\left(1 + a_0 + a_1\right)^2} - C'(a_1) = 0 \end{aligned}$$

We first focus on the direction of change of the equilibrium advertising defined by the two equations above as β changes from $\beta = 1$ to $\beta > 1$. We may think of β as a measure of brand 0's quality vis-a-vis its closest therapeutic substitute. In other words, if $\beta > 1$, then brand 0 ought to be the market leader given that, ceteris paribus, it will have a market share larger than its rival as the median non-loyal patient will always prefer brand 0 to brand 1.

Proposition 3 *i) If $\beta > 1$ then the equilibrium is asymmetric with $a_0 > a_1$.*

ii) Aggregate advertising increases with β.

iii) If Condition U holds then a_0, a_1 and p_0, p_1 increase as β increases.

Proof. See Appendix B. ■

Proposition 3 provides a number of insights. First, the market leader will undertake more advertising than its rival (part i). The asymmetric equilibria emerge as brand 0 becomes more responsive to changes in β: its market share is larger, so that $\frac{\partial p_i}{\partial a_i} S_i$ is larger for brand 0 than it is for brand 1. Second, since both reaction functions shift outwards, aggregate expenditures increase and they become larger for those drugs with a larger quality differential (part ii). The impact on each equilibrium advertising is less clear. On the one hand there is a direct effect: as β increases both best replies shift outwards. On the other hand, there is also an indirect effect: as firm i increases its advertising intensity, firm j wants to reduce its own. The overall effect therefore depends on the relative strength of these two opposing forces. We show that the direct effect dominates so that as the quality differential increases both firms react by increasing both their equilibrium advertising and their prices (part iii).

To gain some insight on the importance of quality differentials on advertising and prices, we provide a numerical resolution of the game when the advertising cost function takes the form $C(a_i) = a_i^2$. As Table 1 illustrates the market leader (the firm whose drug produces lower side effects, i.e. firm 0 whenever $\beta > 1$ and firm 1 otherwise), charges a higher price than its rival, undertakes more advertising, and obtains larger profits. Hence, advertising encourages the prescription of more innovative drugs.

	Advertising		Prices		Market Shares		Profits	
β	a_0	a_1	p_0	p_1	S_0	S_1	Π_0	Π_1
0.8	0.573	0.581	2.92	3.03	0.491	0.509	1.10	<u>1.20</u>
0.9	0.607	0.611	3.21	3.27	0.496	0.504	1.22	<u>1.27</u>
1	0.641	0.641	3.51	3.51	0.5	0.5	1.34	1.34
1.1	0.675	0.671	3.81	3.76	0.504	0.496	<u>1.46</u>	1.41
1.2	0.709	0.701	4.13	4.02	0.507	0.493	<u>1.59</u>	1.49

Table 1: The impact of changes in β when $t_0 = 1$, $u = 4$, $c = 0$ and $k = 0.65$.

We next turn to comparative statics with respect to t. Our interest lies in analyzing the relationship between the degree of product differentiation and the equilibrium level of advertising. We therefore turn to a situation where the relative quality (β) is kept constant while the degree of differentiation is increased.

Proposition 4 *As t increases:*

i) Aggregate advertising increases

ii) The equilibrium becomes more asymmetric, i.e., $a_0 - a_1$ increases for any $\beta > 1$.

iii) If Condition U holds, then, in equilibrium, both firms increase their advertising expenditures and their prices.

	Advertising		Prices		Adv levels		Profits		Profits$_{a_0=a_1=0}$	
t	a_0	a_1	p_0	p_1	$a_0 - a_1$	$a_0 + a_1$	Π_0	Π_1	Π_0	Π_1
0.85	0.573	0.569	2.95	2.91	0.00352	1.4115	1.161	1.120	0.706	0.664
0.9	0.606	0.603	3.22	3.18	0.00358	1.2087	1.258	1.215	0.747	0.704
0.95	0.640	0.636	3.51	3.46	0.00367	1.2758	1.358	1.312	0.789	0.742
1	0.673	0.670	3.80	3.75	0.00375	1.3430	1.461	1.413	0.83	0.781
1.05	0.707	0.703	4.10	4.05	0.00383	1.4101	1.567	1.517	0.872	0.82

Table 2: The impact of changes in t when $\beta = 1.0952$, $u = 4$, $c = 0$ and $k = 0.65$.

Proof. See Appendix B. ■

Proposition 4 provides a number of additional insights. First, it establishes that there is a positive relationship between equilibrium levels of advertising and product differentiation. As t increases, side effects become more important making prices "less effective in capturing consumers", and, by the same token, enhances advertising's attractiveness (part i). At the margin, since price is higher, it is worth spending more on a. The equilibrium becomes more asymmetric as the firm with the lower a_i has the smallest loyal group and prices more aggressively substituting loyal with non-loyal customers. The best reply of its rival is then to invest in the loyal segment. Note that firm i's elasticity of demand decreases with a_i so that, at the asymmetric equilibrium, the firm advertising more intensively faces a more inelastic demand than its rival, thus generating an equilibrium in which firms compete aggressively for different segments of the market, (part ii). Both higher transport cost and increased advertising increases the proportion of customers over which firms can exercise market power resulting in higher equilibrium prices.[20]

Table 2 provides a numerical resolution of the game when t increases for the advertising cost function $C(a_i) = a_i^2$. We have set β constant and equal to 1.0952 so that $t_0 = t$ and $t_1 = \beta t$. Our choice of β makes it possible to compare differences in advertising as a response to quality differentials when T is kept constant. Note that for $t = 1.05$ total transport costs (T) equal 2.2 as in the case $t = 1$ and $\beta = 1.2$ reported in last row of Table 2. In both cases, total advertising equals 1.41. However, the difference in advertising, prices and profits are larger in the former ($t_0 = 1$, $t_1 = 1.2$), where quality differentials are larger, than in the latter ($t_0 = 1.05$ and $t_1 = 1.15$). Another interesting result illustrated in the Table regards the profitability with and without advertising (cf. columns 5 and 6 in Table 2). Firms undertake advertising as this allows them to obtain higher profits. However, these larger profits do not necessarily translate into an increase in the relative profitability of better drugs. On the contrary, given a quality differential the marginal incentive to research for "better drugs" is reduced. For instance, in the last row, where $t = 1.05$, profits are 0.872 and 0.82 without advertising (column 6) while they respectively stand at 1.567 and 1.517 in the presence of promotional effort (column 5). This implies that $\Pi_0(a_0^*, a_1^*) - \Pi_1(a_0^*, a_1^*) < \Pi_0(0,0) - \Pi_1(0,0)$.

Thus, in the context of our model, the fact that better drugs are more intensively advertised does

[20] The result is consistent with the evidence reported in Rizzo (1999) that detailing lowers price sensitivity for antihypertensive drugs.

not necessarily imply that the ability to undertake promotional spend provides incentives to pursue path breaking R&D.[21] However, in the presence of large sunk costs, the ability to undertake promotional effort leading to higher (absolute) profits may induce firms to undertake risky research projects that might be abandoned in the absence of advertising.

We end this Section by analysing changes in the level of co-payments.

Proposition 5 *As k increases:*

i) The equilibrium becomes less asymmetric, i.e., $a_0 - a_1$ decreases for any $\beta > 1$.

ii) Aggregate advertising decreases

iii) If Condition U holds, then, in equilibrium, both firms decrease their advertising expenditures.

Proof. See Appendix B. ■

Proposition 5 establishes a negative correlation between the level of co-payment and firms advertising efforts. The intuition is clear: as k increases, the marginal profitability of advertising efforts decreases and, as a result, firms do less advertising in equilibrium. Note that loyal demand is more sensitive to changes in k than the non-brand loyal one. For the former group an increase in k will reduce $\tilde{x}_i$, while for the latter, *ceteris paribus*, it will not affect $\hat{x}_i$. Since getting a loyal base is more costly for higher values of k, firms will reduce their advertising efforts.

Summing up, the advertising game exhibits increasing differences with respect to both β and t. An increase in either β or t shifts both best replies outwards: keeping a_j constant, firm i finds it optimal to increase its advertising effort as a response to a larger β or t. Thus, the marginal profit increases as either of them increases. By contrast, the advertising game exhibit decreasing differences with respect to k: an increase in k shifts both best replies inwards. As a consequence, aggregate advertising expenditures are larger the larger is β, the larger is t and the lower is k.

6 Empirical analysis

6.1 Data Description

Our model generates several hypothesis concerning prices and advertising strategies that can be tested. After describing the data we have gathered, this section identifies the predictions that are amenable to be empirically tested.

The data we use are from four different sources. Quarterly observations on value and volume sales of all prescription drugs sold in USA during the period 1994q1-2003q4 are retrieved from IMS Health. Product level data on detailing (promotion to physicians) have also been obtained from IMS while the corresponding data on DTCA comes from TNS Media. To avoid confusion, we will refer to the two IMS datasets as IMS-Sale and IMS-Promo. Last, qualitative information about the quality of the drugs have been collected from the Orange Book published by the US Food and Drug Administration (FDA).

[21] This finding is reminiscent to those reported by Ganuza, Llobet and Dominguez (2009), who note that the lack of price-sensitivity of the demand due to marketing effort (or universal public health coverage in EU countries and health insurance in USA) provides an excessive reward for less innovative drugs. Because of this, pharmaceutical firms find relatively more profitable to invest in what they call "small innovations".

IMS-Sale reports value and volume sales for all the different dosages and forms of administration (e.g. 20mg tablets, 50ml vials, etc.), broken down by distribution channel (hospitals and pharmacies). The dataset also reports the therapeutic class of the drug, which is used to define market boundaries (see below for details). In addition, IMS-Sale provides the posology of the active ingredient in each package in a standardised format. Thus, for each package g of drug i, we have total revenues ($R_{g,i,t}$) as well as the quantity of the active ingredient measured in standard units ($SU_{g,i,t}$). This allows us to compute the price of the $g^{th}-$package of product i as $P_{g,i,t} = R^d_{g,i,t}/SU_{g,i,t}$, where the upper script d indicates that revenues are deflated with the Producer Price Index for the pharmaceutical industry available from the US Bureau of Labour Statistics. The (weighted) average price of drug i is then computed using the following formula:

$$\overline{P}_{i,t} = \sum_{g=1}^{N} P_{g,i,t} * \frac{SU_{g,i,t}}{\sum_{g=1}^{N} SU_{g,i,t}}$$

Note that for each drug i or package g we can compute three different prices: one for hospitals (constructed using information on revenues and quantities sold in that channel), one for pharmacies, and an aggregate one.[22]

Drugs in IMS-Sale are classified according to the Anatomical Therapeutic Chemical (ATC) classification system. Following the usual approach in the literature, we define markets at the ATC3 level, which groups drugs that are therapeutic substitutes.

Quarterly data on detailing are retrieved from IMS-Promo. These include the cost of promoting drugs to office-based physicians, hospitals and pharmacists as well as the retail value of product sampling. The cost of detailing at product level is computed by gathering information on the number of representatives' visits and the time spent to discuss a particular drug during these visits. Nominal figures have been deflated with the average of the Producer Price Index for pharmaceutical products and the Producer Price Index for the media industry.

Quarterly data on DTCA expenditure at drug level for the period 1995q1-2004q4 are obtained from TNS Media. Nominal values are deflated using the Producer Price Index for broadcasting and media. TNS monitors advertising efforts for different media, including TV, national and regional newspapers and specialized journals. Traditionally, pharmaceutical firms have promoted their drugs through detailing, often handing-out free samples in the process. DTCA has became more and more important since a change in advertising regulations introduced by the FDA in late 1996.

Finally, the Orange Book published by the FDA provides information on the approval process and the chemical structure of brand drugs that can be used to proxy drug quality. Each drug marketed in the US must go through a detailed FDA review process. The FDA uses a two-tiered review system that assigns a (faster) "priority review" status to drugs that are anticipated to offer major improvements over existing treatments, and a "standard review" status to drugs that offer only minor advances (if any) in treatment.[23] Following the notation used in our theoretical model, "standard review" drugs are assumed to have a transport costs higher than "priority review" drugs (i.e. $\beta > 1$). After a innovative drug with "priority" status has been introduced into the market, follow-on drugs with similar efficacy are less likely

[22] To the best of our knowledge, this is the first study that analyse the price of hospitals and pharmacies separately. Previous studies have used an average price for both channels.

[23] Iizuka (2004) also uses this information to define the quality of a branded drug.

to receive a "priority" status despite being as good as the pioneer drug. The empirical analysis will therefore take into consideration this issue by stipulating that a "standard" drug that enters the market before a "priority" drug is more likely to be of lower quality (i.e. β is higher). By contrast, a "standard" drug launched after a "priority" does not necessarily indicate lower quality.

The orange book also provides information on whether the drug consists of a new chemical entity NCE (i.e. a new active ingredient that has never been marketed before), a new formulation or a new combination of existing molecules. For the purpose of this study, all drugs that are new combination or formulation (NCF) are considered as one group, so that we will have two chemical categories: NCE and NCF. Drugs in the latter group are generally considered improvements on drugs in the former group, either because they combine two active ingredients in a more effective way or because they refer to new formulations that are better tolerated by patients.[24] This distinction allows us to classify the quality of the drugs along a second dimension: for any pair of competing drugs with "priority" and "standard" status, we anticipate that existing asymmetries are larger when the "standard" drugs is a NCE while they are likely to be lower when the "standard" drug is a NCF**.** This reflects the conjecture that even if it is a"standard review" drug, the fact that it is a NCF is indicative of better quality.

The discussion above leads us to classify the drugs along different dimensions, which are summarized in Table 3. First, pairs of drugs are divided according to the review status. The dummy variable *pp* refers to pairs of drugs with priority status while *ss* refers to two drugs which obtained a "standard" review. The most interesting predictions of our model refer to markets where drugs have different qualities or side effects. Accordingly, the empirical analysis focuses on the pairs of drugs characterized by asymmetric qualities, which are defined by the dummy variable *ps* (*i.e.*drug located at point 0 has a "priority" status and the other has a "standard" status). Follow-on drugs with similar efficacy than pioneer drugs are less likely to receive a "priority" status. It is then important to distinguish the case where the "standard" has been launched on the market before the "priority" drug (indicated with the dummy ps^{before}) from the case where the "standard" drug has entered the market after the "priority" drug (dummy ps^{after}). Finally, pairs of drugs where the "standard" drug is a NCE are indicated with the dummy ps^{NCE} while pairs of drugs where the "standard" drug is a NCF are referred as ps^{NCF}. Ideally, we would want to use a continuous variable that could measure or rank the efficacy of various treatments. Given that the medical profession (let alone the industry) does not provide such a continuous ranking, an attempt to construct it would surely be elusive. From that perspective, the use of different dichotomic groups represents a reasonable approximation of quality differentials**.**

INSERT TABLE 3 HERE

Although IMS-Sale provides information on prices and quantities for (almost) the entire universe of prescription drugs, the sample that is used in the empirical analysis only includes those therapeutic areas

[24] For instance, PrandiMet is a drug for type 2 diabetes that combines two older active ingredients: repaglinide and metformin. The two components of PrandiMet help lower blood glucose levels in different ways and scientific studies have shown that this new combination drug is more effective than either molecule alone. Similar findings have been found for treatment of hypertension, where drugs combining ACE inhibitors with diuretics (e.g. Accuretic or Zesturetic) are generally more effective than the corresponding drugs with ACE inhibitors alone (e.g. Accupro or Zestril). Another example consists of drugs targeted at alleviating Chronic Obstructive Pulmonary Disease (COPD), a pathology primarily affecting smokers. Pipeline products for the treatment of COPD primarily consist of combinations of existing molecules.

where at least one of the prescription drugs has positive DTCA. Note that because of the high cost associated with pruchasing IMS-Promo, the number of therapeutic areas for which we have access to detailing expenditure is a subset of the number of markets for which we observe DTCA. This explains the differences in the number of observations available to estimate the empirical specifications either with DTCA or with detailing**.** Moreover, for any therapeutic area, the sample is restricted to those medicines that are successfully matched to the Orange Book. As our model refers to competition among branded drugs, we only consider spells of data where a drug has not faced generic entry. More precisely, we use observations for any branded drug until four quarters before the entry of generic substitutes.[25] Finally, given that DTCA by pharmaceutical company was residual prior to 1997, our sample consists of observations from that year onwards.[26]

6.2 Predictions

Having described the data at hand, we can now define the empirical predictions that are tested with these data. Note that in this section, the term "advertising" will be used to refer to both DTCA and detailing to physicians as both measures will be used to test the theoretical predictions. Recall that part (i) proposition 3 implies that differences in quality leads to asymmetric efforts in advertising, i.e. $(a_0 - a_1) > 0$ with $a = \{DTCA, detailing\}$. Thus,

Prediction 1: *The difference between advertising expenditure in the "priority vs standard" group (ps) are positive while this difference is not significantly different from zero for the "pp" and "ss" groups. Moreover, we expect the difference in advertising expenditure to be higher for the group* ps^{before} *than the group* ps^{after}*. Similarly, we anticipate that the difference in quality and in turn, in advertising efforts is higher for the group* ps^{NCE} *than the group* ps^{NCF}

Next, we estimate a price equation to test proposition (i) of Lemma 1 that states that the price of drug i increases with own and competitors' advertising.

Prediction 2: *The coefficients of own advertising* a_i *and competitor advertising* a_j *in a price equation are positive.*

Doctors working in hospital are exposed to a lower amount of detailing than general practitioners.[27] At the same time, there is well established evidence suggesting that DTCA affects the number of drugs prescribed by family doctors and dispensed by pharmacists (see Mayeroefer and Zuvekas, 2008), but that it does not have such pervasive effect for drugs dispensed in hospitals. Similarly, while the level of co-payment k can affect the decision of what drugs a doctor prescribes (or a patient buys), the cost of drugs administered in hospital forms part of bundle (i.e., even if patients receive a hospital bill that details drug costs, they do not individually chose or pay the drugs used during their entire treatment)**.** This suggests that the dynamics of prices in these two channels might be different. In particular, equation (4) predicts that:

[25] Excluding the last four quarters of data prior to generic entry is motivated by the fact that brand producers typically alter their advertising (and sometimes, pricing) strategies shortly before generic entry.

[26] Recall that the legislation allowing "plain vanilla" advertising was introduced in August 1996.

[27] Pharma promotional expenditures in the USA amounted to $21 billion dollar in 2002. Of this figure, 25% was devoted to physicians detailing and 4% to hospital detailing (see Barfett et al (2004).

Prediction 3: *Firms charge a higher price to pharmacies than hospitals, given that the latter pay the full price of drugs ($k = 1$) and doctors in hospitals are less exposed to firms' marketing (lower value of a).*

In line with the theoretical model, testing the first two predictions above requires to construct a sample of competing drug pairs. To this end, we define a markets as ATC3 classes and match each drug in an ATC3 group with all the other drugs in that ATC3 category. Thus, for a market with n drugs, we construct $n * (n-1)/2$ pairs. Note that a market with only one drug will not enter in our sample. We believe that our approach is justified both on theoretical grounds and for practical reasons. First, this approach can be thought of an extension of our model to a $(n-1)$-dimensional space where the n drugs are located at one unit distance from one another. This is compatible with the idea that doctors know or are used to prescribe only a small fraction of the n drugs available for treating a disease. Therefore, for any therapeutic area, our pairs of drugs identify submarkets populated by consumers with limited information about the available treatments. Second, the positive relationship between differences in qualities and differences in advertising efforts (Prediction 1) and the fact that advertising soften price competition (Prediction 2) do not depend on the number of drugs in the market. For instance, even in a market with n products, it still holds that firms with higher quality products will choose higher level of advertising.[28] Third, creating pairs of drugs leads to an increase in the number of observations that permits us to include a number of control variables, such as market dummies, time dummies and firm dummies. These variables control for variations in prices and advertising efforts due to unobserved heterogeneity between pharmaceutical companies or across therapeutic areas. Finally, the fact that residuals are likely to be correlated within markets is controlled for, as our calculation of standard errors have been obtained on the basis of clustering at the ATC3 level (Moulton (1990)).

Table 4 provides summary statistics. We have 360 drugs, of which 129 have a "priority" status and 285 are NCEs. Products are classified in 87 different markets. The number of drugs in each market varies from 2 to 7, with an average number of about 4 drugs. It is interesting to note that, for the whole sample, average prices weighted by sales are slightly higher in hospital as compared to pharmacies. At first sight this seems to contradict Prediction 3, but an analysis of the data indicates that this is largely due to the fact that hospitals tend to use different posologies of the same drug and more expensive formulations (e.g. injections versus pills). Our empirical analysis controls for this heterogeneity by comparing prices at the level of packages (i.e. drugs with the same posologies and formulations/galenic form) that are used in both channels. The figures at the bottom of Table 4 show that the average expenditure in detailing is ten times higher than in DTCA. In interpreting these figures, it must be recalled that statistics for detailing refer to a smaller number of observations. More precisely, we have the relevant data only for thirty-three of the original eighty-seven therapeutic markets, and these markets include most of the the best selling drugs.

INSERT TABLE 4 HERE

[28] By contrast, our results concerning aggregate advertising effort in a given therapeutic area where drugs have different qualities are not easily generalized to the case of n products. For the intuition, see the discussion of the numerical simulations reported in Table 3.

6.3 Empirical Equations and Results

The empirical specifications used to test Prediction 1 above are the following:

$$(a_{i,t} - a_{j,t}) = \beta_0 + \beta_1 ps + \beta_2 pp + \beta_3 X + u_{i,t} \tag{E1}$$

$$(a_{i,t} - a_{j,t}) = \gamma_0 + \gamma_1 ps^{after} + \gamma_2 ps^{before} + \gamma_3 X + u_{i,t} \tag{E2}$$

$$(a_{i,t} - a_{j,t}) = \delta_0 + \delta_1 ps^{NCE} + \delta_2 ps^{NCF} + \delta_3 X + u_{i,t} \tag{E3}$$

The dependent variable is the difference of the logarithm of advertising expenditures (whether in the form of DTCA or detailing). The logarithmic transformation reduces the effects of outliers in the right tail of the advertising distribution. Given that the logarithm of a non positive number does not exist, our dependent variables are computed as $ln(a_0 + 1) - ln(a_1 + 1)$. This approach has been extensively used in the empirical literature on R&D where many firms reports zero research expenditures (see Klette, 1999).[29] Prices, advertising effort and dichotomous variables identifying quality differences have already been defined in Section 6.1.

The vector X refers to a set of control variables. More specifically, we include a dummy for each ATC3 that controls for time-invariant differences in advertising efforts across markets, firm dummies to control for heterogenous advertising efforts derived from firm level fixed effects (e.g. size), and a set of time dummies that control for time specific shocks that affect advertising efforts (e.g. business cycle effect). Although we only consider the period of time where drugs are under patent protection, advertising efforts could change during the life of a branded drug. Therefore, the specifications include the difference in the number of years that the two drugs have been on the market (AgeDiff) to control for the existence of product life cycle management effects. Finally, the control vector X includes a quarterly count variable that stands for the number of competing drugs in a given therapeutic area. In order to avoid the "dummy variable trap", we use ss (or ss and pp) as base group. Thus, the constant term refers to the omitted quarter, therapeutic area (ATC3 class) and quality pairs (ss).

To estimate equations above we use panel data methods and employ a random effect (RE) estimator.[30] A fixed effect (FE) estimator cannot be applied since the variable of interests do not change over time. Moreover, we believe that the usual concern about unobserved heterogeneity possibly correlated with the right-hand side variables is less important given that our dummy variables refers to primitives (relative quality of the drugs and novelties of the chemical entities) that are determined before price and advertising competition takes place. From an empirical perspective, our main concern is that our dummy variables might be weak proxies of the relative quality of the drugs or the implicit transport costs accounting for side effects. This translates into a problem of measurement errors in the explanatory variables that might lead to estimates that are downward biased (see Wooldridge, pag.75). Thus, if there exists endogeneity problems, they work against our predictions.

[29] One interpretation of these zeros is that they reflect a censoring problem, and that all firms are doing some advertising but not in the form of formal DTCA or promotion.

[30] A similar approach is used by Regan (2008) to estimate the effect of generic entry on the prices of brand products.

Results for the specifications above are reported in Table 5. Columns (1A)-(5A) refer to specifications with DTCA as dependent variable while columns (1B)-(5B) use detailing to physicians. The coefficient of the difference in age is negative in all the specifications (although it is precisely estimated only when DTCA is the dependent variable), thus confirming the well-known fact that newer drugs are more heavily advertised than older drugs. Columns (1A) and (1B) show that the coefficient on ps is positively and significant while the coefficient of pp is also positive but not significantly different from zero (recall that the base group here is the pairs of drugs with "standard" reviews, ss). In line with Prediction 1, these results confirm that asymmetries in the quality of the drugs lead to parallel asymmetries in advertising efforts. Column (2A) and (2B) re-estimate the same model using both pp and ss pairs as control group. Our earlier finding of significant asymmetries for the ps group are confirmed. Prediction 1 also states that the differences in advertising might be higher when the "standard" drugs enters the market before the "priority" drug. Results in column (3A) and (3B) confirm this prediction: the point estimate of the coefficient of ps^{before} is much higher than the coefficient of ps^{after}. A Wald test of the null hypothesis that the two coefficients are equal is rejected at 5% significance level. Columns (4A) and (4B) are consistent zith the prediction that the difference in advertising is higher when the standard drug is a NCE than when it is a NCF, although the coefficient of ps^{NCF} is not precisely estimated. This may due to the fact that there are substantially fewer drugs in the NCF group as compared to the NCE one (see Table 4). Finally, column (5A) and (5B) provide further evidence that differences in advertising expenditures are an increasing function of differences in quality. The coefficient on the interacted dummies $ps^{before} * ps^{NCE}$, which we expect to indicate the pairs of drugs with the largest gap in quality (recall that ps^{NCE} stands for a pair where the the "standard" drug is an NCE), has the highest point estimates among all the dichotomic regressors. The estimates reported below are robust to the exclusion of cases where the pairs of competing products belong to the same company.[31]

INSERT TABLE 5 HERE

We next turn to the analysis of pricing strategies. On the basis of Prediction 2, we estimate a price equation for firm i using own and competitors' advertising (in the form of DTCA or detailing) on the r.h.s. Accordingly, we estimate:

$$p_{i,t} = \beta_1 p_{i,t-1} + \beta_2 p_{j,t} + \beta_3 A_{i,t} + \beta_4 A_{j,t} + \beta_5 X + u_{i,t} \tag{E4}$$

where the subscript j refers to one of the competing drugs that has been matched to drug i. The first term in specification (E4) is firm i's lagged price, which control for dynamic effects in the pricing strategies as well as possible autocorrelation of the errors. Following the first order condition of the price stage, the specification also includes the price of competitor j. Given that price setting can be influenced

[31] The number of these cases is relatively small given that firms do not launch a second generation or entirely new product while they still have a patent protected product on the market. As long as a molecule enjoys exclusivity, large rents can be extracted whenever sales are non-residual. Thus, launching a new drug would cannibalise existing rents. Occurrences of contemporaneous presence of two drugs are primarily found when an originator company attempts to "migrate" its patient base to a new own-drug prior to loss of exclusivity of the old molecule. In general, this can be achieved over a relatively short period of time. The fact that we exclude four quarters prior to generic entry form our sample de facto results in having very few occurrences of two molecules of the same firm on a given market.

not only by present advertising by also by past expenditures, the specification includes the accumulated stock of past advertising efforts A, computed as follows:

$$A_{i,t} = \delta * A_{i,t-1} + a_{i,t} \tag{6}$$

where δ reflects the real depreciation of the accumulated stock over time and $a_{i,t}$ is the flow of (DTCA or detailing) expenditure in the current period.[32] Given that we use the logarithm of A, our estimating equation would collapse for those firms reporting zero advertising. As for the advertising regression above, we apply the standard fix for this problem developed in the R&D literature and define the cumulative advertising as $ln(A_{i,t} + 1)$.

It is possible that firms tend to advertise more and charge higher prices in more concentrated markets. In order to control for the possible positive association between advertising and prices due to the particular structure of a therapeucit area, the control variables X include the Herfindahl index computed using the information in IMS-Sale.[33] As for the advertising equations above, the specification also includes the number of molecule in each market.

As before, we use panel method but in this instance we can employ FE models because the variables of interests change over time. The FE estimation controls for any time-invariant heterogeneity across drugs such as quality and side effects and unobserved differences across firms in their pricing strategies. Accordingly, the vector of control variables X include a set time dummies but not market dummies. With fixed effect, identification of the parameters arises from relating changes in DTCA and detailing to changes in prices for each drug.

There are at least two potential concerns in estimating equation (E4). First, the fact that both DTCA and detailing are measured with errors might lead to an attenuation bias in the corresponding coefficients. Second, although our theoretical model assumes that advertising and prices are chosen sequentially, we cannot rule out instances where the pricing and advertising decisions are simultaneously affected by some aggregate shocks.[34] To deal with these endogeneity problems, we exploit the panel structure of our data and use lags of sales from $t-2$ to $t-3$ as instruments.[35] We have also experimented with a variety of other instruments, such as lags of advertising and the age of the drugs; this additional set of estimations (available upon request) yielded similar results.

Table 6 reports the estimated coefficients of specification (E4). All the coefficients take the expected sign and the most relevant to our research are also statistically significant. Tests of overidentification and weak identification suggest that the instruments are valid (i.e. not correlated with the error term) and relevant (i.e. correlated with the endogenous variable). The high value of the R-squared is due to the inclusion of the lagged depended variable on the right hand side. The coefficient of the Herfindahl index is always positive but it is precisely estimated only in the first specification, when we use a higher

[32] We use a value for δ of 0.7 as in Rizzo (1999). Results are robust to changes in the rate of depreciation. Cumulative advertising is useful to solve the problem of spikes in DTCA over different quarters (i.e. quarters with high DTCA followed by quarters with no DTCA at all). Similarly, Dave and Saffer (2010) use a specification where DTCA is computed as current advertising and a decay-weighted sum of past advertising.

[33] We thank an anonimous referee for drawing our attention to this point.

[34] For instance, the outbreak of a new breed of influenza (e.g. H1N1) might determine a simultaneous increase in prices and DTCA, the later aimed at increasing people's awareness of the problem.

[35] The high correlation between sales and advertising expenditure guarantees that the former are relevant instruments for the latter. The advantage of using sales is that this variable is more precisely measured in our dataset.

number of observations.

Column (1) indicates that the coefficients of own and rival DTCA are positive and statistically significant. Similar results are obtained when we estimate the model for the sub-sample of therapeutic areas for which we observe promotion to physicians. Column (2) shows that an increase in own or rival detailing has a positive effect on prices**.** Finally we estimate a model with both own and rival DTCA and own and rival detailing. All the coefficients of the advertising variables are positive and three out of four are significant at 10% level or more. The results in Table 6 are qualitatively similar to the findings reported in Table 5 of Dave and Saffer (2010). This is rather reassuring given that they use a different sample and a different specification (semi-log model). Overall, these results are consistent with our hypothesis that the price of a drug depends positively on both its own advertising and its rivals' advertising (Prediction 2).

INSERT TABLE 6 HERE

The empirical evidence presented above is based on comparing advertising and prices of pairs of competing drugs. However, it should be noted that convergence to a steady equilibrium price as the one studied in our model can take some time. The study by Lu and Comanor (1998) indicates that prices of new pharmaceutical products are strategically adjusted in the first few years after being launched. In order to test Prediction *3* (*i.e* prices charged to pharmacists are higher than those observed in hospitals), we can compare the prices of newly marketed drugs across the two distribution channels: hospitals and pharmacies.

Comparing the prices in these two channels cannot be done using the average price of a drug, given that hospitals tend to use different posologies and formulations/galenic forms as compared to those sold in pharmacies. For this reason, we compare the price at package (or drug/galenic form) level. In this way, we can attempt to isolate the effect of advertising and co-payments, under the hypothesis that relevant variables other than promotional effort ought have the same effects on the prices set in each these two channels. We therefore selected all drug/galenic-form that are available in hospitals and pharmacies for which we can observe the initial price (at quarter 1) and the price five years after (at quarter 20), yielding a sample of 382 packages.[36] Table 7 shows that the average price for pharmacies is higher than the average price for hospitals and that this difference is increasing over time. Five years after entry, a t-test of the means' equality rejects the null hypothesis that the average price is the same in the two channels.

INSERT TABLE 7 HERE

We further investigate differences between the two channels by looking at the distribution of the growth of prices of the 382 drug-form. The change in price of the $g-th$ package is computed as $\Delta p_g = \ln(p_{g,20}) - \ln(p_{g,1})$, where 1 is the quarter of entry into the market and 20 is the price of the drug twenty quarters after entry. Using growth instead of levels allows us to control for time-invariant

[36] Five years seems a reasonable compromise between having a sufficiently long time window for a firm to adjust its price towards an equilibrium value and having a sufficient large number of observations. Our results are robust to expanding/shrinking the period considered.

unobserved factors that can affect the (level of) prices in the two channels (e.g. differences in cost of distribution or in bargaining power). Figure 2 shows that the empirical c.d.f of Δp for hospitals (black solid line) consistently lies to the left of the c.d.f for pharmacies (grey dotted line), thus implying that there is a higher increase in the price of drugs sold to pharmacies. A formal test of differences in the two pairs of empirical distributions based on the Kolgomorov-Smirnov statistics rejects the null hypothesis of no differences between the two distributions (p-value < 0.01). Overall, these results are consistent with Prediction 3 above and, more generally, with the results of our theoretical model.

INSERT FIGURE 2 HERE

7 Conclusion

This paper develops a theoretical model tailored to the competitive interaction between originator drugs that are still under patent protection. The price competition stage is preceded by strategic decisions on promotional effort in the form of persuasive advertising. The latter endogenously generates two consumer groups: those that are brand loyal and those that are not. While both segments respond to price changes, non-loyal doctor/patient pairs are more sensitive to it. We fully characterize equilibria under parameter constellations that have been chosen to reflect real world conditions. In addition, we allow for asymmetries in terms of drug quality.

The model allows us to make three empirical predictions regarding price and advertising strategies. These predictions are taken to the data, the latter consisting of product level information on price and advertising to consumers and physicians. Quality differences are mapped into the data using a novel taxonomy. We find empirical support for the model's central predictions.

While our results indicate that, for a given quality differential, the better quality drugs are also the ones that are most advertised, this does not warrant the conclusion that the possibility to undertake promotional effort induces higher R&D effort. Indeed, as we do not model the R&D stage, it is not possible to derive conclusions in that respect. If anything, our results suggest that, for a given quality differential, the marginal profit of having developed the superior drug is higher when promotional effort is nil.

Two extensions naturally come to mind. The first would consist in allowing for generic entry. Preliminary work in that direction (De Frutos, Ornaghi, Siotis (2010)) suggest that our model is well suited to reproduce many of the surprising findings that are observed in the data (findings that build on Scherer's famous "generic entry paradox" regarding price). The second line of research would tailor our model's basic architecture to the specificities of Europe's health systems, where a maximum price is negotiated between firms and public authorities.

8 Appendix A

Firms' Best Reply

To provide a complete characterization of firm i best reply we start by noting that whenever its rival is very aggressive and sets low prices, brand i may be better off by attending only to its loyal consumers

at $kp_i = u - t_i$. To capture some of the non-loyal consumers it must set a low price such that the additional quantity (extra consumers) does not compensate the price reduction. As the rival raises its price, it becomes increasingly profitable to attend both segments. In particular, there is $\underline{p}_j$ such that both strategies are equally profitable for firm i. The price $\underline{p}_j$ is determined by the following equality:

$$(\tilde{p}_i - c)\left(\theta_i + (1 - \theta_i - \theta_j)\left(\frac{k\left(\underline{p}_j - \tilde{p}_i\right) + t_j}{t_0 + t_1}\right)\right) = \left(\frac{u - t_i}{k} - c\right)\theta_i \tag{7}$$

where $\tilde{p}_i$ stands for the best reply by firm i when its rival sets $\underline{p}_j$ and i attends both segments. Consequently, $kp_i = u - t_i$ is the best reply to any $p_j \leq \underline{p}_j$. Notice that $\underline{p}_j = 0$ if a_i is sufficiently small. Trivially, if firm i undertakes little advertising, its captive mass of loyal consumers will be low; hence it will prefer competing for both segments.

For $p_j > \underline{p}_j$, serving both segments becomes the best firm i can achieve. Under this condition, total market shares and profits are given by:

$$\begin{aligned} S_i &= \theta_i + (1 - \theta_0 - \theta_1)\left(\frac{k(p_j - p_i) + t_j}{t_0 + t_1}\right),\ kp_j - t_i < kp_i \leq \min\{kp_j + t_j, u - t_i\} \text{ and} \\ \Pi_i &= (p_i - c)\frac{1}{1 + a_0 + a_1}\left(a_i + \left(\frac{k(p_j - p_i) + t_j}{t_0 + t_1}\right)\right). \end{aligned}$$

The first-order condition for profit maximization with respect to p_i yields:

$$\tilde{p}_i = \frac{p_j + c}{2} + \frac{(t_j + a_i T)}{2k}, \tag{8}$$

where T denotes total transport cost $t_0 + t_1$. By substituting $\tilde{p}_i$ into (7), it is possible to derive the value of $\underline{p}_j$:

$$\begin{aligned} k\underline{p}_j &= \left(2\sqrt{a_i T\left(v - t_i\right)} + ck - t_j - Ta_i\right), \text{ and} \\ k\tilde{p}_i(\underline{p}_j) &= \sqrt{a_i T\left(v - t_i\right)} + ck. \end{aligned}$$

Since $\tilde{p}_i$ is a function increasing in p_j, it reaches $(u - t_i)/k$ when the price set by firm j equals $\bar{p}_j$ where:

$$k\bar{p}_j = 2u - ck - t_i - T\left(1 + a_i\right).$$

For prices above $\bar{p}_j$, firm i will not respond by increasing its own price, as the cost of setting a higher price results in losing some of its loyal customers, which is never profitable as the price $kp_i = u - t_i$ strictly dominates any higher price.

Based on the discussion above, the best reply of brand i is given by:

$$R_i(p_j) = \begin{cases} \frac{u - t_i}{k} & \text{if} \quad 0 \leq p_j \leq \max\left\{0, \underline{p}_j\right\} \\ \frac{p_j + c}{2} + \frac{(t_j + a_i T)}{2k} & \text{if} \quad \max\left\{0, \underline{p}_j\right\} \leq p_j < \bar{p}_j \\ \frac{u - t_i}{k} & \text{if} \quad \bar{p}_j \leq p_j. \end{cases}$$

Note that for a_i, a_j sufficiently high, best replies exhibit a discontinuity. As a consequence, a pure strategies equilibrium may fail to exist. This is the case for some parameter values, as the example below illustrates. Let $a_0 = 0.75$, $a_1 = 0.5$, $t_0 = 1$, $t_1 = 1.2$, $k = 0.65$. Firms' best replies for these parameter

values are shown in *Figure 3.a* and *Figure 3.b* below, for two different values of v, $v = 6$ and $v = 4$ ($c = 0$). In both Figures, the thin line depicts firm 0's best reply, while the thick one exhibits firm 1's. Whereas for $v = 6$, best replies fail to intersect, for $v = 4$, they intersect in the increasing segment.

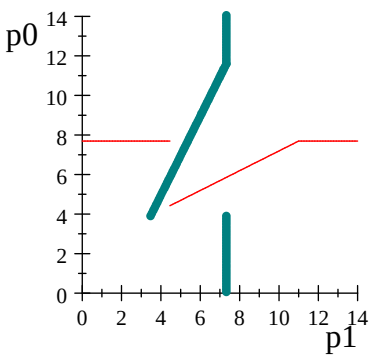

Figure 3.a. Best replies when $v = 6$.

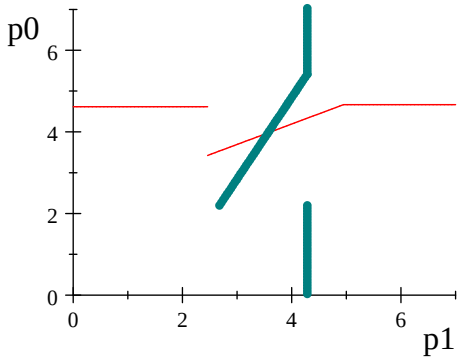

Figure 3.b. Best replies for $v = 4$.

The intuition underpinning the absence of an equilibrium lies in the temptation of setting the maximum price, i.e., $(u - t_i)/k$, even if that implies selling only to the loyal consumers. If firm i does so, then its rival can supply the entire non-loyal segment at a relatively high price. Such a high price makes it attractive for firm i to undercut so as to compete for the non-loyals. This undercutting provokes further undercutting and ends with firm i falling back to the initial price $(u - t_i)/k$, preventing the existence of a pure strategy equilibrium. Note that the non-existence of a pure strategy equilibrium is driven by u being large enough as in *Figure 3.a*. Clearly, this result also depends on a_i being sufficiently large for the loyal base to have enough mass. Nevertheless, one cannot restrict the value of a_i as we are searching for a price equilibrium in every possible subgame, i.e., in every possible continuation induced by (a_0, a_1).

9 Appendix B

Proof of Proposition 1

Proof. *Existence and Uniqueness.* We first show that $v \leq T + 3t_0$ ensures that $\tilde{p}_i\left(\underline{p}_j\right) - \underline{p}_j \geq 0$ holds for $i, j = \{0, 1\}$, $i \neq j$. For low values of advertising effort, the result follows trivially as $\underline{p}_j = \underline{p}_i = 0$. If $\underline{p}_j > 0$, then we have:

$$k\left(\tilde{p}_i\left(\underline{p}_j\right) - \underline{p}_j\right) = t_j + Ta_i - \sqrt{a_i T\left(v - t_i\right)}$$

The right-hand-side difference is a convex function in a_i with a minimum at $a_i = (v - t_i)/4T$. Evaluated at its minimum, the difference becomes $t_j - (v - t_i)/4$ which is non-negative if and only if $v \leq T + 3t_j$.

If $v \geq T + 3t_1$, then $\tilde{p}_1\left(\underline{p}_0\right) < \underline{p}_0$ and $\tilde{p}_0\left(\underline{p}_1\right) < \underline{p}_1$. In the $(x = p_1, y = p_0)-$axes, $\left(\underline{p}_1, \tilde{p}_0\left(\underline{p}_1\right)\right)$ lies below the 45^o line, whereas $\left(\tilde{p}_1\left(\underline{p}_0\right), \underline{p}_0\right)$ lies above it. Consequently, the two best replies never cross (recall counterexample in Appendix A where $\left(\underline{p}_1, \tilde{p}_0\left(\underline{p}_1\right)\right) = (4.4, 4.39)$ and $\left(\tilde{p}_1\left(\underline{p}_0\right), \underline{p}_0\right) = (3.5, 3.8)$).

If $v - T \in (3t_0, 3t_1)$, then $\tilde{p}_1\left(\underline{p}_0\right) < \underline{p}_0$ and $\tilde{p}_0\left(\underline{p}_1\right) > \underline{p}_1$. In the $(p_1, p_0)-$axes, both pairs $\left(\underline{p}_1, \tilde{p}_0\left(\underline{p}_1\right)\right)$

and $\left(\tilde{p}_1\left(\underline{p}_0\right), \underline{p}_0\right)$ lie above the 45^o line. If the latter is above the former then the two best replies never cross. A sufficient condition for $\tilde{p}_1(\underline{p}_0) > \underline{p}_1$ is $t_0 + Ta_1 > t_1 + Ta_0$. To see this, note that:

$$\begin{aligned} k\left(\tilde{p}_1(\underline{p}_0) - \underline{p}_1\right) &= t_1 + Ta_0 - 2\sqrt{a_0 T\left(v - t_0\right)} + \sqrt{a_1 T\left(v - t_1\right)} \\ &> t_0 + Ta_1 - t_1 - Ta_0 > 0 \text{ if } t_0 + Ta_1 > t_1 + Ta_0 \end{aligned}$$

where the first inequality follows from the fact that $\tilde{p}_1\left(\underline{p}_0\right) < \underline{p}_0$ and $\tilde{p}_0\left(\underline{p}_1\right) > \underline{p}_1$ hold.

Assume finally that $v - T \leq 3t_0$ so that the pair $\left(\underline{p}_1, \tilde{p}_0\left(\underline{p}_1\right)\right)$ lies above the 45^o line whereas the pair $\left(\tilde{p}_1\left(\underline{p}_0\right), \underline{p}_0\right)$ lies below. As both firms best replies will cross the 45^o line, they must cross each other so that an equilibrium will exist. Finally note that the crossing is unique and it occurs at a pair in which both firms attend both market segments.

Interiorness. We explore next the nature of this equilibrium. Trivially there are four potential equilibrium configurations depending on wether the best replies intersect at the increasing segment or at the flat one (see Figure 3.b in Appendix A). The crossing point, and hence the associated equilibrium configuration, depends on the location of $\bar{p}_i$ with respect to $p_i^{\max}$ and $p_j^{\max}$, with $\bar{p}_i - p_i^{\max} \geq 0$ iff $v - T \geq Ta_j + (t_j - t_i)$ and $\bar{p}_i - p_j^{\max} \geq 0$ iff $v - T \geq Ta_j$. Assume for instance that $v - T < \min\{Ta_1, Ta_0 - (t_1 - t_0)\}$. Since $\bar{p}_0 < p_1^{\max}$ and $\bar{p}_1 < p_1^{\max}$, firm 0's (firm 1's) best reply lies in its flat stretch when intersecting $p_1^{\max}$ (when intersecting $p_0^{\max}$). Consequently, best replies cross each other at $(p_0^{\max}, p_1^{\max})$. By looking at all the possible cases that can arise, Table 3 maps all potential equilibrium configurations:

$v - T$	$\leq Ta_1$	$\in (Ta_1, Ta_1 + (t_1 - t_0))$	$\geq Ta_1 + (t_1 - t_0)$
$\leq Ta_0 - (t_1 - t_0)$	$(p_0^{\max}, p_1^{\max})$	$(p_0^{\max}, p_1^{\max})$	$(p_0^{\max}, \tilde{p}_1)$
$\in (Ta_0 - (t_1 - t_0), Ta_0)$	$(\tilde{p}_0, p_1^{\max})$	$\{(\tilde{p}_0, p_1^{\max}),\ (\tilde{p}_0, \tilde{p}_1)\}$	$(\tilde{p}_0, \tilde{p}_1)$
$\geq Ta_0$	$(\tilde{p}_0, p_1^{\max})$	$\{(\tilde{p}_0, p_1^{\max}),\ (\tilde{p}_0, \tilde{p}_1)\}$	$(\tilde{p}_0, \tilde{p}_1)$

Table 3: Equilibrium configurations

Note that for $(p_0^{\max}, p_1^{\max})$ to be an equilibrium, $v - T$ must be low enough. Furthermore, if $a_i \in [0, 1]$, then this price pair will never be an equilibrium configuration if $v - T \geq 2t_0$. Similarly, the pair $(p_0^{\max}, \tilde{p}_1)$ will never be an equilibrium configuration if $v - T > 2t_0$. Next, focus on the pair $(\tilde{p}_0, p_1^{\max})$ in the first and second column. Whenever $\bar{p}_0 \in (p_1^{\max}, p_0^{\max})$, second column, the crossing point depends on whether it is above or below $\tilde{p}_0(p_1^{\max})$. Since $k(\tilde{p}_0(p_1^{\max}) - \bar{p}_0) < 0$ whenever $T(2 + a_0 + 2a_1) + 2t_1 < 3v$, in subgames with $k(\tilde{p}_0(p_1^{\max}) - \bar{p}_0) \leq 0$, the unique equilibrium is $(\tilde{p}_0, \tilde{p}_1)$,[37] whereas in those in which $k(\tilde{p}_0(p_1^{\max}) - \bar{p}_0) > 0$, the equilibrium is $(\tilde{p}_0, p_1^{\max})$. Consequently, $(\tilde{p}_0, \tilde{p}_1)$ is the unique equilibrium in every subgame (i.e. firms' best replies intersect at the increasing segment) if $v - T > \frac{4t_1 + 2t_0}{3}$ holds. ■

[37] This corresponds to the case depicted in *Figure 2.b.* where $v = 4$ is larger than $(T(2 + a_0 + 2a_1) + 2t_1)/3$.

Proof of Proposition 2.

Proof. The existence of a Nash equilibrium is immediate from Vives (1990) given that Π_i is an increasing function that satisfies strict submodularity (i.e., $\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j} < 0$), as

$$\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j} = -\frac{2\left(t_j + a_i T\right)\left(t_i + a_j T\right)}{9kT\left(1 + a_0 + a_1\right)^3} < 0.$$

If $t_0 = t_1 = t$, then the advertising game is a strictly symmetric submodular game, and as such it can only have symmetric equilibria. Since the symmetric equilibrium is unique, the result follows.

Assume next that $t_0 < t_1$. Uniqueness requires $-\frac{\partial^2 \Pi_i}{\partial a_i^2} > \left|\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j}\right|$. Note that $\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j} < 0$, $\left|\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j}\right| = -\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j}$, which implies that $-\frac{\partial^2 \Pi_i}{\partial a_i^2} > -\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j}$ or, equivalently, $\frac{\partial^2 \Pi_i}{\partial a_i^2} < \frac{\partial^2 \Pi_i}{\partial a_i \partial a_j}$, where the second-order derivative of Π_i is given by

$$\frac{\partial^2 \Pi_i}{\partial a_i^2} = \frac{2}{9kT}\frac{\left(t_i + a_j T\right)^2}{\left(1 + a_0 + a_1\right)^3} - C''(a_i).$$

Thus, $\frac{\partial^2 \Pi_i}{\partial a_i^2} < \frac{\partial^2 \Pi_i}{\partial a_i \partial a_j}$ holds if

$$C''(a_i) > \frac{2\left(t_j + a_i T\right)\left(t_i + a_j T\right) + 2\left(t_i + a_j T\right)^2}{9kT\left(1 + a_0 + a_1\right)^3} = \frac{2\left(t_i + a_j T\right)}{9k\left(1 + a_0 + a_1\right)^2}$$

Since $2\left(t_i + a_j T\right) < 2T\left(1 + a_j\right)$ and $1 + a_j < 1 + a_0 + a_1$, it follows that $\frac{2(t_i + a_j T)}{9kR^2} < \frac{2T}{9kR}$, $R = 1 + a_0 + a_1$, and consequently that $C''(a_i) > \frac{2T}{9k}$ suffices for uniqueness.

Given that $\frac{2(t_i + a_j T)}{9kR^2} > \frac{2}{9kT}\frac{(t_i + a_j T)^2}{R^3}$, it follows that if *Condition U* holds, then the profit function is strictly concave and an equilibrium must satisfy the first order conditions. ■

Proof of Proposition 3

Proof. i) We first note that the advertising game exhibit increasing differences in β as $\frac{\partial^2 \Pi_i}{\partial a_i \partial \beta} > 0$, meaning that, as β increases, both best replies shift upwards as compared to the case $\beta = 1$. Moreover:

$$\frac{\partial^2 \Pi_0}{\partial a_0 \partial \beta} - \frac{\partial^2 \Pi_1}{\partial a_1 \partial \beta} = \frac{\left(1 + a_0 - a_1\right) t}{9\left(1 + a_0 + a_1\right) k} > 0 \text{ if } 1 + a_0 > a_1 \tag{9}$$

Since the best reply of firm 0 shifts upwards more than the one of its rival it follows that $a_0 > a_1$ if $\beta > 1$.

ii) Since $a_0 + a_1$ has slope -1, as β increases, both best replies increase and so does aggregate advertising.

iii) Total differentiation of the FOC with respect to β yields:

$$\frac{\partial^2 \Pi_i}{\partial^2 a_i} \times \frac{da_i}{d\beta} + \frac{\partial^2 \Pi_i}{\partial a_i \partial a_j} \times \frac{da_j}{d\beta} + \frac{\partial^2 \Pi_i}{\partial a_i \partial \beta} = 0$$

Using Cramer's rule we have:

$$\frac{da_i}{d\beta} = \frac{N_i}{D} = \frac{-\left(\frac{\partial^2 \Pi_i}{\partial a_i \partial \beta}\right)\left(\frac{\partial^2 \Pi_j}{\partial a_j^2}\right) + \left(\frac{\partial^2 \Pi_i}{\partial a_i \partial a_j}\right)\left(\frac{\partial^2 \Pi_j}{\partial a_j \partial \beta}\right)}{\left(\frac{\partial^2 \Pi_0}{\partial a_0^2}\right)\left(\frac{\partial^2 \Pi_1}{\partial a_1^2}\right) - \left(\frac{\partial^2 \Pi_0}{\partial a_0 \partial a_1}\right)\left(\frac{\partial^2 \Pi_1}{\partial a_1 \partial a_0}\right)}.$$

We first note that $D > 0$. The result follows from *Condition U*, as it implies that the slope of the best replies is larger than -1, or, equivalently $\left(\frac{\partial^2 \Pi_i}{\partial a_1 \partial a_0}\right) / \left(\frac{\partial^2 \Pi_i}{\partial a_i^2}\right) < 1$.

Focus next on N_0. Using (9) it follows that $\frac{\partial^2\Pi_0}{\partial a_0\partial\beta} = \frac{\partial^2\Pi_1}{\partial a_1\partial\beta} + \frac{(1+a_0-a_1)t}{9(1+a_0+a_1)k}$, so that

$$N_0 > \frac{\partial^2\Pi_1}{\partial a_1\partial\beta}\left(\frac{\partial^2\Pi_0}{\partial a_0\partial a_1} - \frac{\partial^2\Pi_1}{\partial a_1^2}\right) = \frac{\partial^2\Pi_1}{\partial a_1\partial\beta}\left(\frac{\partial^2\Pi_1}{\partial a_1\partial a_0} - \frac{\partial^2\Pi_1}{\partial a_1^2}\right) > 0$$

where the first inequality follows from $-\left(\frac{(1+a_0-a_1)t}{9(1+a_0+a_1)k}\right)\left(\frac{\partial^2\Pi_1}{\partial a_1^2}\right) > 0$ given that $\frac{\partial^2\Pi_1}{\partial a_1^2} < 0$. The first equality stems from the fact that the second-order cross derivatives are identical, while the last inequality **is derived** from *Condition U*, as it implies $\frac{\partial^2\Pi_1}{\partial a_1\partial a_0} - \frac{\partial^2\Pi_1}{\partial a_1^2} > 0$.

Consider finally N_1. Using (9), and decomposing Π_0 into $R_0 - C(a_0)$ it follows that

$$\begin{aligned} N_1 &= \left(\frac{\partial^2\Pi_1}{\partial a_1\partial\beta}\right)\left(C''(a_0) - 2\frac{t\left(a_1\left(1+\beta\right)+1\right)}{9kR^2}\right) + \left(\frac{\partial^2\Pi_1}{\partial a_0\partial a_1}\right)\left(\frac{\left(1+a_0-a_1\right)t}{9kR}\right), \text{ with} \\ \frac{\partial^2\Pi_1}{\partial a_1\partial\beta} &= \frac{t\left(\beta\left(\beta+2\right)\left(a_0+2a_1+1\right)\left(3a_0+2a_1+3\right)+3a_0\left(a_0+2\right)+8a_0a_1+4\left(a_1+1\right)^2\right)}{9\left(a_0+a_1+1\right)^2\left(\beta+1\right)^2k} \end{aligned}$$

The first term is positive if *Condition U* holds, whereas the second term is negative as the second-cross partial derivative is negative. To establish the sign, we substitute $C''(a_i)$ by $\frac{2t(1+\beta)}{9k}$ (recall that $C''(a_i) > \frac{2t(1+\beta)}{9k}$ by *Condition U*), and obtain that the difference is a strictly positive function, which establishes our result. ■

Proof of Proposition 4

Proof. i) We first show that the advertising game exhibits increases differences in t. To see this note that

$$\begin{aligned} \frac{\partial^2\Pi_0}{\partial a_0\partial t} &= \frac{\left(1+\left(1+\beta\right)\left(2+2a_0+3a_1\right)\right)\left(\beta+\left(\beta+1\right)\left(2a_0+a_1+1\right)\right)}{9k\left(\beta+1\right)\left(a_0+a_1+1\right)^2} > 0 \\ \frac{\partial^2\Pi_1}{\partial a_1\partial t} &= \frac{\left(\beta+\left(1+\beta\right)\left(2+2a_1+3a_0\right)\right)\left(1+\left(\beta+1\right)\left(a_0+2a_1+1\right)\right)}{9k\left(\beta+1\right)\left(a_0+a_1+1\right)^2} > 0 \end{aligned}$$

Moreover,

$$\frac{\partial^2\Pi_0}{\partial a_0\partial t} - \frac{\partial^2\Pi_1}{\partial a_1\partial t} = \frac{\left(a_0-a_1\right)\left(1+\beta\right)-\left(1-\beta\right)}{9k\left(a_0+a_1+1\right)} \gtrless 0 \text{ iff } \left(a_0-a_1\right) \gtrless \frac{1-\beta}{1+\beta} \tag{10}$$

ii) If $\beta = 1$ then $a_0 = a_1$ and $\frac{\partial^2\Pi_0}{\partial a_0\partial t} = \frac{\partial^2\Pi_1}{\partial a_1\partial t}$ so that as t increases, both best replies shift upwards and the new (symmetric) equilibrium entails larger advertising efforts. As $\beta > 1$, then by Proposition 3 part i) we have that $a_0 > a_1$ and, consequently, $\frac{\partial^2\Pi_0}{\partial a_0\partial t} > \frac{\partial^2\Pi_1}{\partial a_1\partial t}$ so that the difference $a_0 - a_1$ **widens**.

iii) Using Cramer's rule, it follows from the proof of Proposition 3 part iii) that the sign of da_0/dt and da_1/dt equals the sign of the numerator N_i^t with:

$$N_i^t = -\left(\frac{\partial^2\Pi_i}{\partial a_i\partial t}\right)\left(\frac{\partial^2\Pi_j}{\partial a_j^2}\right) + \left(\frac{\partial^2\Pi_i}{\partial a_0\partial a_1}\right)\left(\frac{\partial^2\Pi_j}{\partial a_j\partial t}\right).$$

Using (10), we can rewrite N_0^t as

$$N_0^t = \left(\frac{\partial^2\Pi_0}{\partial a_0\partial t}\right)\left(\frac{\partial^2\Pi_0}{\partial a_0\partial a_1} - \frac{\partial^2\Pi_1}{\partial a_1^2}\right) - \left(\frac{\left(a_0-a_1\right)\left(1+\beta\right)-\left(1-\beta\right)}{9k\left(a_0+a_1+1\right)}\right)\left(\frac{\partial^2\Pi_0}{\partial a_0\partial a_1}\right) > 0$$

the first term is positive under *Condition U*, and the second one is positive as $-\partial^2\Pi_0/\partial a_0\partial a_1 > 0$.

Focus next on N_1^t. Using (10) again, we can rewrite N_1^t as:

$$N_1^t = \left(\frac{\partial^2\Pi_1}{\partial a_1 \partial t}\right)\left(\frac{\partial^2\Pi_1}{\partial a_0 \partial a_1} - \frac{\partial^2\Pi_0}{\partial a_0^2}\right) + \frac{(a_0 - a_1)(1+\beta) - (1-\beta)}{9k(a_0 + a_1 + 1)}\left(\frac{\partial^2\Pi_0}{\partial a_0^2}\right)$$
$$= \frac{\partial^2\Pi_0}{\partial a_0 \partial t}\left(C''(a_0) - \frac{2t(a_1 + \beta a_1 + 1)}{9k(a_0 + a_1 + 1)^2}\right)$$
$$-\frac{(a_0 - a_1)(1+\beta) - (1-\beta)}{9k(a_0 + a_1 + 1)}\left(C''(a_0) - \frac{2t(a_1 + \beta a_1 + 1)^2}{9k(1+\beta)(a_0 + a_1 + 1)^3}\right) > 0$$

where the result follows from the fact that $\frac{2t(a_1+\beta a_1+1)}{9kR^2} < \frac{2t(a_1+\beta a_1+1)^2}{9k(1+\beta)R^3}$, $R = a_0 + a_1 + 1$, so that the first term in brackets is larger than the second one, and the fact that $\frac{\partial^2\Pi_0}{\partial a_0 \partial t} > \frac{(a_0-a_1)(1+\beta)-(1-\beta)}{9kR}$. ■

Proof of Proposition 5

Proof. As it follows the same steps as previous proofs, we only report the proposition specific steps.

i) and ii) follow from $\frac{\partial^2\Pi_i}{\partial a_i \partial k} < 0$, and:

$$\frac{\partial^2\Pi_0}{\partial a_0 \partial k} - \frac{\partial^2\Pi_1}{\partial a_1 \partial k} = -\frac{1}{9}t\frac{(a_0 - a_1)(1+\beta) - (1-\beta)}{k^2(a_0 + a_1 + 1)} < 0 \text{ as } \beta > 1 \tag{11}$$

Since $a_0 > a_1$ by Proposition 3, the differences $a_0 - a_1$ shrinks, both best replies decrease, and so does aggregate advertising.

iii) N_0^k is negative as it is the sum of two negative terms, with:

$$N_0^k = \left(\frac{\partial^2\Pi_0}{\partial a_0 \partial k}\right)\left(\frac{\partial^2\Pi_0}{\partial a_0 \partial a_1} - \frac{\partial^2\Pi_1}{\partial a_1^2}\right) + \left(\frac{(a_0 - a_1)(1+\beta) - (1-\beta)}{9k^2(a_0 + a_1 + 1)}\right)\left(\frac{\partial^2\Pi_0}{\partial a_0 \partial a_1}\right) < 0.$$

Regarding N_1^k, note that it is the sum of a positive and a negative term. We compute the difference by exploiting that $C'' > \left(\frac{2t(1+b)}{9k}\right)$, and find the sum to be negative. ■

10 References

Anderson, S.P., de Palma A. and J.-F. Thisse, Discrete choice theory of product differentiation, MIT Press, Cambridge, MA (1992).

Azoulay, P., Do Pharmaceutical Sales Respond to Scientific Evidence?, *Journal of Economics and Management Strategy* 11 (4), (2002), pp. 551-594.

Bagwell, K., The Economic Analysis of Advertising," Handbook of Industrial Organization, vol. 3, (2007). Elsevier.

Bala, R. and P. Bhardwaj, Detailing vs. Direct-to-Consumer Advertising in the Prescription Pharmaceutical Industry, *Management Science* 56, (2010), pp. 148-160

Baye, M. R. and J. Morgan, Brand and price advertising in online Markets, *Management Sciences* 55 (2009), pp. 1139-1151.

Brekke, K. and M. Kuhn, Direct to consumer advertising in pharmaceutical markets, *Journal of Health Economics* 25 (1), (2006), pp. 102-130.

Chioveanu, I., Advertising, brand loyalty and pricing, *Games and Economic Behavior* 64(1), (2008), pp. 68-80.

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

Dave, D. and H. Saffer, The impact of Direct-to-Consumer advertising on pharmaceutical prices and demand, *NBER Working Paper 15696* (2010).

De Frutos, M.A., Ornaghi C. and G. Siotis "How do brand producers react to generic entry?", *mimeo*

von der Fehr, N.-H. and K. Stevin, Persuasive advertising and product differentiation, *Southern Economic Journal* 65, (1998), pp. 113-126.

Frank R.G. and D.S. Salkever, Pricing, patent loss and the market for pharmaceuticals, *Southern Economic Journal* 59 (2), (1992), pp. 165-179.

Ganuza, J.J., Llobet G. and B. Dominguez, R&D in the Pharmaceutical Industry: A World of Small Innovations, *Management Science* 56, (2009), Vol. 55, No. 4, pp. 539-551

Grossman G. and S. Shapiro, Informative advertising with differentiated products, *Review of Economic Studies* 51, (1984), pp. 63-81.

Hollon M., Direct-to-Consumer Advertising: A Hazard Approach to Health Promotion, *Journal of American Medical Association* 293, No.16 (2005), pp.2030-2033

Hotelling, H., Stability in competition, *Economic Journal* 39, (1929), pp. 41-57.

Iizuka, T., What explains the use of Direct-to-Consumer advertising of prescription drugs?, *Journal of Industrial Economics* 52 (3), (2004), pp. 349-379.

Iizuka, T. and G. Jin, Direct-to-Consumer advertising and prescription choice, *Journal of Industrial Economics* 55 (4), (2007), pp. 771-771.

Johnson, J. and D. Myatt, On the Simple Economics of Advertising, Marketing and Product Design, *American Economic Review* 96 (2006), pp. 756-784.

Klette, T.J., Market Power, Scale economies and productivity: Estimates from a panel of establishment data, *The Journal of Industrial Economics* 48, (1999), pp. 451-476.

Konigbauer, I., Advertising and generic market entry, *Journal of Health Economics* 26 (2), (2007), pp. 286-305

Linnosmaa, I. E., Advertising, free-riding, and price differences in the market for Prescription Drugs, *The B.E. Journal of Economic Analysis & Policy* 8 Iss. 2, (2008), Article 8.

Lu, Z. J. and C. S. Comanor, Strategic Pricing of New Pharmaceuticals. *The Review of Economic and Statistics,* 80 (1998), pp. 108-118.

Mayerhoefer, C. D. and S. H. Zuvekas, The Shape of Demand; What Does it Tell Us about Direct-to-Consumer Marketing of Antidepressant, *The B.E. Journal of Economic Analysis and Policy,* 8 Iss. 2, (2008), article 4.

Moulton, B., 'An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit', *Review of Economics and Statistics,* 72 (1990), pp. 334-38.

Regan, T. L., Generic Entry, Price Competition, and Market Segmentation in the Prescription Drug Market, *International Journal of Industrial Organization* 26, (2008), pp.930-948.

Rizzo, J.A., Advertising and competition in the ethical pharmaceutical industry: The case of antihypertensive drugs, *Journal of Law and Economics* 42, (1999), pp. 89-116.

Scott Morton, F.M., Barriers to entry, brand advertising, and generic entry in the US pharmaceutical industry. *Rand Journal of Economics* 30, (2000), pp. 421-440.

Tullock, G., The welfare cost of tariffs, monopolies, and theft, *Western Economic Journal* 5, (1967), pp. 224-232.

Vives, X., Nash Equilibrium with strategic complementarities, *Journal of Mathematical Economics* 19, (1990), pp. 305-321.

Vives, X., Complementarities and Games: New Developments, *Journal of Economic Literature* 43, (2006), pp. 437-479.

Wooldridge, J. M., Econometric Analysis of Cross Section and Panel Data, *The MIT Press,* (2001)

Wosinska, M., Just what the patient ordered? Direct-to-Consumer advertising and the demand for pharmaceutical products (October 2002). *HBS Marketing Research Paper No. 02-04*. Available at SSRN: http://ssrn.com/abstract=347005.

Table 3 – Dummy Variables

Review Status of the Two drugs	Time of entry [a]	Chemical Type [b]	Name of Dummies
Priority (P) vs Standard (S)			*ps*
	Standard enters before Priority		ps^{before}
	Standard enters after Priority		ps^{after}
		Standard is a NCE	ps^{NCE}
		Standard is a NCF	ps^{NCF}
Standard (S) vs Standard (S)			*ss*
Priority (P) vs Priority (P)			*pp*

The two letters of the dummy variable indicate the review status of the pairs of drugs (e.g. *ps* refers to a priority drug compared to a standard drug)

[a] A standard drug approved in year *t* is not necessarily of a lesser quality than a priority drug approved in period $s<t$ while a standard drug approved in year *t* is more likely to be of lower quality than a priority drug approved in period $s>t$

[b] NCE refers to new chemical entities (new active ingredients). NCF refers to new combination or new formulation of existing active ingredients.

Table 4 – Data Summary

Observations		
Number of products:		360
"Priority" review	129	
"Standard" review	231	
NCE	285	
NCF	75	
Number of ATC3 level markets		87
Number of products in each ATC3	Mean	4.13
	Sd	1.70
	Min	2
	Max	7
Variables	Mean	Sd
Price (sales weighted)	$18.8	$84.7
Price Hospitals (sales weighted)	$19.0	$84.9
Price Pharmacists (sales weighted)	$18.4	$83.5
DTCA Expend. ('000):	$1177	$4445
Detailing Expend. ('000):	$12107	$22635

Statistics for price, DTCA and detailing are computed using drug/quarter observations during the period of effective exclusivity (7206 observations for prices and DTCA and 1700 observations for detailing).

Table 5 – Advertising Regression

Estimation Strategy: Random Effects										
Dep Vbl:	*($DTCA_0$-$DTCA_1$)*[a]					*($Detailing_0$-$Detailing_1$)*[a]				
Variables:	(1A)	(2A)	(3A)	(4A)	(5A)	(1B)	(2B)	(3B)	(4B)	(5B)
Ps	1.537**	1.401**				1.244**	1.701**			
	(0.717)	(0.569)				(0.524)	(0.419)			
ps^{before}			2.185***					1.365**		
			(0.695)					(0.683)		
ps^{after}			0.966*					0.841		
			(0.560)					(0.567)		
ps^{NCE}				1.660***					1.461***	
				(0.580)					(0.515)	
ps^{NCF}				0.357					0.331	
				(1.240)					(0.521)	
ps^{before}*ps^{NCE}					2.706***					1.701**
					(0.659)					(0.743)
ps^{before}*ps^{NCF}					1.138					1.007
					(1.795)					(0.971)
ps^{after} *ps^{NCE}					1.172**					1.304*
					(0.594)					(0.728)
ps^{after}*ps^{NCF}					-0.162					0.109
					(1.056)					(0.601)
pp	0.390	ref. group	ref. group	ref. group	ref. group	0.565	ref. group	ref. group	ref. group	ref. group
	(0.999)					(0.681)				
ps	ref. group	ref. group	ref. group	ref. group	ref. group	ref. group	ref. group	ref. group	ref. group	ref. group
Age Difference	-0.060**	-0.060**	-0.048*	-0.061**	-0.047**	-0.033	-0.033	-0.027	-0.036	-0.032
	(0.026)	(0.026)	(0.027)	(0.026)	(0.027)	(0.025)	(0.025)	(0.029)	(0.026)	(0.032)
# of Observat.	9347	9347	9347	9347	9347	2866	2866	2866	2866	2866
R-squared	0.245	0.245	0.254	0.245	0.265	0.464	0.464	0.464	0.466	0.466

Heteroskedasticity robust S.E. in parentheses, computed by clustering observations at ATC3 market level. Significance level: *** < 0.01; ** < 0.05; * < 0.10. All the specifications include market (ATC3) dummies, time dummies and firm dummies, and number of molecules in each market.

[a] Dependent variable is computed using the difference of the logarithm of DTCA and Detailing. Given that the "log" of zero is not defined, we actually use $\ln(a_0+1)$- $\ln(a_1+1)$.

Table 6 – Price Regression

Estimation Strategy: IV Fixed Effect				
Dependent Variable:		$\ln(p_{i,t})$		
Variable:	*Name Vbl:*	(1)	(2)	(3)
Lag of Own price	$\ln(p_{i,t-1})$	0.881***	0.941***	0.936***
		(0.026)	(0.018)	(0.018)
Rival price	$\ln(p_{j,t})$	0.011	0.007	0.019
		(0.023)	(0.017)	(0.016)
Own (cumulative) DTCA	$\ln(DTCA_{i,t})$	0.084***		0.025*
		(0.029)		(0.015)
Rival (cumulative) DTCA	$\ln(DTCA_{j,t})$	0.035*		0.004
		(0.019)		(0.005)
Own (cumulative) detailing	$\ln(DET_{i,t})$		0.037**	0.014**
			(0.018)	(0.007)
Rival (cumulative) detailing	$\ln(DET_{j,t})$		0.031**	0.013*
			(0.015)	(0.008)
Herfindahl Index (for ATC3 market)	$hhi_{m,t}$	0.057*	0.087	0.031
		(0.032)	(0.065)	(0.047)
Number of Observations		9190	2666	2666
Centered R-squared		0.602	0.854	0.864
Overidentification Test (Sargan Statistics *p*-values)[a]		0.646	0.226	0.222
Weak Identification Test [b] (Gragg-Donald Wald F Statistics *p*-values)		<0.01	<0.01	<0.01

Heteroskedasticity robust S.E. in parentheses. Significance level: *** < 0.01; ** < 0.05; * < 0.10. All specifications include time dummies and number of molecules in each market. To control for endogeneity of DTCA and detailing, we use lagged values of sales from t-2 to *t-3*.

[a] The null hypothesis is that the instruments are valid, i.e. uncorrelated with the error term

[b] The null hypothesis is that the estimator is weakly identified (i.e. instruments are weakly correlated with included endogenous variables).

Table 7 – Prices in hospital and pharmacies at drug-form level

		Numb of Observation	Mean	Stand Dev.
Price at entry (qrt=1)	Hospital	382	20.73	4.43
	Pharmacies	382	20.86	4.47
	Test of Equality [p-values]		[0.551]	
Price 5 years after entry (qrt=20)	Hospital	382	20.20	4.01
	Pharmacies	382	22.79	4.59
	Test of Equality [p-values]		[0.0004]	

Figure 2 – Distribution of price growth (over five years) at drug-form level