Empirically Characterizing Informational Barriers to Entry in the Anti-Ulcer Drug Market *

Andrea Coscelli
National Economic Research Associates,
London

Matthew Shum
Department of Economics
University of Toronto

October 9, 1998

Abstract

We empirically quantify and characterize informational barriers to entry into the anti-ulcer drug market by studying the diffusion process of the molecule *omeprazole* in the first 30 months after it entered the market. Using a novel panel dataset tracking doctors' complete prescription histories, we specify and estimate a learning model in which doctors, initially uncertain about the quality differential between *omeprazole* and the incumbent molecules, update their beliefs about this differential from *first-hand experience* after observing noisy signals from patients to whom they have prescribed the molecule. We find strong evidence that doctors' uncertainty about *omeprazole*'s quality is resolved by first-hand experience rather than through the pharmaceutical companies' marketing activities, so that learning through first-hand experience explains almost *all* of *omeprazole*'s diffusion path over our three-year sample period. This casts some doubt on previous researchers' findings, utilizing aggregate data, regarding the importance of marketing in securing market share for an entrant brand.

JEL Classification Codes: I10, L10.

Keywords: pharmaceutical industry, barriers to entry, simulated maximum likelihood, diffusion models.

*We thank Drs. Giuseppe Traversa and Roberto Da Cas at the *Istituto Superiore di Sanita'* for making the data available. Shum thanks the Alfred Sloan Foundation for financial support. We thank Tim Bresnahan, Costas Meghir, Fiona Scott Morton, Peter Reiss, Frank Wolak, and participants in seminars at Stanford University and University College, London for their suggestions.

Address correspondence to Shum at the Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada M5S 3G7. Tel: (416) 978-6999; Fax: (416) 978-6713; e-mail: mshum@chass.utoronto.ca.

1 Introduction

First-mover advantage is a well-documented phenomenon in many differentiated product markets (see Urban, Carter, Gaskin, and Mucha (1986) for a survey of the evidence). Economists have tended to attribute this phenomenon to lack of information among consumers about the quality or attributes of an entrant's product; for example, Shapiro (1982), pg. 7 states that

> ...the fundamental source of the entry barrier is an information one: consumers have better information about established brands than about new ones [...] information is the basic barrier to be overcome by a new product...

The doctor/patient relationship is fraught with uncertainty. Doctors have incomplete information on the medical condition of a patient, and which treatment is best for the patient. Oftentimes, doctors learn about the quality of alternative treatments only through direct experience – actual prescriptions of the new molecule. This *first-hand information* accumulates slowly, however, and is confounded by heterogeneity across diagnoses: what works for diagnosis X may not work as well for diagnosis Y. Otherwise, doctors can obtain *second-hand information* from advertising and sales pitches by the manufacturers of medical treatments, articles in medical journals, or recommendations from other doctors. This uncertainty has significant implications for competition and market structure in pharmaceutical markets. If new, innovative medical treatments face informational hurdles in gaining acceptance among doctors, then existing treatments enjoy a "pioneer advantage".

The disadvantages faced by entrants' products due to doctors' lack of information about them are the focus of our analysis. While a quite sizeable theoretical literature (Schmalensee (1982) and Bagwell (1990) are two examples) have developed this idea of "informational barriers to entry", this paper provides an empirical characterization of these barriers in the anti-ulcer drug market, the largest therapeutic drug market worldwide. We focus not only on measuring the extent of informational barriers to entry but, more importantly, explore how these barriers are overcome. The main goal and contribution of this paper is to quantify the importance of information derived from actual ("first-hand") experience with the entrant product relative to informative derived from "second-hand" sources in overcoming barriers to entry.

Using a novel panel dataset of complete prescription histories for a sample of doctors in the Rome (Italy) metropolitan area, we study the diffusion process of a new anti-ulcer molecule (*omeprazole*) during a three-year period (1990-1992). In order to distinguish between first- and second-hand sources of information, we specify a learning model in which doctors, initially uncertain about the quality differential between *omeprazole* and the incumbent molecules, update their beliefs about this differential after observing noisy signals of it from patients to whom they have prescribed the molecule. We call this learning via actual prescriptions "first-hand experience", and compare the effect of information gained thereby on doctors' willingness to prescribe *omeprazole* with information derived from other "second-hand" sources.

In the pharmaceuticals industry, the most important of these second-hand sources of information is advertising and marketing. Schweitzer (1997) p. 46 notes that:

> It is estimated that the [American pharmaceuticals] industry spends at least $5 billion annually on advertising and marketing — more than *$8000* [author's emphasis] for every physician in the United States... In fact, the drug industry spend approximately $1 billion more on marketing and advertising than on research...

The goal of this paper is to determine the importance of first-hand experience relative to more aggregate sources of information (such as marketing) in predicting *omeprazole*'s diffusion during the 1990-1992 period. We find strong evidence that doctors' uncertainty is resolved by first-hand experience (actual prescriptions) rather than through marketing activities and the other potential sources of information captured by time dummies. Even keeping advertising's effects and the time dummies at their estimated levels, the predicted prescription probability remains very low (around 2%) and virtually unchanged over the sample period for doctors without first-hand experience, whereas for a doctor who prescribes *omeprazole* with "average frequency" (relative to the sample), it rises from about 2% to more than 20% over the sample period. Learning through first-hand experience explains almost *all* of *omeprazole*'s observed diffusion into the market over the sample period.

The next section provides some background on the international and Italian anti-ulcer drug markets. Section 3 describes the doctor-level learning model, and section 4 derives the estimating equations associated with this model. Section 5 describes the panel dataset of complete prescription histories that we employ. Results from several specifications of the learning model are presented and interpreted in section 6, and in the last section we draw our main conclusions.

2 Background

2.1 The Anti-Ulcer Drug Market

The anti-ulcer drug market is the largest therapeutic drug market worldwide. It is naturally segmented along different diagnoses (sub-markets), with preferred treatment differing across segments depending on the severity of the diagnosis. These segments are summarized in figure 1 and discussed below.

The two most common diagnoses requiring treatment using anti-ulcer drugs are peptic ulcers and

Figure 1: **Diagnoses and Treatments**

Diagnosis	**Treatment**[a]	**Preferred drugs**[b]
1—Minor heartburn	Drugs or no Prescription	Anti-acids
2—Pathological hypersecretory conditions	Anti-ulcer drugs	*omeprazole*
3—Attack therapy for GERD or Peptic Ulcer	Anti-ulcer drugs	*omeprazole*
4—Maintenance therapy for GERD or Peptic Ulcer	Anti-ulcer drugs	H_2-antagonists

[a]Medical Economics Co. (1997)
[b]Medical Economics Co. (1997)

Gastroesophagal Reflux Disease (GERD). **Peptic ulcers** strike more than 10% of all Americans at one time or another during their lives. An ulcer is a focal area of the stomach (*gastric*) or duodenum (*duodenal*) that has been destroyed by digestive juices and stomach acid. **Gastroesophagal Reflux Disease (GERD)** refers to the backward flow of acid from the stomach up into the esophagus.

During the 1990-92 period covered in our sample, traditional treatments for these two diagnoses consisted in attacking the ulcer either with H_2- receptor antagonists (the market leader *ranitidine*, as well as *cimetidine*, *nizatidine*, *famotidine* and *roxatidine* belong to this class), or protonic pump inhibitors (*omeprazole* was the first approved molecule in the latter class). After healing occurred, the patient would either stop the treatment hoping that no recurrence would occur, or continue taking an anti-ulcer drug at lower dosage as part of "maintenance" therapy.

The third main diagnosis is a **pathological hypersecretory condition**. Different ailments are defined under this label, including Zellinger-Ellison syndrome. Finally, on occasion anti-ulcer drugs are also prescribed for **minor heartburns** (which are, however, often symptomatic of more serious gastrointestinal disorders).

Omeprazole Drugs based on the *omeprazole* molecule entered the Italian market in mid-1990. The Swedish pharmaceuticals giant Astra holds the patent on *omeprazole*, and markets it worldwide under the brand name LOSEC (but in the U.S. as PRILOSEC). LOSEC became the world's best selling drug in 1996, grossing more than US$2 billion, and gaining about half of the worldwide anti-ulcer market.[1] However, this market ascendance was gradual; during the sample period, drugs based on *omeprazole* managed to garner only 20% of the Italian market, and even in mid-1995, it still clearly lagged behind the H_2-antagonist molecules in terms of market share, measured as percentage of total anti-ulcer drug revenues (see figure 2).[2]

Omeprazole inhibits the gastric acid ("protonic") pump, thereby directly blocking acid production irrespective of the acid-producing stimulus. The older H_2-receptor antagonist molecules (such as *ranitidine*, the active ingredient in then-market leader *Zantac*), only inhibit the acid-producing stimuli of histamine at H_2-receptor cells, and are therefore not as effective in suppressing acid production arising from other stimuli.

Indeed, *omeprazole* was found in clinical trials to be more effective than H_2-receptor antagonists in treating serious ulcers and severe esophagitis (a GERD-related condition). On the other hand, other clinical studies on animals found that *omeprazole* produced a dose-related increase in gastric carcinoid tumors. Astra-Merck Inc. (1996) recommends that a standard treatment involving Prilosec last no longer than 8 weeks, making it inappropriate for preventive maintenance therapies. Continuous use of *omeprazole* was advised only for the pathological conditions (e.g., Zellinger-Ellison) where H_2- receptor antagonists were ineffective (Medical Economics Co. (1997)). Furthermore, the molecule is too strong for heartburns.

Recent advances Beginning around 1993 (just after the sample period of our data), researchers found that many duodenal ulcers[3] are due to the presence of a bacterium–*Helicobacter Pylori*. The use of antibiotics to fight the *H.Pylori* infection is a major scientific advance; recent studies show that antibiotics can permanently cure 80-90 percent of peptic ulcers (Graham (1993)). Despite

[1]This information comes from the websites of Astra-Merck and Glaxo.

[2]Each molecule in the Italian market is usually marketed by more than one firm. Usually one of the firms is the patent-holder and the others are licensees. *Omeprazole* was introduced and jointly marketed by three firms (*Bracco*, *Malesci* and *Schering Plough*), none of which was the patent-holder. Curiously enough, *Astra*, the patent-holder, only entered the Italian market in May 1993 (after the end of our sample period), because it did not have a marketing division in Italy at the time *omeprazole* was patented. Therefore *Astra* decided to license the product before directly entering the market.

There are interesting competitive issues between the patent-holder and the licensee to be analyzed, but we abstract away from them and focus on competition *across* molecules. We do not distinguish between different brands of a particular molecule, i.e., between RANIDIL and ZANTAC, which are two different brands of the *ranitidine* molecule. We only focus on the *molecules* prescribed by the doctors. Coscelli (1998a) investigates competition among sellers of the same molecules.

[3]Approximately 90%, according to Astra-Merck Inc. (1996)

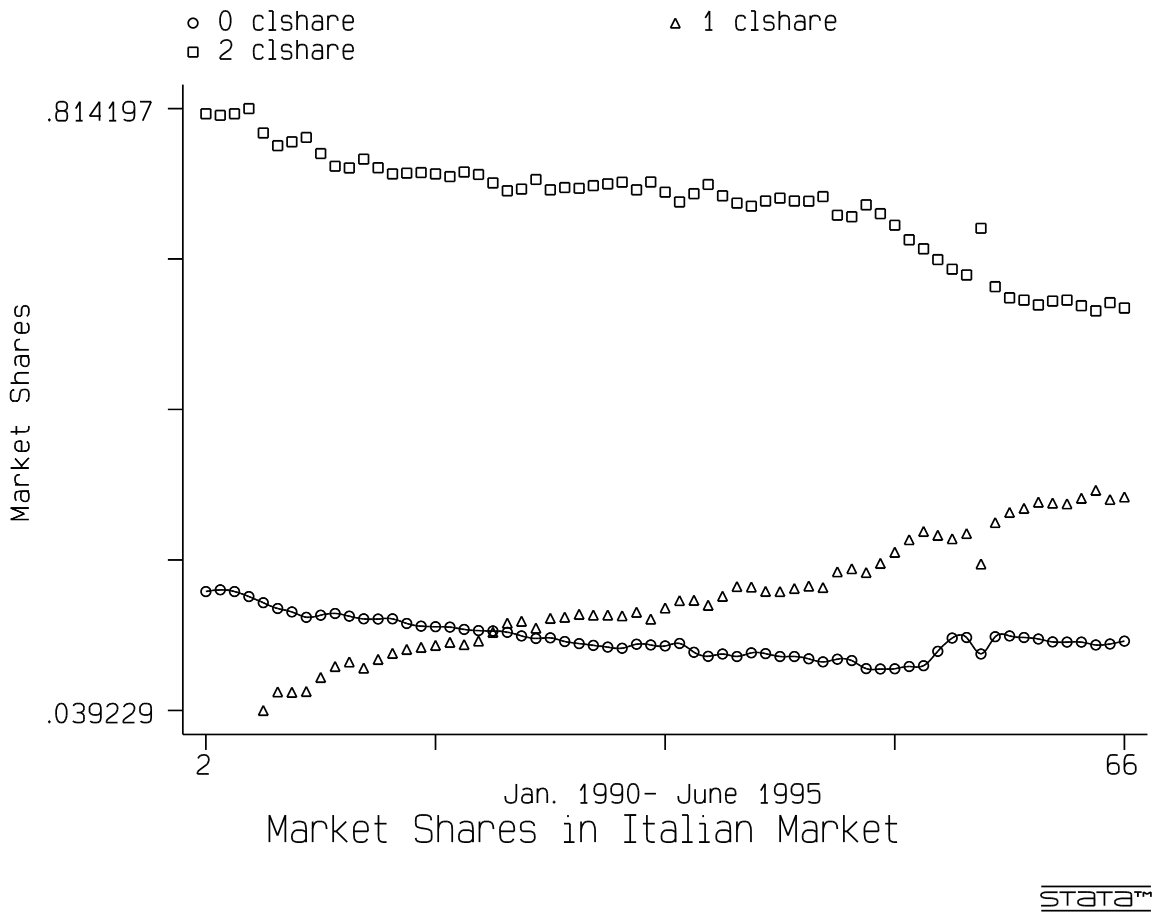

Figure 2: Molecule Market shares in the Italian anti-ulcer drug market
Monthly market shares of drugs based on particular molecules, measured as percentage of total anti-ulcer drug revenues.
1: drugs based on the *omeprazole*molecule
2: drugs based on H_2-antagonist molecules
0: all others (including anti-acids).
Note: *Omeprazole*-based drugs entered the market in June, 1990
Source: IMS Italy.

these advances in curing ulcers, the anti-ulcer drug market has not shrunk appreciably in recent years; in fact, antibiotics are often used in conjunction with anti-ulcer drugs in treating ulcers.[4] Furthermore, antibiotics are not effective against maladies not caused by *H.Pylori*; these include not only gastric ulcers, but also GERD and pathological hypersecretory conditions. According to Astra-Merck Inc. (1996), 50% of new PRILOSEC prescriptions (in the U.S.) were for GERD, 16% for ulcers, and 34% for other conditions.

2.2 Previous Research on Pharmaceutical Markets

Several studies have documented the existence and, more importantly, the nature of barriers to entry into pharmaceutical markets. Early research by Bond and Lean (1977) investigated the phenomenon of pioneer advantage, and reasons under which it persists or erodes in two therapeutic drug markets: diuretics and antiaginals. Not surprisingly, they find substantial pioneer advantage. More interestingly, however, they also find that late entrants offering products containing some therapeutic novelty managed to gain large market shares when backed by heavy advertising campaigns. On the other hand, late entrants offering products of similar quality (i.e., "me-too" brands) failed to gain large market shares, even after expensive advertising campaigns.

Like Bond and Lean, Berndt, Bui, Reiley, and Urban (1997) measure the importance of *marketing* in overcoming pioneer advantage in the anti-ulcer drug market. They perform an econometric analysis of the rise to market dominance of Glaxo's late entrant ZANTAC (*ranitidine*), especially vis-a-vis the pioneer and early market leader TAGAMET (*cimetidine*, marketed by Smith-Kline). Their findings clearly show that product advances (such as those embodied in ZANTAC) don't necessarily translate into success in therapeutic drug markets (in terms of market share) without tremendous marketing muscle. As striking as this result is, however, they never explain the causes of pioneer advantage. The availability of doctor-level prescription histories allows us a unique opportunity to compare the effectiveness of first-hand experience versus promotional efforts in convincing doctors to prescribe the entrant drug, which is important given the magnitude of marketing expenditures in this industry.

A related literature has investigated the extent of competition in pharmaceutical markets by estimating cross-price elasticities between the competing drugs in a market. Unlike the work of Stern (1995) and Ellison, Cockburn, Griliches, and Hausman (1997), however, we abstract away from competition between existing anti-ulcer drugs, and focus squarely on how the substitutability between these drugs, as a whole, and the entrant drugs based on *omeprazole* varied during the sample period.

[4] In fact, Prilosec in combination with the antibiotic Biaxin (clarithromycin) was the first FDA-approved treatment for *H.Pylori* infections Astra-Merck Inc. (1996).

Finally, there has also been an extensive empirical literature in marketing examining diffusion patterns for new products (see Bass, Mahajan, and Muller (1990) for a review). Our analysis links these models – mostly formulated at the macro-level – to micro-level demand data. In the learning model described below, we focus on the accumulation of first-hand experience at the micro- (individual doctor) level to explain the macro-level diffusion pattern. Ackerberg (1996a) has estimated a similar learning model to study consumers' purchase patterns of a new brand of yogurt (Yoplait 150), while Berndt and Pindyck (1998) have recently estimated a diffusion model using aggregate market share data to study the importance of network externalities in determining the diffusion patterns of H_2-antagonist drugs into the anti-ulcer drug market.

3 The Learning Model

3.1 Doctor's Problem

Doctor i's decision problem involves two steps: first a diagnosis ($j \in J$) is formulated, then a treatment (molecule) is chosen. Since our focus is on a doctor's choice between the entrant and the incumbent drugs, we model the second step. We assume that doctor i distinguishes between *two* alternatives: the new molecule, *omeprazole* (alternative 1), and *any* of the other molecules (alternative 0). The utilities from each alternative (for patient k with diagnosis j during period t) are:

$$U^i_{1jkt} = \alpha^*_1 x^i + \beta p_{1t} + \gamma \mathrm{ADV}_{1t} + \xi^*_{1t} + \delta^*_{1j} + \epsilon^i_{1jkt} \qquad \text{if new molecule} \tag{3.1}$$

$$U^i_{0jkt} = \alpha^*_0 x^i + \beta p_{0t} + \gamma \mathrm{ADV}_{0t} + \xi^*_{0t} + \delta^*_{0j} + \epsilon^i_{0jkt} \quad \text{if any other molecule} \tag{3.2}$$

where

- the vector x^i contains doctors' characteristics
- p_{1t} and p_{0t} are, respectively, the price of *omeprazole* and a weighted average of the prices of the incumbent drugs weighted by their market shares at time t.
- ADV_{1t} and ADV_{0t} are, respectively, advertising for *omeprazole* and the *sum* of advertising by the incumbent brands, during period t.
- ξ^*_{1t} and ξ^*_{0t} parameterize period t factors which affect the attractiveness of, respectively, *omeprazole* and the incumbent drugs. These are the same over all doctors, patients, and diagnoses. In particular, ξ^*_{1t} proxies for aspects of the learning process which we don't explicitly model, such as word of mouth and journal articles.
- δ^*_{1j} and δ^*_{0j} parameterize the "unobserved quality" of *omeprazole* and the incumbent drugs, respectively. These are unobserved by the econometrician. Doctors, however, are presumed to

know δ^*_{0j}, and have imperfect information about δ^*_{1j}. How doctors learn about δ^*_{1j} is described further below.

- ϵ^i_{1jkt} and ϵ^i_{0jkt} are *i.i.d.* (over doctors, patients, diagnoses, and time periods) shocks associated with, respectively, *omeprazole* and the incumbent drugs. They are observed by the doctors, but not by the econometrician.

By distinguishing between different diagnoses, we allow the entrant and incumbent anti-ulcer drugs to differ in their effectiveness and suitability across diagnoses. In other words, we allow for "segmentation" or "horizontal differentiation" in the market for anti-ulcer drugs on the basis of diagnosis.

3.1.1 Choice rule

The process whereby the doctor learns about the quality of the new molecule is specified as restrictions on doctor i's choice problem: at time t, doctor i decides whether or not to prescribe *omeprazole* to a patient k with diagnosis $j \in J$. In making this choice, doctor i chooses the option with the higher *per-period*[5] utility. Throughout, agency problems between the doctor and the patient are abstracted away, so that we assume the doctor maximizes the *patient's* utility from the prescription.[6]

The choice rule for the doctor is:

$$\begin{gathered}\text{prescribe } \textit{omeprazole} \text{ iff} \\ E_t(U^i_{1kjt}) > U^i_{0kjt},\end{gathered} \tag{3.3}$$

which involves an expected value because δ_{1j} is a random variable for doctor i, for all diagnoses j.

[5]Doctors are *myopic* in our model — at any given time period, a doctor chooses the molecule with the highest per-period utility based solely on the information that she has at that period. If she were forward-looking, she would choose the molecule with the highest *present discounted utility* and thereby take into account the information that she would gain about *omeprazole* by prescribing it this period. As Shapiro (1982) notes, with forward looking agents, products of uncertain quality can be *more* attractive than products of known quality because there is no "down side risk" associated with the uncertain brand: "if it is a failure [the agent] simply returns to the old brand".

Forward-looking decision making would involve early experimentation with new brands, so that the adoption path associated with a forward-looking learning model may not be adequate to explain the observed gradual rise in market share. We see no evidence of such early experimentation in the data. Current work by Crawford and Shum (1998) examines issues of uncertainty and matching in pharmaceutical demand in a fully forward-looking framework.

[6]The reputation effects concomitant with the long-term nature of many patient–doctor relationships in Italy (the National Health Service requires each enrollee to list a general practitioner) minimize the divergence between doctors' and patients' objective function which form the basis of agency problems.

Since U^i_{1kjt} is linear in δ_{1j}[7], the choice rule (3.3) for the doctors can be rewritten as:

$$\begin{aligned}
&\text{prescribe } \textit{omeprazole} \iff \\
&\alpha_1^* x^i + \beta p_{1t} + \gamma \text{ADV}_{1t} + \xi^*_{1t} + E_t\delta^*_{1j} + \epsilon^i_{1jkt} \; > \alpha_0^* x^i + \beta p_{0t} + \gamma \text{ADV}_{0t} + \xi^*_{0t} + \delta^*_{0j} + \epsilon^i_{0jkt} \longleftrightarrow \\
&\left(\epsilon^i_{1jkt} - \epsilon^i_{0ijkt}\right) \; > \\
&\qquad - \left((\alpha_1^* - \alpha_0^*)x^i + \beta(p_{1t} - p_{0t}) + \gamma(\text{ADV}_{1t} - \text{ADV}_{0t}) + (\xi^*_{1t} - \xi^*_{0t}) + (E_t\delta^*_{1j} - \delta^*_{0j})\right).
\end{aligned} \tag{3.4}$$

Given our assumption that ϵ^i_{1jkt} and ϵ^i_{0ijkt} are independent, and each distributed extreme value, the probability that doctor i prescribes *omeprazole* takes the binary logit form:

$$\begin{aligned}
&Prob(\text{prescribe } \textit{omeprazole}) = \\
&\frac{\exp\left((\alpha_1^* - \alpha_0^*)x^i + \beta(p_{1t} - p_{0t}) + \gamma(\text{ADV}_{1t} - \text{ADV}_{0t}) + (\xi^*_{1t} - \xi^*_{0t}) + (E_t\delta^*_{1j} - \delta^*_{0j})\right)}{1 + \exp\left((\alpha_1^* - \alpha_0^*)x^i + \beta(p_{1t} - p_{0t}) + \gamma(\text{ADV}_{1t} - \text{ADV}_{0t}) + (\xi^*_{1t} - \xi^*_{0t}) + (E_t\delta^*_{1j} - \delta^*_{0j})\right)} \equiv \\
&\qquad \frac{\exp\left(\alpha x^i + \beta \Delta p_t + \gamma \Delta \text{ADV}_t + \xi_t + E_t\delta_j\right)}{1 + \exp\left(\alpha x^i + \beta \Delta p_t + \gamma \Delta \text{ADV}_t + \xi_t + E_t\delta_j\right)}
\end{aligned} \tag{3.5}$$

where we have substituted $\alpha \equiv \alpha_1^* - \alpha_0^*$, $\xi_t \equiv \xi^*_{1t} - \xi^*_{0t}$, and $E_t\delta_j \equiv E_t\delta^*_{1t} - \delta_{0t}$. α, γ, and the ξ_t's are the coefficients to be estimated. The central question in this paper can now be rephrased as whether *omeprazole*'s gradual increase in market share over time is better explained by first-hand doctor-specific learning through prescriptions (which enters the model through changes in $E_t\delta_j$) or through secondhand sources of information (captured by ADV and the ξ_t's). The learning process is explained in detail in the next section.

3.2 Learning Process

We assume that the learning processes are independent across doctors.[8] Therefore, we describe the learning process for doctor i, omitting the superscript i in most of the equations below for expositional clarity. We assume that, at time $t = 0$ (i.e., at *omeprazole*'s entry), she (doctor i) has the following *initial* beliefs about $\vec{\delta}$, the J-dimensional vector of quality differentials between *omeprazole* and the incumbent molecules:

$$\vec{\delta} \sim N\left(\begin{bmatrix} E_0\delta_1 \\ \vdots \\ E_0\delta_J \end{bmatrix}, \begin{bmatrix} \sigma^2_{\delta,1} & 0 & \ldots & 0 \\ 0 & \sigma^2_{\delta,2} & \ldots & 0 \\ 0 & \cdots & \cdots & \vdots \\ 0 & 0 & \ldots & \sigma^2_{\delta,J} \end{bmatrix}\right). \tag{3.6}$$

All the information that she has about *omeprazole*'s efficacy in treating the various diagnoses should be summarized in the initial estimates $E_0\delta_1, \ldots, E_0\delta_J$, while the initial variances $\sigma^2_{\delta,1}, \ldots, \sigma^2_{\delta,J}$

[7] We do not allow risk aversion in the utility function, because it would be difficult to separately identify the risk aversion parameter and the mean of the initial distribution of the "true" quality.

[8] Informational spillovers across doctors ("word of mouth") at the aggregate level are captured by the time dummies.

parameterize her "confidence" in the accuracy of her initial estimates (higher confidence $\longleftrightarrow$ smaller initial variance).

3.2.1 Updating process

In this section, we describe the updating process in some detail. Assume that doctor i begins time period t with beliefs that

$$\vec{\delta} \sim N\left(E_t\vec{\delta}, \Sigma_{\delta,t} \equiv E_t\vec{\delta}\vec{\delta}' - (E_t\vec{\delta})(E_t\vec{\delta})'\right) \tag{3.7}$$

During period t, the doctor prescribes *omeprazole* to k_j of her patients with diagnosis j, and observes k_j noisy signals of δ_j. We assume that these k_j signals ($\mu_{jtk},\ k = 1, \ldots, k_j$) take the following form:

$$\mu_{jtk} = \delta_j + \nu_{jtk} \tag{3.8}$$

where ν_{jtk} is normally distributed, with zero mean. In other words, each prescription yields a draw from a normal distribution centered around the true (but unknown) δ_j. Doctors attempt to form estimates of δ_j from observations of the noisy signals μ_{jtk}'s.

Correlation structure We assume that the correlation structure among the ν's for the different diagnoses are induced by a "single-factor" variance components structure for each ν:

$$\nu_{jtk} = \rho_j\theta_t + \eta_{jtk},\ j = 1, \ldots, J, \tag{3.9}$$

where:

- θ, the common component (across j and k) is distributed $N(0, \sigma_\theta^2)$, i.i.d. over t
- the η_{jtk}'s are independent over j, t, and k, and distributed $N(0, \sigma_{\eta j})$, $j = 1, \ldots, J$
- $\rho_1, \ldots, \rho_J$ are time-invariant parameters (the factor-loadings).

This one-factor variance components specification reduces the number of parameters which we must estimate, while placing mild (and, we believe, acceptable) restrictions on the structure of correlation.[9]

[9] One can extend the current model by allowing for a 2-factor variance components structure, but we have experienced problems identifying some of the parameters in that case.

Given these assumptions, then, the following correlation structure among all the signals (μ's) that doctor i observes in period t emerges:[10]

1. $\mathrm{Var}(\nu_{jtk},\ \nu_{jtk}) = \rho_j^2\sigma_0^2 + \sigma_{\eta j}^2$
2. $\mathrm{Var}(\nu_{jtk},\ \nu_{jtk'}) = \rho_j^2\sigma_0^2$, for $k \neq k'$
3. $\mathrm{Cov}(\nu_{jtk},\ \nu_{j'tk'}) = \rho_j\rho_{j'}\sigma_0^2$, for $j \neq j'$ and $\forall\ k,\ k'$.

Appendix A illustrates this model with an example.

Period-by-period updating Given the normality assumptions on the signals μ as well as on the δ's, a doctor's posterior beliefs about $\vec{\delta}$ given $\vec{\mu}_t$ are described by a normal distribution whose mean and variance can be derived using the multivariate normal conditional mean and variance formulas (cf. Amemiya (1985), p. 3). The computed posterior distribution in period t serves as the prior distribution for period $t+1$. In this way, we derive the sequence of a doctor's posterior distributions over all the periods by repeatedly applying the conditional mean and variance formulas for jointly normally distributed random variables.

In order to do this, we must first characterize the joint distribution of $(\vec{\delta}, \vec{\mu}_t)$, during period t:

$$\begin{pmatrix} \vec{\delta} \\ \vec{\mu}_t \end{pmatrix} \sim N\left(\begin{bmatrix} E_t\vec{\delta} \\ \vec{\delta}_{\mu t} \end{bmatrix}, \begin{bmatrix} \Sigma_{\delta,t} & \Sigma_{\delta,\mu,t} \\ \Sigma'_{\delta,\mu,t} & \Sigma_{\mu,t} \end{bmatrix} \right) \tag{3.10}$$

where $\vec{\delta}_t$ and $\Sigma_{\delta,t}$ are, respectively, the mean and variance-covariance matrix of $\vec{\delta}$ conditional on all the signals received before period t. $\vec{\delta}_{\mu,t}$ and $\Sigma_{\mu,t}$ are the mean and variance-covariance matrix of the vector of signals $\vec{\mu}_t$ (given in equations A.20 and A.21 in the appendix), and $\Sigma_{\delta,\mu,t}$ is the matrix of covariance terms between $\vec{\delta}$ and $\vec{\mu}_t$ (which is easy to derive given equations A.19, 3.8 and 3.9).

The first two moments of the posterior distribution of $(\vec{\delta} \mid \vec{\mu}_t)$ are

$$\vec{\delta}_{t+1} \equiv E_t(\vec{\delta} \mid \vec{\mu}_t) = \vec{\delta}_t + \Sigma'_{\delta,\mu,t}\Sigma_{\mu,t}^{-1}(\vec{\mu}_t - (\vec{\delta}_t)) \tag{3.11}$$

and

$$\Sigma_{\delta,t+1} \equiv \Sigma_t(\vec{\delta} \mid \vec{\mu}_t) = \Sigma_{\delta,t} - \Sigma'_{\delta,\mu,t}\Sigma_{\mu,t}^{-1}\Sigma_{\delta,\mu,t}. \tag{3.12}$$

[10] Note that we allow for correlation among the signals only within a given time period t. As yet, we do not allow these signals to be correlated over time. Furthermore, by allowing for correlation among the disturbances to the signals (i.e., the ν's), the variance-covariance matrices describing a doctor's beliefs about $\vec{\delta}$ become non-diagonal, even though in period 0 (i.e., before *omeprazole*'s entry) this matrix is diagonal (see equation 3.6).

Equation 3.11 yields the means of the posterior distribution of the δ's which are substituted into the logit prescription probabilities. These probabilities form the basis for our likelihood function, which is described in the next section.

The procedure we have just outlined to obtain a doctor's estimates of the δ's given her prescriptions during each period may appear cumbersome, especially given the much simpler posterior mean and variance formulas for sampling from a normal distribution with unknown mean and known variance: DeGroot (1970), ch. 9, contains formulas for both the univariate and multivariate cases.[11] However, our problem fits into neither of these cases. The difference arises because doctors don't observe complete vectors of draws from the multivariate distribution which they are trying to estimate the unknown mean of.[12] For that reason, we need to sequentially derive the posterior means and variance-covariance matrices for the unknown δ's for each period t ($t = 1, \ldots, 31$) in the manner explained above.

Estimated parameters We estimate the elements of the period zero initial mean vector $(E_0\delta_1, \ldots, E_0\delta_J)$ and the diagonal elements of the initial variance covariance matrix $(\sigma^2_{\delta,1}, \ldots, \sigma^2_{\delta,J})$. The true values $\delta_1, \ldots, \delta_J$ are parameters, as are the parameters of the correlation structure $\rho_1, \ldots, \rho_J$ and $\sigma^2_{\eta 1}, \ldots, \sigma^2_{\eta J}$.

4 Empirical model

4.1 Grouped logit model

In this section, we derive the likelihood function which we estimate. Given distributional assumption on the *i.i.d.* ϵ disturbances, the doctor's static expected utility maximization problem described earlier (3.3) yields choice probabilities for each observed prescription in the data. However, since

[11] A regression interpretation is also possible. Essentially, each doctor tries to estimate an unknown population mean using draws (or "signals") from a distribution centered at this mean, with known variance. Her "initial estimate" also constitutes an observation, but has a different variance (the "initial variance", parameterizing her confidence in her pre-sample information) than the other observations.

For the univariate case (i.e., one unknown population mean), the unbiased, minimum variance estimate of the population mean is the "GLS estimator", which is essentially a weighted average of the initial mean and the sample mean of the signals, where the weight given each observation is the inverse of the standard deviation (the "precision") of the distribution from which the observation is drawn. Analogously, for the multivariate case, where an estimate of a *vector* of unknown population means is desired using vectors of draws from a multivariate distribution centered around these means, the efficient estimator is a "joint-GLS" (seemingly unrelated regression) procedure.

[12] To be precise, there are four diagnoses, but during some periods, doctors could observe, for example, two signals for diagnosis 1, 1 signal for diagnosis 3, and none for diagnoses 2 and 4. The theorems in DeGroot (1970), chapter 9, are not suitable for this setting.

we assume that, conditional on diagnosis, all the patients are homogeneous,[13] we can aggregate up to the (doctor, month, diagnosis) level. Therefore a single observation in this model is how many patients of doctor i in the diagnostic cell j are prescribed *omeprazole* in month t (r^i_{jt}) out of the total number of patients with diagnosis j (N^i_{jt}) who are prescribed *any* anti-ulcer drug in month t by doctor i. [14]

Assuming that the ϵ's are distributed *i.i.d.* (over patients and time) extreme value, the likelihood function for a given observation takes the *grouped logit* form. Omitting the i superscript (and defining $V^{i*}_t \equiv \alpha x^i + \beta \Delta p_t + \gamma \Delta \mathrm{ADV}_t + \xi_t$),

$$\mathrm{Prob}(r_{jt} \mid N_{jt}) = \binom{N_{jt}}{r_{jt}} \left(\frac{\exp(V^{i*}_t + E_t\delta_j)}{1+\exp(V^{i*}_t + E_t\delta_j)} \right)^{r_{jt}} \left(\frac{1}{1+\exp(V^{i*}_t + E_t\delta_j)} \right)^{N_{jt}-r_{jt}} \tag{4.13}$$

Second, the expected value of the "true" quality of *omeprazole* for diagnosis j in period t, $E_t\delta_j$, needs to be included in the grouped logit choice problem defined in equation (4.13). The main problem we encounter in the estimation process is that while the doctor observes the realization of the signal, μ_{jt}, we as econometricians don't. This implies that the likelihood function for a given sequence of prescription frequencies for a given (doctor-diagnosis type) combination involves an integral over the distribution of the unobserved signals:

$$\begin{aligned} &\mathrm{Prob}(r_{11}, \ldots, r_{1T}, , \ldots, r_{J1}, \ldots, r_{JT} | N_{j1}, \ldots, N_{jT}, , \ldots, N_{J1}, \ldots, N_{JT}) = \\ &\int \left[\prod_{t=1}^{T} \prod_{j} \binom{N_{jt}}{r_{jt}} \left(\frac{\exp(V^{i*}_t + E_t\delta_j)}{1+\exp(V^{i*}_t + E_t\delta_j)} \right)^{r_{jt}} \times \left(\frac{1}{1+\exp(V^{i*}_t + E_t\delta_j)} \right)^{N_{jt}-r_{jt}} \mid \vec{\mu} \right] dF(\vec{\mu}) \end{aligned} \tag{4.14}$$

where $\vec{\mu}$ is the vector of signals observed by doctor i for patients with diagnosis j (but unobserved by the econometrician). Since a doctor observes one signal for each patient to whom she prescribes *omeprazole* , the dimensionality of $\vec{\mu}$ is $R_T \equiv \sum_{j=1}^{J} \sum_{t=1}^{T} r_{jt}$, the number of prescriptions of *omeprazole* which the doctor wrote during the entire 31-month sample period to *all the* patients. In our data, R_T can run into the hundreds, making the integral intractable using traditional quadrature methods. We use simulation techniques (see Stern (1997)) to evaluate these integrals.

[13] This assumptions implies that, among all the patients receiving diagnosis j by doctor i at time t, the unobserved *iid* component drives whether or not a particular patient receives *omeprazole*. This is a standard assumption in group data models. More general treatments would require multiple integration over patient heterogeneity distribution within each cell.

On the other hand, since previous work on this dataset has shown the importance of patient level characteristics in the choice process, in future work, we would like to generalize the current model by assuming a parametric distribution of patient heterogeneity within cells. (Some recent work by Berry, Levinsohn, and Pakes (1997) has explored ways of pooling micro and macro data in estimating demand models.) This would allow us to estimate knowledge spillovers across patients for the doctor.

[14] This is equivalent to assuming that the doctor follows a random treatment strategy; that is, the doctor every month decides the proportion of her patients who will be treated with *omeprazole*, and randomizes among them.

4.2 Simulated maximum likelihood estimation

We will estimate our model using simulated maximum likelihood. This procedure contains several steps. First of all, for each (doctor-diagnosis type) combination, we draw M sequences of disturbances with mean 0 and variance σ_j^2. Each sequence consists of R_T signals. Note that by definition (given above), R_T varies across doctors. In the results reported in this paper, we used $M = 10$.

Secondly, for each sequence of signals, we calculate the 31 (one for each month) means $E_t\delta_j$, $t = 1, \dots, 31$ using the standard normal conditional mean formula (reported in equation 3.11). Furthermore, the posterior mean of δ_j in month t will serve as the prior mean for month $t+1$. Having calculated the $E_t\delta_j$'s for $t = 1, \cdots, 31$ for a given drawn sequence of signals, we can calculate the grouped logit probabilities (in equation (4.13)) by substituting the calculated $E_t\delta_t$'s. Finally, we averaged the calculated grouped logit likelihood function (one for each drawn sequence) over all the drawn sequences.

Hence, the log-likelihood function for all (doctor-diagnosis type) observations is:

$$\begin{aligned}\text{LogL} = \sum_{i=1}^{326} \log \frac{1}{M} \sum_{m=1}^{M} \Bigg[\prod_{t=1}^{31} \prod_{j=1}^{J} \binom{N_{jt}}{r_{jt}} \left(\frac{\exp(V_t^{i*} + E_t\delta_j)}{1 + \exp(V_t^{i*} + E_t\delta_j)} \right)^{r_{jt}} \times \\ \left(\frac{1}{1 + \exp(V_t^{i*} + E_t\delta_j)} \right)^{N_{jt} - r_{jt}} \mid \vec{\mu^m} \Bigg] \end{aligned} \tag{4.15}$$

where 326 is the total number of doctors in the sample and J, the number of diagnosis groups, is 4 (note that the signals $\vec{v}$ are now indexed by m, which labels the drawn sequences).[15]

4.3 Controlling for Unobserved heterogeneity across Doctors

In the learning model described above, the rate of learning depends mostly on the diagonal elements — the variances — of the variance-covariance matrix of the signals (equation A.21 in the Appendix), which parameterize the "noisiness" of the signals. Roughly speaking, noisier signals (larger variances) slow down learning. In the data, this rate of learning is identified off of *serial correlation* in omeprazole prescriptions over time, at the individual doctor level [16]. A high degree

[15] Note that in the above equation we take averages of the simulated probabilities before taking the logarithm. It would have been much easier to average over the simulated log probabilities, since in that way we avoid the possible computational problems (specifically, numerical underflows) which can arise when taking the product over t and j of the prescription probabilities. The likelihood function that we maximize is $\sum_i \log E_{\vec{v}} \text{Prob}(p_{ijt} \mid N_{ijt})$. Simulation is necessary simply to calculate the expectation, which occurs *inside* the log operation. Since $\sum_i \log E_{\vec{v}} \text{Prob}(p_{ijt} \mid N_{ijt}) \neq \sum_i E_{\vec{v}} \log \text{Prob}(p_{ijt} \mid N_{ijt})$ the easier alternative (simulating the log probabilities) would be inappropriate.

[16] The time dummies should soak up serial correlation at the aggregate level.

of serial correlation at the individual doctor level would imply "fast" learning, while absence of serial correlation would imply slow learning.

As such, one possible interpretation of the learning model is that it provides a structural explanation for serial correlation (conditional on covariates like *omeprazole*'s price and advertising) at the individual doctor level in prescriptions of *omeprazole*. We must also be sure to control for other, simpler explanations of this micro-level serial correlation. The most obvious one is *doctor heterogeneity*. To see this more clearly, assume that doctors know the true quality of *omeprazole*, but vary in their (time-invariant) propensity to prescribe it. What kind of prescription paths will emerge here? Doctors with high propensities will always prescribe a lot of *omeprazole*, while those with low propensities always a small amount — leading to a high degree of serial correlation in *omeprazole* prescription at the individual doctor level. Therefore we estimate specifications which control in varying degrees for both both observed (using doctor-level covariates) and unobserved heterogeneity across doctors.

We assume that unobserved heterogeneity is summarized in the variable κ_i, which shifts the intercept of doctor i's net per-period utility from prescribing *omeprazole*at time t. The time-invariant κ_i parameterizes doctor i's unobserved propensity to prescribe the *omeprazole*-based drugs over the incumbent drugs. This unobserved propensity may arise due to, for example, benefits that doctor i receives from manufacturer's of *omeprazole*-based drugs which are unobserved by the econometrician.

κ^i is modeled in *random effect* fashion; i.e., we specify a distribution for it (assumed identical over all doctors i) and integrate out over this distribution in evaluating the likelihood function. Due to the complexity of the likelihood function (equation 4.15) we restrict κ^i to have a 2-point distribution:

$$\kappa^i = \begin{cases} \kappa^1 & \text{with probability } q \\ \kappa^2 & \text{with probability } 1-q \end{cases} \quad \text{for all } i. \tag{4.16}$$

Furthermore, we normalize κ^i to have zero mean; i.e., $E\kappa^i = 0 \longleftrightarrow \kappa^2 = \frac{-q\kappa^1}{1-q}$, so the specification adds two parameters (κ^1, q) to our model.

After controlling for unobserved heterogeneity in this fashion, the resulting simulated log-likelihood function for the prescriptions written by doctor i takes a multinomial form:

$$\log l^i = q \log(l^{i,s} \mid \kappa^1) + (1-q)\log(l^{i,s} \mid \kappa^2) \tag{4.17}$$

where (similar to equation 4.15)

$$\log(l^{i,s} \mid \kappa) =$$
$$\log \frac{1}{M} \sum_{m=1}^{M} \left[\prod_{t=1}^{31} \prod_{j=1}^{J} \binom{N_{jt}}{r_{jt}} \left(\frac{\exp(V_t^{i*} + E_t\delta_j + \kappa)}{1 + \exp(V_t^{i*} + E_t\delta_j + \kappa)} \right)^{r_{jt}} \times \left(\frac{1}{1 + \exp(V_t^{i*} + E_t\delta_j + \kappa)} \right)^{N_{jt} - r_{jt}} \mid \vec{\mu^m} \right] \tag{4.18}$$

5 Data

We employ data collected by the Italian National Institute of Health which records, for a 10% sample[17] of the doctors in the metropolitan area of Rome, all prescriptions of anti-ulcer drugs (therapeutic class A02B[18]) to all their patients during a three-year period (1990-1992). A prescription episode is the unit of observation in the dataset. The dataset contains more than 660,000 observations, each of which records the identity of the patient, the prescribing doctor, the drug prescribed, and the year and month of the prescription; 326 doctors, and 174,000 patients are represented in the data. The median number of anti-ulcer drug prescriptions for the 326 doctors is just under 2,000 prescriptions. Only a few doctors fill less than 1,000, or more than 3,500 prescriptions for anti-ulcer drugs during the sample period. Therefore, the doctor sample is fairly homogeneous in terms of practice size.

Figure 3 shows the in-sample market shares during the three-year period for the *omeprazole*-based drugs and two large groups of competitors: H_2- receptor antagonists and all other drugs (including anti-acids). The latter two groups of drugs constitute the non-*omeprazole* alternative (or "outside good") in our model. Comparing with figure 2, which shows market shares in the Italian national market, one notes that the rise in *omeprazole* is more pronounced in the national market than in our sample. Naturally, this is accompanied by steeper drops in the national market shares of the molecules in the H_2-antagonist and "all other" categories.

5.1 Patient heterogeneity: different diagnoses

As discussed in section 2, the anti-ulcer drug market is naturally segmented on the basis of diagnoses, so that a given molecule differs in its usefulness in treating different diagnoses (see figure 1).

[17]The doctors in the sample were not chosen following any sampling technique, since the only information available was their "id" number.

[18]This four digit code, the ATC code, is an international classification system according to which drugs are divided into different categories by target organ, mechanism of action, and chemical and therapeutic characteristics.

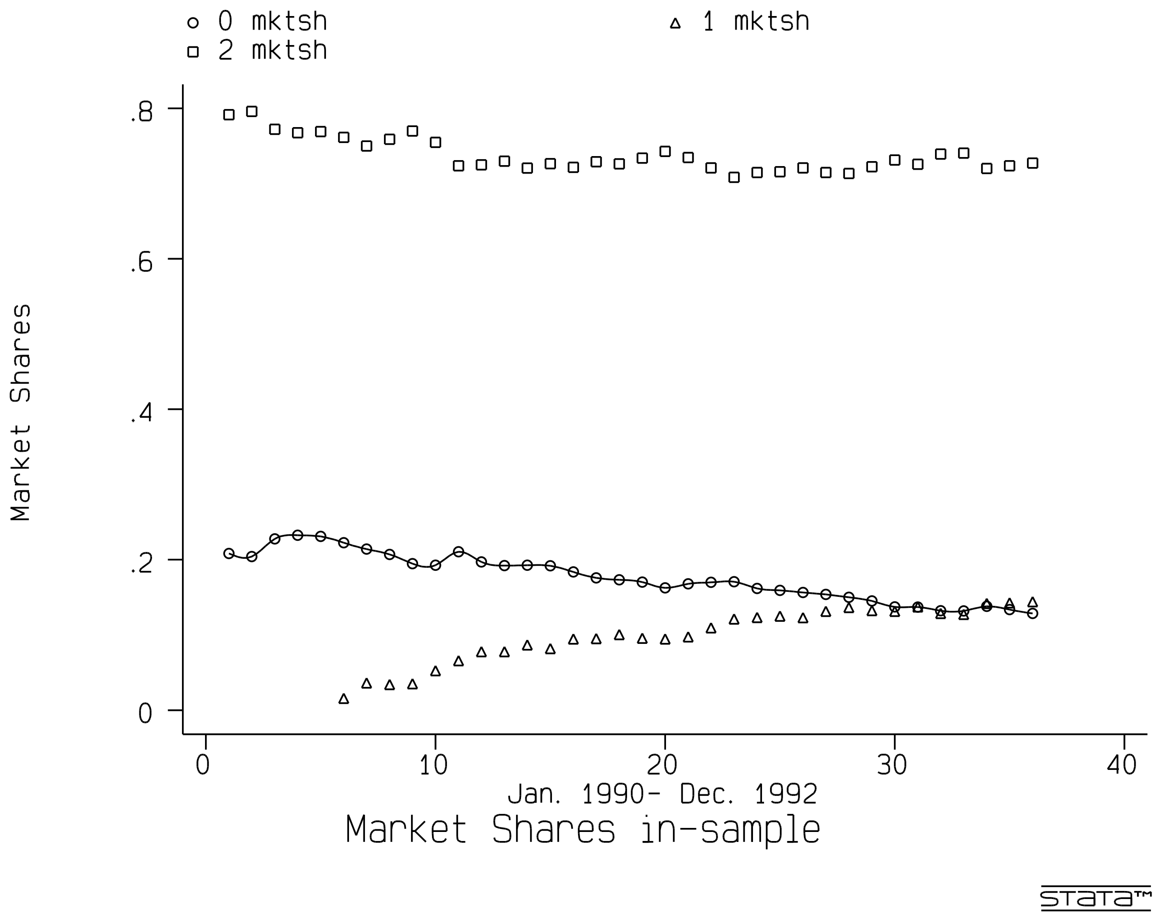

Figure 3: In-sample monthly market shares: t=1,...,36–January 1990-December 1992.
1: drugs based on the *omeprazole*molecule
2: drugs based on H_2-antagonist molecules
0: all others (including anti-acids).
Note: *Omeprazole*-based drugs entered the market in June, 1990
Source: authors' computations using the data described in section 5.

This consideration motivates the multivariate learning model presented in section 3, in which we permit *omeprazole*'s quality differential δ to vary across diagnoses. However, we do not observe a patient's actual diagnosis in the data: how can we allocate our observations to the four different diagnoses listed in figure 1? We observe the quantity of anti-ulcer drugs prescribed, which varies according to the diagnosis. Therefore, after careful perusal of the medical literature, we have classified observations into the four diagnosis groups based on the *length* and *intensity* of treatments. Figure 4 describes the diagnoses, the possible treatments, the preferred drugs and the criteria we use to map prescriptions into diagnoses (Medical Economics Co. (1997)). More details are given in Appendix B.

Figure 4 shows the proportion of the observations allocated to the 4 diagnostic classes. Cell 2 contains less than 5% of the prescriptions, while cell 4 contains more than 50% of the total number of prescriptions in the sample. These proportions are consistent with the incidence in the overall Italian population (peptic ulcer 59%, GERD 5%, minor heart-burns 23%, and "others" 12%) according to a survey of Italian doctors in 1991 cited in Pertile (1995).

Figure 4: Observations in the 4 diagnostic cells

Diagnostic cell	**Empirical Distinction**[a]	**Frequency**	**Percent**
1–Minor heartburn	Patient has <= 2 in-sample prescriptions	135,466	20.49
2–Pathological hypersecretory conditions	$Q >= 133\%$ average monthly quantity for ulcer in *every* month	32,362	4.89
3–Attack therapy for ulcer or GERD	$Q >= 133\%$ average monthly quantity for ulcer in the month	132,361	20.02
4–Maintenance therapy for ulcer or GERD	$Q < 133\%$ average monthly quantity for ulcer in the month	361,054	54.60

Sample: N=661,243.

Source: data described in section 5.

[a]Prescription assigned to diagnoses by the authors using the daily dosage for the average patient requiring an ulcer treatment as suggested by Medical Economics Co. (1997).

The large incidence of prescriptions for diagnosis 1 ("minor heartburns") evidenced in figure 4 is troubling. It suggests that "heartburn" may be a mischaracterization of these patients' gastrointestinal woes — as discussed before, standard treatment for ulcers calls for a (short) period of intense "attack" therapy (using a powerful molecule such as *omeprazole*) followed either by a maintenance period or cessation of treatment. Diagnosis 1 may well encompass many patients who underwent attack therapy followed by cessation, for whom *omeprazole* would have been an appropriate molecule.[19]

Section C in the Appendix provides more details on the data, in particular describing the covariates which we constructed for estimating the learning model.

6 Estimation Results from Learning Model

6.1 Three different specifications

Figure 5 contains three specifications of the 4-Diagnosis model described above, where we have categorized the various patient-visit observations into four diagnosis cells and the noisy signals of *omeprazole*'s quality to be drawn from four different distributions, characterized by diagnosis-specific mean ($\delta_i,\ i = 1, \ldots, 4$) and variance ($\sigma_{\eta i},\ i = 1, \ldots, 4$). We estimate a No-Spillover version (**MODEL A**) in which signals are not allowed to be correlated across diagnoses, and a Spillover version (**MODEL B**) where this correlation is allowed to be nonzero. Finally, in **MODEL C** we augment the Spillovers model to control for doctor unobserved heterogeneity, as given by the likelihood function in equation 4.17.[20]

Across all specifications, we set σ_0^2 (the variance of the common component θ's) equal to 1; our attempts to estimate it have not been successful. On the basis of the likelihood function values alone, the Spillover model with doctor heterogeneity (**MODEL C**) provides, by far, the best fit of the data.

[19]Furthermore, "survival bias" may be present here: perhaps those patients whose treatment lasted only two months or less (and therefore classified in diagnosis 1) are exactly those for whom *omeprazole* had a high success rate. More generally, if *omeprazole* is a very effective molecule, leading to less frequent prescriptions and shorter treatment lengths, then our classification scheme can lead to selection bias because, essentially, inclusion in a particular diagnostic cell is endogenous. Therefore, as a specification check, (see footnote 20) we also estimated a model assuming patient homogeneity, without attempting to distinguish among diagnoses. In future work, we plan to explicitly model the effect of treatment option on treatment length ("success").

[20] In order to gauge the statistical superiority of the 4-Diagnosis model, which is prone to the criticism that the criteria whereby we group patients in diagnosis cells may be endogenous, we also estimated a 1-diagnosis model in which all the signals which a doctor observes about *omeprazole*'s quality are $i.i.d.$ draws from a distribution with mean δ and variance σ_η^2. The estimates (which are available from the author) vary little qualitatively and quantitatively from the those obtained from the 4-diagnosis models (models A, B, C), thus providing some indirect evidence that this type of endogeneity is not a big problem here.

Figure 5: **Results for 4-Diagnosis models**

Number of observations: 10,106 (326 doctors/31 months)

Dependent variable: $r^i_{jt} = k$ if doctor i prescribes *omeprazole* k times to patients with diagnosis j during month t, when her total number of anti-ulcer prescriptions to patients with diagnosis j is N^i_{jt}

	MODEL A		**MODEL B**		**MODEL C**	
	No Spillovers		*W/ Spillovers*		*W/ Spillovers*	
	estimates	s.e.	estimates	s.e.	estimates	s.e.
Choice characteristics						
Δp: $(\text{PRICE}_{omep} - \text{PRICE}_{og})/1000$	-0.033	0.092	-0.021	0.028	-0.010	0.420
ADVERTISING (FLOW)	1.957	0.247	1.890	0.064	1/025	0.861
Doctors' characteristics						
HERFINDAHL PRODUCT LEVEL	0.009	0.015	0.009	0.003	-0.009	0.001
HERFINDAHL MOLECULE LEVEL	-0.361	0.013	-0.349	0.012	-0.728	0.025
QUANTITY/100	-0.027	0.0004	-0.033	0.001	-0.009	0.001
Learning process parameters						
Initial Distribution						
INITIAL MEAN – DIAG. 1 $E_0\delta_1$	-3.917	0.095	-3.391	0.041	-3.497	0.257
INITIAL MEAN – DIAG. 2 $E_0\delta_2$	-3.293	0.098	-3.226	0.033	-2.784	0.253
INITIAL MEAN – DIAG. 3 $E_0\delta_3$	-3.662	0.094	-3.550	0.037	-3.123	0.253
INITIAL MEAN – DIAG. 4 $E_0\delta_4$	-4.376	0.094	-4.242	0.034	-3.903	0.249
INITIAL VAR-COV MATRIX	see Appendix, section D					
Other parameters						
True quality diff.: diag. 1 (δ_1)	-1.097	0.098	-0.898	0.034	-1.148	0.254
True quality diff.: diag. 2 (δ_2)	-0.041	0.096	-0.047	0.024	-0.086	0.253
True quality diff.: diag. 3 (δ_3)	-0.975	0.096	-0.432	0.038	-0.839	0.251
True quality diff.: diag. 4 (δ_4)	-1.384	0.091	-1.815	0.039	-1.846	0.254
VAR-COV MATRIX OF SIGNALS	see Appendix, section D					
Unobs'd Het. Dist'n Parms.						
κ^1					-0.223	0.016
q					0.614	0.066
TIME PERIOD DUMMIES	yes		yes		yes	
Log-likelihood fxn	-54700.59		-54343.49		-52888.52	
M (# sim. draws)	10		10		10	

6.2 Results and interpretation

Informational barriers to entry According to all the sets of results, doctors initially assume that the basket of incumbent drugs is better than *omeprazole* in treating every diagnosis (we estimate the initial mean of the quality differential between *omeprazole* and the basket of incumbent drugs to be large and negative).[21] Our estimates of the δ's, are also all negative but, for all the specifications, much higher than the initial means, indicating that, relative to their beliefs later in the sample period, doctors are very pessimistic about *omeprazole*'s quality at the time of its entry.

This pessimism initially biases doctors against prescribing *omeprazole*, and is the basis of informational barriers to entry into this market. Note that uncertainty by itself is not enough to lead to informational barriers to entry — we require also, as in this case, initial estimates of the new product's quality which are biased downwards (pessimistic). As Shapiro (1982) pointed out, if agents choose the product yielding highest *expected* utility (as our doctors do here), uncertainty about product quality, by itself, can cause a new product's expected utility to be higher than an incumbent's (known) utility, thus actually *facilitating* entry. Without something else which shifts down agents' *ex ante* utility from trying the new product — pessimism in our analysis here, and "brand loyalty" in Shapiro's case — informational barriers to entry will not arise.[22]

Initial estimates and true values of δ's The spillover parameters are small in magnitude (as demonstrated by the small off-diagonal elements in the second matrix in figure 9), and not very precisely estimated. Therefore it is not surprising that the results for the two models with (B) and without (A) spillovers are very similar.

The only way they differ is in what the estimates of the initial means imply about doctors' initial ranking of *omeprazole*'s efficacy relative to the basket of incumbent drugs for the four different diagnoses. For model A, the doctors start with an initial ranking of (from highest to lowest efficacy): 2,3,1, and 4. For model B, this ranking is 2,1,3,4 (which differs from the model A results in implying that doctors believe *omeprazole* to be more effective in treating heartburns (diagnosis 1) than attacking peptic ulcers (diagnosis 3)). However, the "true" ranking (from the estimates of the true quality differentials) is the same in all three models: 2,3,1,4. This agrees with the initial ranking in both models A and C implying that, while doctors systematically underestimate *omeprazole*'s quality upon its entry into the market, they are correct in their assessments of *omeprazole* 's relative

[21] We assume that *all* doctors have the same initial beliefs about $\vec{\delta}$, *omeprazole*'s "true" quality differential versus the incumbent drugs. This assumption is consistent with the idea that doctors' initial beliefs are shaped by advertising taking place before entry, and publications in medical journals.

[22] Let's quote Shapiro again: "Basically, the argument that entrants are penalized because of subjective uncertainty about their products is wrong, although they are penalized *if their products are on average expected to be worse than established brands*" (Shapiro (1982), pp. 5-6, our emphasis.)

efficacy across different diagnoses.

The actual magnitudes of the estimated initial means and "true" quality differentials are difficult to interpret, since they depend highly on our choice of the other covariates to include in the utility specification. In section 6.3, we present some computational results which illustrate the extent to which doctors update their beliefs concerning *omeprazole*'s quality differential, as implied by these estimates.

Rate of learning Figure 8 (in Appendix D) contains the estimated initial (i.e., before *omeprazole*'s entry) variance-covariance matrices for models B and C, while figure 9 contains the estimated variance-covariance matrices for a hypothetical vector of disturbances in the signals (the ν_{jtk}'s) for a time period t. The estimated magnitudes for the initial variance-covariance matrices, calculated in figure 8, are extremely small (roughly three orders of magnitude smaller than the estimates of the initial means), implying a great deal of confidence in these initial (and, as we saw above, pessimistic) estimates.

For model C, the smallest element in the initial variance-covariance matrix (the bottom one in figure 8) is for diagnosis 3, "attack" therapy, implying that doctors are most confident about their initial estimate of *omeprazole*'s efficacy in treating this diagnosis. This is not surprising since, as discussed before, the medical evidence seemed unambiguously clear about *omeprazole*'s effectiveness in attacking ulcers.

On the other hand, the finding that doctors are very confident in their pessimistic estimates is troubling: if we were to find, in addition, that the signals which doctors receive are drawn from distribution with *large* variances (i.e., "noisy" signals), the learning process would be slow, since the initial confidence will exacerbate the downward bias in doctors' subsequent updated estimates of the δ's.

Fortunately, this is not true. The magnitudes of the diagonal elements of the estimated variance-covariance matrices of the signals (reported in figure 9) are also small, similarly implying that the signals are very precise. The model C results indicate that the signals from diagnosis 3 patients — those undergoing "attack" therapy, according to our classification scheme — are the most precise.

Finally, the model C estimates of the off-diagonal elements for the variance-covariance matrices for the signals indicate mostly positive spillovers, except for the signals associated with diagnosis 4 ("maintenance" therapy). The negative values for the correlation terms involving this diagnosis indicate that a positive outcome in prescribing *omeprazole* to a patient for diagnoses 1–3 leads doctors to deem *omeprazole* is less attractive for maintenance purposes — not surprising, since (as described before) maintenance therapy typically requires longer than the eight weeks which is the appropriate length of a treatment involving *omeprazole*. This confirms the last column of figure 1 (derived independently from medical sources) that the anti-ulcer market is *horizontally differentiated*: since a doctor will rank the two alternatives (*omeprazole*-based drug versus incumbent drugs) differently depending on the diagnosis of the patient who she's prescribing to.

Advertising The coefficient on advertising is positive and significant in all the models (the point estimate is 1.890 for the Spillovers model), implying that doctors are more likely to prescribe *omeprazole* in those quarters when Italian detailing activity on *omeprazole*'s behalf is relatively intense. This confirms the findings in Berndt, Bui, Reiley, and Urban (1997) that marketing was key in ensuring success in the anti-ulcer market. The question then arises: how effective is advertising vis-a-vis direct first-hand experience in convincing doctors to prescribe *omeprazole*? This will be explored further below.

Price The coefficient on Δp, the price differential between *omeprazole* and all the other anti-ulcer molecules, is negative but not significant across all three models (-0.010 for model D). This finding is not surprising, since the Italian health care institutions (all anti-ulcer drugs are paid for by the National Health Service) do not force doctors to internalize the costs of the drugs that they prescribe. Furthermore, there is not much price variation in the sample, and the price of the alternative drugs is a weighted average of the prices of *all* the incumbent drugs.

Doctor-specific covariates Finally, we turn to the doctor-specific covariates which capture observed time-invariant heterogeneity across doctors in their willingness to prescribe *omeprazole*. The results described here are robust across all three specifications of the learning model. The two covariates which measure dispersion in prescription patterns prior to *omeprazole*'s entry are generally statistically significant[23] but have different signs: we find that greater dispersion at the *molecular* level (lower value for the molecular-level herfindahl index) increases the probability of prescribing *omeprazole*, but not greater dispersion at the product level. Furthermore, larger quantity prescribed is associated with *lower* propensity to prescribe *omeprazole*. Clearly these three covariates will be quite correlated at the doctor-level, so that these changes in signs across different specifications is not surprising.

6.3 Importance of first-hand versus second-hand information

Next, we turn to the main question we address in this paper: how important is first-hand experience versus other second-hand sources of information, such as advertising in overcoming barriers to entry in this market? Some answers are given in figure 6, which decomposes the across-time changes

[23] The sole exception is the coefficient on HERFINDAHL PRODUCT LEVEL in model B.

in the predicted market shares of *omeprazole* into (1) a part due to outside sources of learning (parameterized by advertising and the 31 time dummies in our specifications); and (2) a part due to learning by first-hand experience (i.e., due to the rise in doctors' estimates of the δ's) for the model C results. The curve marked avg represents the average doctor in our sample (whose evaluation of the quality of *omeprazole* changes in every period), and the curve marked nopres represents a doctor who never prescribes *omeprazole*, who therefore learns about it only on the basis of advertising and "word of mouth" (proxied by the 31 period-specific dummies) alone. For purposes of comparison, the actual (in-sample) market shares of *omeprazole* for each diagnosis are also plotted, by the curve marked Actual.

It is clear that the actual observed market share path is remarkably well predicted for "average" doctors (the avg lines), but severely underpredicted by doctors who don't prescribe at all (the nopres lines). This implies that the likelihood that a doctor prescribes *omeprazole* does not change much over time unless the doctor actually starts prescribing it, thereby receiving direct verification of its quality relative to the incumbent brands: *first-hand experience plays a* critical *role in overcoming doctors' initial pessimism about the quality of the new entrant.*

Recall that these results obtained even though we "stacked the cards" against it by including a full set of time-period dummies to soak up any factors left unexplained by either the learning process or the included covariates. These graphs give very clear evidence that first-hand experience is much more effective than marketing at overcoming informational barriers to entry into the anti-ulcer drug market: learning through first-hand experience accounts for almost *all* of the rise in *omeprazole*'s in-sample market share over time! [24]

The height of these barriers to entry can be quantified in terms of the extra marketing costs which manufacturers of the entrant *omeprazole*-based drugs had to incur in order to secure a given market share for these drugs. The tru_noad lines give us some idea of the magnitude of these marketing costs. It shows what the path of market shares would have been if the average doctor were perfectly informed about δ, in the absence of advertising (here measured in the number of doctors visits made on behalf on a particular drug). For all the diagnoses, the tru_noad line lies above the other lines, which is not surprising, given the finding above that almost all the rise in market share is attributable to the learning process. Clearly, given our estimates, *marketing efforts would be superfluous if doctors were perfectly informed* about the quality differentials of the *omeprazole*-based drugs versus the incumbent drugs.

[24] This finding that advertising is an (albeit imperfect) "substitute" for first-hand experience of a product has been documented in several consumer non-durables product markets, such as cereals (Shum (1998)) and yogurt (Ackerberg (1996b)).

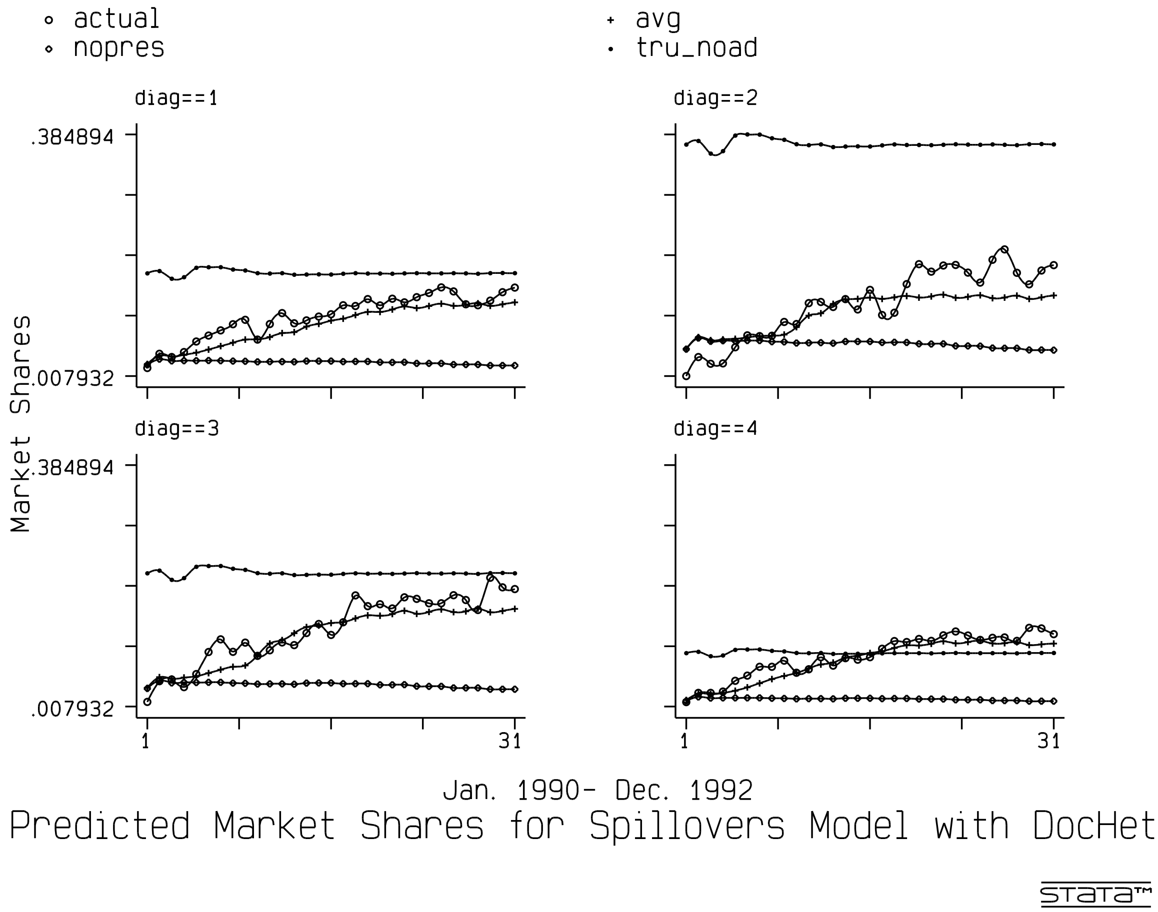

Figure 6: **Predicted Market Shares for Spillovers Model controlling for unobserved doctor heterogeneity (Model D)**

On the y-axis, market share of *omeprazole* in treating different diagnoses over time for 3 scenarios:

(1) **Actual:** actual (in-sample) market share of *omeprazole*

(2) **avg:** predicted market shares for doctor who prescribes the "average" amount of *omeprazole*

(3) **nopres:** predicted market share for doctor who never prescribes *omeprazole* (i.e., who learns about *omeprazole* on the basis of advertising and the time dummies alone)

(4) **tru_noad:** predicted market share for doctor in perfect information setting (who knows true δ), but without advertising

Only time dummies and δ's vary over time; the other covariates (price and doctor characteristics) are set at their sample means.

7 Conclusions

We have used unique panel data on doctor-level prescription histories to explicitly quantify informational barriers to entry into the anti-ulcer drug market. We focused on the entry of the new molecule *omeprazole* into the Italian market in 1990. Although *omeprazole* represented an innovation upon previous treatments, its market share grew only gradually during our sample period (1990-1992), but by 1996 it became the best selling prescription anti-ulcer drug in the world.

According to our model, *omeprazole*-based brands suffered a competitive disadvantage upon their entry into the Italian anti-ulcer market due to doctors' initial pessimism about the quality of the drug. Our main finding is that even after controlling for time effects, advertising and doctors' heterogeneity, we find that most of the observed gradual growth in demand for *omeprazole* is predominantly explained by doctors' accumulation of positive first-hand experience of these drugs: through prescription of these drug to their patients, doctors become more optimistic about *omeprazole*'s quality, and this increasing optimism drives the gradual rise in *omeprazole*'s market share.

This finding appears at odds with the findings in Berndt, Bui, Reiley, and Urban (1997) that marketing was key in allowing ZANTAC to overcome the pioneer advantage of TAGAMET in the H_2-antagonist market. The major advantage of using micro-data is the ability to separately identify the effects of first-hand experience (which differs across doctors) from the aggregate effects of marketing.

Our finding that first-hand information weighs in most heavily in doctors' decision processes implies that there is little that manufacturers of new drugs can do to overcome the informational advantages enjoyed by incumbent producers except potentially for offering free samples which, from this point of view, represent a direct attempt by pharmaceutical companies to convince doctors to try the drug and accumulate the requisite first-hand experience. However, providing free samples constitutes only about 9% of total promotional expenditures in the American pharmaceuticals industry. In contrast, "detailing" activities (office visits) constitute over 55% (Schweitzer (1997), pp. 48-49).

On the other hand, the informational barriers to entry which we have characterized are not "malevolent" in any way, in the sense that incumbents do not strategically exploit informational advantages to deter entry into markets. These barriers to entry affect all firms equally upon their entry: presumably, the incumbents had to overcome similar informational hurdles when they first entered (or created, as the case may be) the market (Berndt, Bui, Reiley, and Urban (1997) documented how ZANTAC, the incumbent in our analysis, had to overcome similar hurdles when it entered the market).

Therefore, equilibrium market structure is the same as that which would arise in a market characterized by large fixed costs and free entry — while fewer products would exist in equilibrium relative to the perfect information case, there is no clear reason that the range of products would be suboptimal in any way[25], or that entry barriers are more severe for the pioneer than for later entrants. In other words, informational barriers to entry don't imply "path dependence". This contrasts with the recent literature on network externalities in which path dependence develops and the market can be "locked-in" to an equilibrium characterized by suboptimal products. Therefore, informational barriers to entry, while increasing fixed costs into an industry, need not imply dynamic market inefficiencies.

[25]Free entry in general may lead to a suboptimal *number* of products, as pointed out by Mankiw and Whinston (1985).

References

Ackerberg, D. (1996a): "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Investigation," mimeograph, Boston University.

——— (1996b): "Empirically Distinguishing Informative and Prestige Effects of Advertising," mimeograph, Boston University.

Amemiya, T. (1985): *Advanced Econometrics.* Harvard University Press.

Astra-Merck Inc. (1996): *1996 Annual Report.*

Bagwell, K. (1990): "Informational Product Differentiation as a Barrier to Entry," *International Journal of Industrial Organizaation*, 8(2), 207–223.

Bass, F., V. Mahajan, and E. Muller (1990): "New product diffusion models in marketing," *Journal of Marketing*, 54, 1–26.

Berndt, E., L. Bui, D. Reiley, and G. Urban (1997): "The Roles of Marketing, Product Quality and Price Competition in the Growth and Composition of the US Anti-Ulcer Drug Industry," in *The Economics of New Goods*, ed. by T. Bresnahan, and R. Gordon.

Berndt, E., and R. Pindyck (1998): "Network Externalities and Diffusion in Pharmaceutical Markets: Antiulcer Drugs," MIT Sloan School, mimeo.

Berry, S., J. Levinsohn, and A. Pakes (1997): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: Autos Again," Yale University, mimeo.

Bond, R. S., and D. Lean (1977): "Sales Promotion and Product Differentiation in Two Prescription Drug Markets," Discussion paper, Federal Trade Commission.

Coscelli, A. (1998a): "Are Market Shares in Drug Markets affected by Doctors' and Patients' Preferences for Brands?," University College, London.

——— (1998b): "Entry of New Drugs and Doctors' Prescriptions," University College, London.

Crawford, G., and M. Shum (1998): "An Empirical Model of Matching between Patients and Drugs," mimeo.

DeGroot, M. (1970): *Optimal Statistical Decisions.* McGraw-Hill Book Company.

Ellison, S., I. Cockburn, Z. Griliches, and J. Hausman (1997): "Characteristics of Demand for Pharmaceutical Products: An Examination of four Cephalosporins," *Rand Journal of Economics*, 28(3).

Graham, D. (1993): "Treatment of Peptic Ulcers caused by Helicobacter Pylori [edi torial; comment]," *New England Journal of Medicine*, 328(5), 349–50.

Mankiw, G., and M. Whinston (1985): "Free entry and social inefficiency," *Rand Journal of Economics.*

Medical Economics Co. (ed.) (1997): *Physicians' Desk Reference (Print ed.).* Oradell, N.J.

Pertile, M. (1995): "ZANTAC. Il Successo ... e' "un Qualcosa di Piu"'," M.Sc. Thesis, CUOA—Consorzio Universitario per gli studi di Organi zzazione Aziendale.

Schmalensee, R. (1982): "Product Differentiation Advantages of Pioneering Brands," *American Economic Review*, 72(3), 349–365.

Schweitzer, S. (1997): *Pharmaceutical Economics and Policy.* Oxford University Press.

Shapiro, C. (1982): "Advertising as a barrier to entry?," Federal Trade Commission Bureau of Economics Working Paper #74.

Shum, M. (1998): "Advertising and Competition with Habit Persistent Consumers: Estimates from a Probabilistic Nested Logit Model," University of Toronto, mimeo.

Stern, S. (1995): "The Demand for Pharmaceuticals," Stanford University mimeo.

Stern, S. (1997): "Simulation-Based Estimation," *Journal of Economic Literature*, 35(4), 2006–39.

Urban, G. L., T. Carter, S. Gaskin, and Z. Mucha (1986): "Market Share Rewards to Pioneering Brands: An Empirical Analysis and Strategic Implications," *Management Science*, 32(6), 645–659.

A Illustrative example of the learning model

To illustrate the model described in section 3.2.1, let us assume that $J = 4$, and that in period t, two signals are observed for the first two diagnoses, and one each for diagnoses 3 and 4 (in other words, we observe $\vec{\mu}_t \equiv (\mu_{1t1}, \mu_{1t2}, \mu_{2t1}, \mu_{2t2}, \mu_{3t1}, \mu_{4t1})$). Also assume that, conditioning on all the signals received up to (but not including) period t, the doctor believes that the vector $\vec{\delta}$ has the following distribution:

$$\begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix} \sim \left(\begin{bmatrix} E_t\delta_1 \\ E_t\delta_2 \\ E_t\delta_3 \\ E_t\delta_4 \end{bmatrix} \equiv E_t\vec{\delta}, \begin{bmatrix} \gamma_{1t}^2 & & & \\ \gamma_{12t} & \gamma_{2t}^2 & & \\ \gamma_{13t} & \gamma_{23t} & \gamma_{3t}^2 & \\ \gamma_{14t} & \gamma_{24t} & \gamma_{34t} & \gamma_{4t}^2 \end{bmatrix} \equiv \Sigma_{\delta t} \right). \tag{A.19}$$

Given these assumptions, then, in period t, doctor i believes that the vector of signals $\vec{\mu}_t$ has mean

$$\vec{\delta}_{\mu,t} \equiv \begin{bmatrix} E_t\delta_1 \\ E_t\delta_1 \\ E_t\delta_2 \\ E_t\delta_2 \\ E_t\delta_3 \\ E_t\delta_4 \end{bmatrix} \tag{A.20}$$

and variance-covariance matrix

$$\Sigma_{\mu,t} \equiv \begin{bmatrix} \gamma_{1t}^2+\rho_1^2\sigma_\theta^2+\sigma_{\eta 1}^2 & \gamma_{1t}^2+\rho_1^2\sigma_\theta^2 & \gamma_{12t}+\rho_1\rho_2\sigma_\theta^2 & \gamma_{12t}+\rho_1\rho_2\sigma_\theta^2 & \gamma_{13t}+\rho_1\rho_3\sigma_\theta^2 & \gamma_{14t}+\rho_1\rho_4\sigma_\theta^2 \\ & \gamma_{1t}^2+\rho_1^2\sigma_\theta^2+\sigma_{\eta 1}^2 & \gamma_{12t}+\rho_1\rho_2\sigma_\theta^2 & \gamma_{12t}+\rho_1\rho_2\sigma_\theta^2 & \gamma_{13t}+\rho_1\rho_3\sigma_\theta^2 & \gamma_{14t}+\rho_1\rho_4\sigma_\theta^2 \\ & & \gamma_{2t}^2+\rho_2^2\sigma_\theta^2+\sigma_{\eta 2}^2 & \gamma_{2t}^2+\rho_2^2\sigma_\theta^2 & \gamma_{23t}+\rho_2\rho_3\sigma_\theta^2 & \gamma_{24t}+\rho_2\rho_4\sigma_\theta^2 \\ & & & \gamma_{2t}^2+\rho_2^2\sigma_\theta^2+\sigma_{\eta 2}^2 & \gamma_{23t}+\rho_2\rho_3\sigma_\theta^2 & \gamma_{24t}+\rho_2\rho_4\sigma_\theta^2 \\ & & & & \gamma_{3t}^2+\rho_3^2\sigma_\theta^2+\sigma_{\eta 3}^2 & \gamma_{34t}+\rho_3\rho_4\sigma_\theta^2 \\ & & & & & \gamma_{4t}^2+\rho_4^2\sigma_\theta^2+\sigma_{\eta 4}^2 \end{bmatrix}. \tag{A.21}$$

B Classifying prescription observations to the four diagnoses

We followed several steps in classifying our prescription observations into the four diagnosis groups described in the main text. First of all, from our data, we distinguish between three mutually-exclusive patient types.

- **Heartburn patients** are those with <= 2 in-sample prescriptions (regardless of the quantity prescribed); all the observations associated with these patients are classified as Diagnosis 1.
- Patients with **pathological hypersecretory conditions** require large dosages of anti-ulcer medicine each month; we distinguish them in our data as those who are prescribed at least 133% of the average daily dose of anti-ulcer drugs in *every* month during which they are prescribed anti-ulcers drugs. The observations from these patients fall into diagnosis group 2.
- Patients who are not identified as either of the first two types are defined as suffering from either **GERD or peptic ulcers**.[26] Observations for these patients can be classified into one of two diagnosis groups, corresponding to "attack" (diagnosis 3) and "maintenance" (diagnosis 4) therapy. For observations belonging to these patients, we distinguish among *patient-months*: if the amount of the drug prescribed in a month is greater or equal than 133% of the average daily dose for maintenance in that month, we classify the patient to diagnosis 3 (attack therapy) *for that month*; otherwise, the patient is classified to diagnosis 4.

  During the sample period, (i.e., before the advent of antibiotic treatments for duodenal ulcers), the usual treatment would consist of an attack therapy, followed either by no prescription or maintenance. If there was no maintenance therapy, the proportion of recurrences was very high (70-80% according to Graham (1993)). A recurrence would be followed by a new attack therapy, and so on. Therefore, a typical patient suffering from either a peptic ulcer or GERD would cycle through periods of attack therapy, maintenance, no prescription, recurrence, attack, and so on — a regimen which we try to capture by our classification scheme.

[26] It would have been interesting to further distinguish between peptic ulcer and GERD, since *omeprazole* was proven in clinical trials to be more effective at treating GERD, but it is not possible to distinguish between the two in our data since the prescribed daily doses for a treatment regimen for either peptic ulcers or GERD is the same.

C Additional details regarding the data

Figure 7 presents the summary statistics for the variables used in the estimates. The first block summarizes the number of observations we have for each (doctor, month) cell. There is an obser-

Figure 7: **Summary Statistics**

	4 diagnoses cells (N=10,106: DOCTORS=326*31=T)				
Dependent variables–number of prescriptions in each (doctor,month) cell					
Variable	**# Obs.**	**Mean**	**St.Dev.**	**MIN**	**MAX**
MINOR HEART-BURNS					
omeprazole	10015	1.157	1.747	0	17
Total	10015	11.466	6.86	1	54
HYPERSECRETORY CONDITIONS					
omeprazole	5567	.706	1.613	0	26
Total	5567	5.342	6.373	1	70
ATTACK THERAPY					
omeprazole	8989	1.62	2.934	0	42
Total	8989	13.557	17.8	1	148
MAINTENANCE THERAPY					
omeprazole	10092	2.748	3.092	0	33
Total	10092	30.651	12.94	1	101
Doctors' Characteristics					
Herfindahl Brand level	326	.252	.068	.13	.564
Herfindahl Molecule level	326	.503	.113	.232	.805
Monthly quantity	326	798.37[a]	234.15	241	1682.08
Choice Characteristics					
Price diff.	31	1,188.724[b]	30.661	1131.17	1237.15
Ratio of ADV of *omeprazole*[c] *over all* the incumbent drugs	31	.427	.094	.247	.565

[a] The unit of measurement is *defined daily doses* (days of therapy).

[b] Italian Lire; 1US$=1,600 ItL; therefore, the average price difference between a day of treatment with *omeprazole*, and a day of treatment with the outside good is about 75 cents (this is approximately 25% of the price of a day of treatment with *omeprazole*).

[c] This data was provided by IMS Italy; the measurement unit is number of visits by the "detailers" during the quarter.

vation in a (doctor, month, diagnosis) cell if at least one patient receives the particular diagnosis by the doctor in that month. Therefore, the upper bound on the number of observations we have for any diagnosis is 10,106 (326 doctors × 31 months). The least frequent is diagnosis 2: in only 5,567 (doctor-months) combinations at least one patient was diagnosed with a pathological hypersecretory condition. The "Mean" column indicates the average number of patients who respectively received a prescription of *any* anti-ulcer drug (the *total* row), or a prescription of *omeprazole* across the (doctor-month) cells *conditional on at least one* prescription of anti-ulcer drugs taking place. For example, for diagnosis 3 (attack therapy): on average there are 1.62 prescriptions of *omeprazole*, and 13.55 prescriptions of anti-ulcer drugs, conditional on at least one prescription of anti-ulcer drugs by doctor i in month s.

Doctor-level covariates The second block summarizes the values for the covariates proxying for observed heterogeneity across doctors. They are: (i) the monthly average of herfindahl index[27] at *brand* level; (ii) the monthly average of the herfindahl index at the *molecule* level; and (iii) the average monthly quantity of anti-ulcer drugs prescribed. All of these are calculated based on doctors' observed prescription behavior *before omeprazole's entry.*

The first two variables — the two herfindahl indices — control for factors (either due to the doctor or to her patient pool) which lead her to prescribe a greater variety of drugs. Previous work on the Italian pharmaceutical market (Coscelli (1998b)) shows that when the entering drug is identical or very similar to existing ones, the higher these variables are (so the fewer the number of drugs that a doctor tends to prescribe to her patients), the less likely the doctor is to prescribe a new drug to her patients. These findings might not hold here, though, because *omeprazole* is a truly innovative drug, therefore its therapeutic quality vis-a-vis the quality of the competing brands becomes a crucial factor in the choice process. The third doctor-specific variable (average monthly quantity) controls for unobserved factors which make her prescribe more anti-ulcer drugs to her patients in general. We would expect that unconditionally doctors who prescribe a larger quantity of anti-ulcer drugs would be more likely to prescribe a new drug, but it is uncertain if this relationship holds when we condition in addition on doctors' willingness to prescribe across drugs using the two other doctor-specific variables described above.

Drug-level covariates The third block summarizes drug-specific covariates. *Omeprazole* had on average a 25% price mark-up over its competitors. [28] For our measure of advertising, we use the number of office visits made on behalf of each anti-ulcer molecule.[29] We have observations on the number of office visits in Italy made on behalf of each molecule on a quarterly basis; for the purposes of this model, we have created variables which are the *ratio* of the number of visits made on behalf of *omeprazole* versus visits made on behalf of all other molecules.[30] In addition, we use

[27]The herfindahl index is a standard measure of concentration. It is computed as the sum of the squares of the market shares. A high value of the herfindahl index indicates that the doctors are concentrated in their prescriptions across brands.

[28]Over the sample period, *omeprazole*'s price differential decreases, but since prices are tightly regulated in Italy, this decrease is due less to price decreases in the non-*omeprazole* molecules, than to shifts in the composition of the non-*omeprazole* "basket" towards cheaper therapeutic alternatives.

[29]This choice of number of visits, rather than other measures of marketing such as advertisements in journals or number of free samples distributed, was dictated to large extent by the data (from IMS Italy) — which for these other variables is spotty at best. Only for office visits do we have complete data over the sample period.

[30]From equation 3.5, the appropriate covariate to include in the utility specification should be ΔADV_t, the *difference* in advertising between *omeprazole* and the incumbent molecules in period t. However, this latter covariate moves very

a flow measure of advertising (ADV) as opposed to a stock measure. [31] The ratio of advertising of *omeprazole* over advertising for all the other anti-ulcer drugs available ranges between 24.7% and 56.5% during the sample period.

D Variance-covariance matrices

Figure 8: Estimated Initial Variance-Covariance Matrices

- For Model A:

$$10^{-3} * \begin{bmatrix} 0.2806 & 0 & 0 & 0 \\ 0 & 0.1663 & 0 & 0 \\ 0 & 0 & 0.2314 & 0 \\ 0 & 0 & 0 & 0.2463 \end{bmatrix}$$

- For Model B:

$$10^{-3} * \begin{bmatrix} 0.4335 & 0 & 0 & 0 \\ 0 & 0.2276 & 0 & 0 \\ 0 & 0 & 0.3588 & 0 \\ 0 & 0 & 0 & 0.5064 \end{bmatrix}$$

- For Model C:

$$10^{-3} * \begin{bmatrix} 0.1522 & 0 & 0 & 0 \\ 0 & 0.1679 & 0 & 0 \\ 0 & 0 & 0.1250 & 0 \\ 0 & 0 & 0 & 0.1733 \end{bmatrix}$$

erratically over time, with large swings within short intervals of time. We preferred the ratio because its variation was much less erratic, and also because it is unitless.

[31] Advertising stocks, however, will already be captured by the time effects $\xi_2, \ldots, \xi_{31}$.

Figure 9: Estimated Var-Covariance Matrices for the Signals

Hypothetical example: Two signals each for diagnoses 1 and 2; one signal each for diagnoses 3 and 4

- For Model A:

$$\begin{bmatrix} 0.0036 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0036 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0018 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0018 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0031 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0047 \end{bmatrix}$$

- For Model B:

$$\begin{bmatrix} 0.0053 & 0.0011 & 0.0001 & 0.0001 & 0.0013 & -0.0020 \\ 0.0011 & 0.0053 & 0.0001 & 0.0001 & 0.0013 & -0.0020 \\ 0.0001 & 0.0001 & 0.0022 & 0.0000 & 0.0001 & -0.0002 \\ 0.0001 & 0.0001 & 0.0000 & 0.0022 & 0.0001 & -0.0002 \\ 0.0013 & 0.0013 & 0.0001 & 0.0001 & 0.0053 & -0.0024 \\ -0.0020 & -0.0020 & -0.0002 & -0.0002 & -0.0024 & 0.0097 \end{bmatrix}$$

- For Model C:

$$\begin{bmatrix} 0.0012 & 0.0002 & 0.0001 & 0.0001 & 0.0003 & -0.0004 \\ 0.0002 & 0.0012 & 0.0001 & 0.0001 & 0.0003 & -0.0004 \\ 0.0001 & 0.0001 & 0.0014 & 0.0000 & 0.0001 & -0.0002 \\ 0.0001 & 0.0001 & 0.0000 & 0.0014 & 0.0001 & -0.0002 \\ 0.0003 & 0.0003 & 0.0001 & 0.0001 & 0.0010 & -0.0006 \\ -0.0004 & -0.0004 & -0.0002 & -0.0002 & -0.0006 & 0.0018 \end{bmatrix}$$