Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs after Patent Expiration

Ching, Andrew

Rotman School of Management, University of Toronto

16 February 2008

Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs After Patent Expiration

Andrew T. Ching*
Rotman School of Management
University of Toronto

First draft, April 2003

This draft, February 16, 2008

*I would like to thank Michael Keane and Thomas Holmes for their encouragement and advice. I would also like to thank Stephen Schondelmeyer, Gautam Gowrisankaran, Ig Horstmann, Chris Ferrall, Antonio Merlo, John Ham, Lung-Fei Lee, Iain Cockburn, and in particular, two anonynomous referees and Rob Porter for their helpful comments. I am grateful to Stephen Schondelmeyer for providing the IMS data, and to Fiona Scott Morton for providing her data on patent expiration dates. I am also grateful for comments received from participants in numerous seminars and conferences. Masakazu Ishihara provides excellent research assistance. I am responsible for all errors, omissions and interpretations. Address: Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON, CANADA M5S 3E6. Email: aching@rotman.utoronto.ca.

Abstract

This paper investigates whether aggregate consumer learning together with consumer heterogeneity in price sensitivity could explain why (i) there is a slow diffusion of generic drugs into the market, and (ii) brand-name originators keep increasing their prices over time even after the number of generic entrants has become fixed. To examine these hypotheses, I estimate a structural demand model that incorporates consumer learning and heterogeneity in price sensitivity. By conducting a counterfactual experiment, which eliminates the uncertainty of generics, I find that learning plays a role in explaining the slow diffusion. By simulating the model, I find that the branded pricing pattern could be explained by: (a) the diffusion rate of generics for price-sensitive patients is faster than that for price-insensitive patients, causing the proportion of price-insensitive patients faced by brand-name firms to slowly increase over time; (b) the brand-name price elasticities of demand (evaluated at the observed prices) are often less than one and increase over time, suggesting that brand-name firms may set their prices lower than what they would do if they were myopic, in order to slow down the learning process for generic qualities. But such an incentive diminishes over time as the uncertainty slowly resolves.

JEL Classification Numbers: C15, D12, D83, I11, L15, L65

1 Introduction

In 1984, the U.S. Congress passed the Waxman-Hatch Act which has significantly simplified the approval process for low-cost generic drugs.[1] Since then, many generic firms have produced generic versions of brand-name originals after their patents have expired. The large number of products that faced generic competition, at well-defined initial dates, has provided an unique opportunity to study the choice between brand-name drugs and generic drugs. In particular, it leads to two interesting observations on the prescription drugs market in the 80s: (i) there is a slow diffusion of generic drugs into the market, though generics typically cost 50 to 75 percent less than the brand-name originals, and (ii) many brand-name originators keep increasing their prices while generic firms reduce their prices over time, even after the number of generic entrants have been stabilized.

In this paper, I investigate whether aggregate consumer learning (i.e., learning from others) together with consumer heterogeneity in price sensitivity could explain the slow diffusion of generics and the branded pricing pattern. Conceivably, learning could influence price-sensitive and price-insensitive patients differently. In particular, the diffusion rate of generics for price-sensitive patients could be faster than that for price-insensitive patients, causing the proportion of price-insensitive patients faced by brand-name firms to slowly increase over time. In addition, brand-name firms may price their products lower than what they would do if they were myopic, in order to slow down the learning process for generic qualities. But such an incentive diminishes over time as the uncertainty slowly resolves. These could explain why brand-name firms increase their prices even after the number of generic entrants has become fixed over time.

To examine these hypotheses, I estimate the importance of aggregate consumer learning and the extent of heterogeneity in price sensitivity, using a structural demand model and aggregate market share data. Measuring the importance of these two features in demand for prescription drugs is of fundamental importance in policy implications. Consider, for example, the FDA expands their consumer education program to promote consumer confidence in generic drugs. Patients who are very price-insensitive may still choose to use the brand-name drugs, even if the education program improves their prior belief of generic qualities. If there is a large number of price-insensitive patients, such a campaign may have little impact on the overall generic substitution rate. A cost and benefit

[1] A generic drug is essentially an imitation of an original brand-name drug. When the patent protection on the original drug expires, other manufacturers can make copies and reproductions of the drug.

analysis of this kind of education program will depend critically on the importance of learning and the extent of consumer heterogeneity.

However, estimating a structural demand model with these two features using aggregate market share data is challenging. First, prices could be endogenous. As a result, if one simply measures the rate of learning by the change in generic market share over time, conditional on the price difference between the brand-name drug and generic drugs, the price coefficients might be biased towards zero. Even if there are good instruments available for prices, the complexity of the learning model has limited the application of recent techniques developed by Berry[6], Berry et al.[7]. Second, it is difficult to measure the heterogeneity in price sensitivity. In addition to income levels, another major factor that generates heterogeneity in price response is the diversity of insurance coverage for prescription drugs (Office of Technology Assessment[54]). An ideal way to capture heterogeneity in price sensitivity is to incorporate the structure of insurance coverage for prescription drugs into a structural model. However, the complicated coverage structure, which involves deductible, formulary, coinsurance rate, copayment, is hard to model. Moreover, to my knowledge, there is no publicly available datasets that describe details of prescription drug coverages in the U.S. population.

Given these practical hurdles, my approach is as follows. I first specify a structural demand model with consumer learning and heterogeneity in price sensitivity. In modeling the heterogeneity in price sensitivity, I allow for two types of patients who have different price coefficients in their indirect utility functions. I propose a new approach to address the potential endogeneity problem of prices. I estimate the demand model jointly with a pseudo-pricing policy function, which is a function of observed and unobserved state variables. Since some of the state variables are unobserved to the econometrician, I obtain parameter estimates by using simulated maximum likelihood. Although this method increases the number of parameters to be estimated (for the pseudo-pricing policy function), it is computationally feasible and does not impose strong assumptions about the process by which the pricing policy functions are formed. Using this framework and a data set detailing the evolution of prices and market shares for 14 drugs from 1984-1990, I estimate the

price coefficients, the proportion of each type of patient, and the structural learning parameters which describe how the demand side evaluates risks and perceived attribute levels of drugs.[2]

My estimation method relies on an important identification assumption: the timing of generic entry is exogenous. This assumption allows me to use the number of generic entrants as an exclusion restriction in the pseudo-pricing policy function. Although generic firms' entry decisions are endogenous, they usually do not have complete control on the exact timing of entry. Generic firms need to obain an approval from the FDA before they can sell the generic verison of a chemical. It is common that the FDA requires a generic firm to revise its application several times before approving it. The uncertainty about adopting the manufacturing process and how to meet the standard set by the FDA have introduced exogenous randomness in their entry time. This in turn generates exogenous variations in prices and choice sets, which help identify the price coefficients and the heterogeneity in price sensitivity, respectively. In section 2, I will discuss the FDA approval process in details.

My results are summarized as follows. I find evidence that learning plays a role in explaining the slow diffusion of generics. The estimates show that the price coefficients of the two segments of patients are very different: their ratios are 3 for heart disease drugs, and 4 for anti-depressants and anti-psychotic drugs. The proportion of price-sensitive patients is roughly 0.28 for drugs that mainly treat irregular heartbeat and chest pain, and 0.35 for the rest of the drugs. Patients are risk-averse and on average patients have pessimistic prior expectation about the quality of generics. Although I only observe aggregate market share data, I show that generic diffusion rates vary systematically across consumer types using the estimated model. In particular, I find that the diffusion rate of generics for price-sensitive consumers is much faster than that for price-insensitive consumers. Moreover, I find that the static brand-name price elasticities of demand, evaluated at the observed prices, are often less than one at the beginning and increase over time. This suggests

[2]An alternative way to estimate learning and heterogeneity is to use full-information maximum likelihood. This approach, which requires one to fully specify the demand and supply side, and an equilibrium, will take the endogeneity of price into account. Moreover, by explicitly incorporating the firms' behavior, this method should also help identify the extent of consumer heterogeneity. The caveat of this approach is that if consumer learning is important, it is likely that firms will be forward-looking. Modeling the supply side will then involve using a dynamic oligopoly model. Unfortunately, the computational burden of solving a stochastic version of such a model has hindered the application of full solution maximum likelihood.

that brand-name firms may be forward-looking and take learning into account when setting prices. The brand-name firms may try to price their products lower than what they would do if they were myopic, in order to slow down the learning process for the generic quality. The increase in the magnitude of the brand-name price elasticities over time may reflect that the uncertainty about the generic quality has been slowly resolved, causing the forward-looking incentive to diminish over time.

The rest of the paper is organized as follows. Section 2 provides some background on the U.S. pharmaceutical industry and discusses the related literature. In particular, it discusses the regulatory delay in approving generic drugs, and other competing hypotheses that could explain the slow diffusion of generics. Section 3 describes the demand model. Section 4 describes the data set and explains the estimation strategy. Section 5 presents the results. The last section concludes by discussing limitations of this study and directions for future research.

2 Background and Literature Review

2.1 Slow Diffusion of Generic drugs and Pricing Pattern

To illustrate the slow diffusion of generics, I consider the co-movements of average market share of generics and average relative price of generics. Figure 1 and 2 plots the average relative price of generics against time and the average market share of generics against time, respectively. I use my data set, which consists of 14 drugs, to obtain average market shares of generics and average relative prices of generics. The data is quarterly and period 0 refers to the first quarter that generics enter the market. Although generic drugs and brand-name drugs are made of the same chemicals and the average initial price of generics is 40% cheaper, the average initial market share of generics is only 0.05. It then slowly increases to 0.3 in period 5 even though the average relative price of generics remains fairly stable at around 0.6 for the first five periods.

I illustrate the pricing pattern of my sample (altogether 14 markets) in Figure 3, which plots the average wholesale price per patient day (AWP) for the brand-name drug and generic drugs against time. Price is measured at 1990 dollars. I also plot the number of generic manufacturers in that figure. It shows that brand-name prices increase after generic entry for all but one market. Haloperidol is the only market which experiences decline in brand-name prices after generic entry.

In contrast, average generic prices are consistently decreasing over time. Moreover, it should be emphasized that the increasing (decreasing) trend for brand-name (generic) prices continues even after the number of generic entrants has been stabilized.

2.2 Possible Explanations for the Slow Diffusion of Generics

There is evidence which suggests that learning with risk-aversion may be important in explaining observed diffusion patterns. Several studies surveyed opinions from physicians, pharmacists, and patients regarding the factors that determine their choices between a brand-name drug and its generic counterparts (e.g., Strutton et al.[52], Carroll and Wolfgang[10], and Mason and Bearden[38]). Their results indicate that the perceptions of the generic quality and related risk concerns, were important determinants for adopting generic drugs during the 80s. Although brand-name drugs and their generics use the same active ingredients, other ingredients such as tablet fillers, binders, coatings, or flavors may be different. The public (i.e., physicians, pharmacists and patients) may worry that these factors could affect the efficacy of generic drugs. The fact that brand-name drugs retain a substantial market share despite the large price differentials between the brand-name drugs and the generic drugs provides further support for the uncertainty hypothesis. This view is shared by other researchers (e.g., Caves et al.[11], Frank and Salkever[24], Griliches and Cockburn[27]). The "generic scandal" disclosed in 1989 also suggests that some generic drugs might not have been as good as brand-name drugs during the 80s.[3] Prior to the 1984 Waxman-Hatch Act, generic manufacturers were required to repeat all clinical tests, which were very costly. After the passage of the 1984 Act, the clinical tests requirements for approving generics drugs have been dropped and generic firms are only required to conduct test to show bioequivalence. As a result, the availability of generic drugs has become much more prevalent after 1984. The sudden surge in the availability of generic drugs, together with the new approval procedures, may lead patients, physicians and pharmacists to feel uncertain about the quality for generic drugs during 1984-1990, the period that my data set covers.

[3]Investigations by the U.S. Attorney's office during 1988-89 discovered that: (i) there were several cases of bribery in the generic drug approval process, (ii) some generic firms obtained the FDA approval for marketing new generic drugs by submitting false data, and (iii) some generic firms were found violating good manufacturing practices.

Another interesting feature of the prescription drug market is that learning from others seems to be relatively more important compared with other markets. Physicians or pharmacists, who are in contact with many patients, often talk to their colleagues in conferences or educational meetings. They may serve the function of information pooling. In addition to communications among physicians and pharmacists, there are institutions like the HMO and FDA's MedWatch, which keep track of the past experiences of a drug product and update the industry's perceived efficacy and safety of drug products.

Other than aggregate learning, there are three factors that could also explain the slow diffusion of generics: (i) learning about the existence of generics; (ii) the slow expansion in sales force or local sales offices; and (iii) delay in the production process. It might seem plausible that the first factor is relevant. However, pharmacists usually are quite knowledgeable about whether a generic is available, partly because generic firms market their products directly to pharmacists. Also, Medicaid's prescription drug program, which accounts for roughly 50% of third-party payments for prescription drugs in the 80s (Masson and Steiner [39]), sets reimbursement ceiling. Usually for drugs that have generics available, the ceilings are not high enough to cover the brand-name drug costs. Moreover, the retail dollar gross margin on the generic is typically higher than the brand-name (Masson and Steiner [39]). This should give them incentives to keep generics in stock as soon as they become available.[4] In addition, when the patent of a brand-name drug is going to expire, the news would usually be heavily reported by the media. Given these institutional details, it seems that the health care professionals should learn about the existence of generics for a particular chemical within a fairly short time frame. Therefore, I expect that this factor, if present, would affect the initial part of the diffusion process more. To control for this factor, I estimate one version of the model by withholding the first four periods of observations since the first generic enters the market. I will discuss the details of this econometric specification in section 5.1.

As for the second factor, there are certainly some new entrants in the generic drug business after the passage of the 1984 Waxman-Hatch Act. The incumbent generic firms have also entered many more markets which previously have prohibitively high entry fixed costs. Therefore, one

[4]It should be pointed out that most of the states allow pharmacists to substitute generics for brand-name drugs unless physicians explicitly prohibit it. In 1980, the only exceptions are Hawaii and Alaska. By 1989, all pharmacists are free to substitute generics as long as physicians allow it.

may conjecture that the slow expansion of sales forces or distribution channels could be another explanation for the slow diffusion. However, the costs of marketing generics are much lower than brand-name drugs (James[31]). In particular, generic firms do not rely on large sales force to visit physicians. Instead, they contact pharmacies directly, either in person or by telephone. Since quality, packaging and labeling of the products have already been approved by the FDA, what a pharmacy typically needs to know is the firm's product line, and the corresponding prices. Such information could be communicated over the phone, or obtained by referring to the company's catalog. This suggests that the nature of generic marketing is characterized by significant economies of scale. Even if a generic firm needs to increase its sales force, the cost of training new sales agents is relatively low, as they do not need to explain drugs' efficacy and side-effects profiles in details. Also, note that some generic firms have already been active for years prior to 1984, the year my dataset begins. These firms should already have their sales forces and local sales offices set up before the passage of the Waxman-Hatch Act.

Nevertheless, the second factor might still play a role in explaining the slow diffusion. Unfortunately, data on sales forces is proprietary and difficult to obtain. I therefore explore its implication on what I observe in the data. One implication of the slow expansion of sales forces is that there should be an overall improvement in the initial diffusion rates over time, assuming the same set of generic manufacturers are in the market during 1984-1990. Moreover, this implication could also be due to the general gain in confidence about generics drugs over time. This leads me to examine whether the initial diffusion rates of generics vary across markets with different patent expiration dates. As a robustness check, I also use data on the number of distributors to control for the slow expansion of distribution channels. I will explain in details how I control factor (ii) in section 5.1 and 5.6.

The third factor, though generally relevant, may not be very important in the pharmaceutical markets. In order to gain an approval from the FDA, a generic firm has to show the FDA that it has the requisite manufacturing capabilities. Moreover, pharmaceutical manufacturing processes usually have a high degree of automation. They use milling and micronizing machines to pulverize bulk chemicals into extremely fine particles. They then process these finished chemicals further in mixing machines. The mixed ingredients will be mechanically and automatically capsulated, pressed into tablets, or made into solutions (Bureau of Labor Statistics[43]). Although delay in

production is still possible because some machines could be occupied in manufacturing other drugs, it seems likely that the delay would only be a matter of months given the nature of mass production. But the observed slow diffusion in generics lasts for years. I therefore have decided not to take this factor into account in this study.[5]

2.3 Does Price Matter?

One may argue that due to the presence of health insurance, decision makers need not pay the cost of drugs, and hence they may not take prices into account when making choices. However, this claim may not be warranted in the U.S. Although the majority of the U.S. population has health insurance coverage, it is uncommon for insurance plans to cover the drug costs in full. For instance, Medicare does not provide any prescription drug coverage in the 80s. Moreover, HMOs and PPOs in the 80s are not as popular as the 90s or today. Most private health insurance providers had "major medical" plans with an overall annual deductible and some coinsurance rate applied to all covered services, including prescription expenses.[6] Hence, even if many argue that physicians do not have an incentive to learn drug prices, it seems plausible that a majority of patients, who are responsible for paying part of the prescription expenses, have an incentive to find out the brand-generic price differential, although the insured patients' price sensitivity should be significantly lower than the uninsured.

2.4 Timing of Generic Entry

The crucial identification assumption of exogenous timing of generic entry deserves more explanation. In most industries, entry decisions of firms are endogenous and so is the timing of entry. But in the US pharmaceutical industry, generic firms usually do not have complete control over the entry time, even though the entry decisions are endogenous (Scott Morton [48], Ching [13]). To enter a market, a generic firm needs to submit an application to the FDA. The application is

[5]To the best of my knowledge, none of the structural modeling papers that study the diffusion of new prescription drugs take this factor into account.

[6]In 1989, about 70% of the non-elderly population had private health insurance coverage. Among those, about 61% had an overall annual deductible and some coinsurance rate applied to prescription expenses. The rest usually required a fixed copayment for prescription drugs instead of including them in the overall deductible (Office of Technology Assessment[54]).

called an Abbreviated New Drug Applications (ANDA). In order to obtain an approval, a generic firm has to prove that its product contains the same active ingredients, strength, dosage form, route and is bioequivalent. The time it takes to adopt the manufacturing technology and obtain approval from the FDA is quite uncertain from generic firms' point of view.[7] Depending on the formulation of the drug, the resource constraint, the experience of the firm, and the availability of raw materials, it could take up to a few years for a generic firm to adopt the manufacturing technology. Although the approval process has been simplified after 1984, it is still quite involved. It includes bioequivalence review, chemistry/microbiology/labeling review, plant inspection, and independent laboratory tests of preliminary batches of the product. Meadows[41] reports that "(after the initial ANDA application is submitted) it takes more than 20 months on average for a new generic drug to be approved by the FDA, and it usually involves multiple review cycles. Only about 7 percent of applications are approved on the first cycle and about a third are approved on the second cycle." This is consistent with my data set. As shown in Figure 3, instead of entering the market immediately after the patent expiration, generic firms enter at different points in time in most of the markets.

Withdrawal of an ANDA before its final approval seems unlikely as the cost of preparing for an initial ANDA is sunk, and it is a major cost component of producing a generic drug (Scott Morton [48]). The marginal cost of production is very low. Therefore, generic firms usually enter the market as soon as they receive the approval. Most importantly, the uncertainty of adopting manufacturing process and the FDA's approval process appear to have little direct relationships with the unobserved characteristics in the demand model. Hence, it seems plausible that the timing of generic entry is relatively exogenous.

2.5 Literature Review

There is a growing interest in modeling the demand for prescription drugs. Stern[51] estimates a two-level nested logit model using aggregate market share level data from four therapeutic classes (Minor Tranquilizers, Gout, Oral Diabetics and Sedatives), where consumers choose among drugs of the same therapeutic class at the first level, and then choose between a brand-name drug and generics at the second level. Ellison, Cockburn, Griliches and Hausman[21] estimate an Almost

[7] This view is also shared by Caves et al.[11].

Ideal Demand System using product level data on four anti-infective drugs. Berndt, Bui, Reiley and Urban[3] estimate the effect of detailing and journal advertising in the US anti-ulcer drugs market. Hellerstein[30] estimates a physician's prescription choice model using individual level data. All these studies ignore learning, or more generally speaking, state dependence in demand.[8] However, if state dependence is present, estimating a model without it could potentially lead to bias in the estimates and give misleading policy implications (Heckman[28]). For example, if aggregate consumer learning is important, a price promotion will not only affect the quantity sold in the current period, but also have a long-term impact on demand. A demand model without state dependence will not be able to predict such a long-term effect.

In the context of the choice between brand-name drugs and their generic counterparts, learning could also play an important role in explaining the branded pricing pattern. Caves et al.[11], Grabowski and Vernon[26], and Frank and Salkever[24]) argue that the increase in branded prices after generic entry is due to consumer heterogeneity in price-sensitivity. Their explanation is that some price-sensitive consumers switch to generics, causing the average price sensitivity of demand faced by the brand-name firms to decrease. Although this argument explains why branded prices increase right upon generic entry, it cannot explain why they continue to increase *even after the number of generic entrants has been stabilized.* As far as I know, the literature has not provided any explanations about this stylized fact yet. In this paper, I argue that learning, together with consumer heterogeneity in price-sensitivity, could explain the branded pricing puzzle. I argue that learning causes price-sensitive consumers slowly switch from brand-name drugs to generic drugs. As a result, the average price-sensitivity of consumers faced by the brand-name firms gradually decreases over time. In addition, brand-name firms have an incentive to charge prices lower than what they would do if they were myopic, in order to slow down the learning process. But such a dynamic incentive becomes smaller over time as the uncertainty about the generic quality slowly resolves. This could explain why the brand-name firms keep raising their prices over time even after the number of generic entrants has become fixed.

There are three other studies that are closely related to mine. Crawford and Shum[16] estimate a forward-looking learning model for the Italian anti-ulcer drug market. Using the same

[8]State dependence in demand refers to any causal relationships between past purchasing behavior and current purchasing behavior. Consumer learning could be one of these causal relationships.

data set, Coscelli and Shum[15] estimate a myopic learning model for omeprazole (one of the anti-ulcer drugs), where they allow for informational spillovers across all patients of a given physician. The focus of Coscelli and Shum[15] and Crawford and Shum[16] is to study how physicians choose a brand-name drug that matches their patients' need. Berndt, Pindyck and Azoulay[4] estimate a reduced form diffusion model for the US anti-ulcer drug market, and focus on measuring the diffusion rates of new brand-name drugs. Unlike my paper, none of them studies the role of learning from other physicians, and consumer heterogeneity in price sensitivity, in the choice between the branded drugs and their generic counterparts.[9]

3 The Model

I now turn to describe my model in details. I extend the individual Bayesian learning demand model developed by Erdem and Keane[22] to an aggregate learning model where consumer preferences are allowed to be heterogeneous with respect to their price-sensitivity.

Product characteristics can be distinguished as p_j, A_j, and ξ_j, where p_j is the price of product j, A_j is the mean attribute level of product j, and ξ_j represents an unobserved demand shock for product j (e.g., media coverage, promotional campaigns by brand-name firms or the FDA/insurance companies, etc.). Physicians/patients in the model are perfectly informed about p_j and ξ_j, but are imperfectly informed about A_j.

At the beginning of each period, patients make their choices based on the public perception about the quality of each product. After taking their drugs, some patients reveal their experience signals to the public when revisiting their physicians. Then physicians, who act as an information aggregator, update the public information of the mean attribute for each product in a Bayesian fashion.[10] The model can be broken up into two components: (1) learning about product attributes, and (2) demand. I now describe them in turn.

[9] A work-in-progress by Currie and Park[17] also estimate a Bayesian learning model for anti-depressant drugs using aggregate market share data in the U.S. In contrast to my study, they neither allow for consumer heterogeneity in price sensitivity, nor attempt to control for the endogeneity problem of prices. A priori, it is not clear whether their exogeneity assumption of prices is justified because pharmaceutical firms are free to set their prices in the U.S.

[10] As discussed in the previous section, this is motivated by the feature of learning from others in prescription drug markets.

3.1 Learning about Product Attributes

A drug is an experience good. Each patient i's experience of the attribute of product j at time t ($\tilde{A}_{ijt}$) may differ from its mean attribute level A_j. The difference between $\tilde{A}_{ijt}$ and A_j could be due to the idiosyncrasies across people in reacting to drugs. For instance, when different patients take the same pain-relief drug, the time that they need to wait before their headache disappears may vary, simply because they have different severity of illness. Even when a patient takes the same drug at different points in time, the waiting time may still change, as body conditions may vary (it may depend on how much sleep one had, how much one ate, and what other drugs one is concurrently taking, etc.). In addition, there might be some intrinsic quality differences on the production side – different batches of drugs may have different qualities.

Let n_g be the number of generic entrants. The actual experience of consuming a generic drug $j = 1, \cdots, n_g$ is assumed to be as follows,

$$\tilde{A}_{ijt} = A_g + \delta_{ijt}, \tag{1}$$

where $\tilde{A}_{ijt}$ is the experience signal that patient i receives when consuming generic drug j at time t. The noisy error term, δ_{ijt}, is treated as an $i.i.d.$ random variable with zero mean.[11] Similarly, the actual experience of consuming a brand-name drug is assumed to be as follows,

$$\tilde{A}_{ibt} = A_b + \delta_{ibt}. \tag{2}$$

The initial period of my model ($t = 0$) is the period right before the patent expires. Since brand-name products have typically been on the market for six to ten years by the time their patents expire, I assume that the public has already accumulated a sufficient number of experience signals to infer their true mean attribute levels. Hence, the public is only uncertain about mean attribute levels of generic drugs in my model.

In order to facilitate the construction of Bayesian updating rules, the signal noise δ_{ijt} and the initial prior on A_g are assumed to be normally distributed. Letting $t = 0$ be the initial period of the model, I have

$$\delta_{ijt} \sim N(0, \sigma_\delta^2), \tag{3}$$

[11] Here I assume that $A_j = A_g, \forall j$. I make this assumption because physicians/patients usually cannot choose which manufacturer's generic drug to be dispensed.

$$A_g \quad \sim \quad N(A, \sigma^2_{A_g}(0)), \tag{4}$$

where $\sigma^2_{A_g}(0)$ is the initial variance (at $t = 0$) or uncertainty about A_g. According to (3) and (4), when a generic drug is first introduced, the initial prior is that its mean attribute level (A_g) is normally distributed with initial prior mean A and initial prior variance $\sigma^2_{A_g}(0)$. Thus, letting $I(0)$ denote the initial prior information about generic drugs, $E[A_g|I(0)] = A$.

Let S_t be the set of experience signals that are revealed to physicians at time t. Since not every patient revisits his/her physician, the cardinality of S_t ($card(S_t)$) is generally smaller than the quantity of generics consumed at time t (q_{gt}), which is the total number of experience signals revealed to patients. Let κ be the fraction of experience signals revealed to physicians in each period. Then $card(S_t) = \kappa q_{gt}$. Intuitively, one can interpret κ as the probability that a patient revisits a physician and discuss his/her experiences about generics.

Physicians as a whole use information revealed to them over time (i.e., S_t) to update the prior expectation of A_g. The updating of the public information set will not occur until the end of the period (i.e., until the experience signals are revealed in that period). Let $\bar{A}_{gt}$ be the sample mean of the set of experience signals revealed to physicians at time t.[12] Then according to the Bayesian rule (DeGroot[18]):

$$E[A_g|I(t+1)] = E[A_g|I(t)] + \beta_g(t)(\bar{A}_{gt} - E[A_g|I(t)]), \tag{5}$$

where $\beta_g(t)$ is a Kalman gain coefficient, which is a function of experience variability (σ_δ), perceived variance ($\sigma^2_{A_g}(t)$). The Kalman gain coefficient can be expressed as:

$$\beta_g(t) = \frac{\sigma^2_{A_g}(t)}{\sigma^2_{A_g}(t) + \frac{\sigma^2_\delta}{\kappa q_{gt}}}. \tag{6}$$

The β_g coefficient can be interpreted as the weight associated with the new information when updating the prior expectation of A_g. Each time $\sigma^2_{A_g}(t)$ is updated, the β_g coefficient will be updated accordingly.

The perception variance at the beginning of time $t+1$ is given by (DeGroot[18]):

$$\sigma^2_{A_g}(t+1) = \frac{1}{\frac{1}{\sigma^2_{A_g}(0)} + \frac{\kappa Q_{gt}}{\sigma^2_\delta}}, \tag{7}$$

[12] Let A_g be the true mean attribute level of generic drugs. Then, $\bar{A}_{gt}|(\kappa q_{gt}, I(t)) \sim N(A_g, \frac{\sigma^2_\delta}{\kappa q_{gt}})$.

where $Q_{gt}(=\sum_{\tau=1}^{t} q_{g\tau})$ is the cumulative consumption of generics, or,

$$\sigma^2_{A_g}(t+1) = \frac{1}{\frac{1}{\sigma^2_{A_g}(t)} + \frac{\kappa q_{gt}}{\sigma^2_\delta}}. \tag{8}$$

Equations (7) and (8) suggest that the perceived variance associated with A_g (and consequently the perceived variance of $\tilde{A}_{ij}$) will be lower, ceteris paribus: (a) the lower the experience variability of the product; and (b) the more experiences the public has with generic drugs.[13] Also, it should be noted that it is the ratio between the experience variability (σ^2_δ) and the fraction of experience signals revealed (κ) that determines the rate of learning.

3.2 Demand

The demand for prescription drugs is complex. The principal-agent relationship among patients, physicians and pharmacists certainly plays an important role in determining demand in this market. However, with only product level data (i.e. prices, quantities and measurable characteristics of the products) available, it would be very difficult, if not impossible, to identify the parameters of a demand model with multiple decision makers. Thus, my demand model abstracts away from this multiple-decision making process. Coscelli and Shum[15] and Crawford and Shum[16] also use a similar approach.

The demand system is obtained by aggregating a discrete choice model of individual patient behavior. In each market, a patient's choice is modeled as a two-stage nested process. The choice set J is partitioned into subsets J_l, where $l \in \{0, b, g\}$. The choice set J_0 only consists of an "outside" alternative, J_b only consists of the brand-name drug (b), and J_g consists of the generic drugs $(1, \cdots, n_g)$. Patients select the subset J_l first, then they select an alternative in that subset, in each of T discrete periods of time, where T is finite.[14]

[13] It is implicitly assumed that patients, who do not have their experience signals revealed to the public, will not use their own experience signals in updating their priors. Notice that the initial slow diffusion of generic sales exhibited in the data suggests that learning is a slow process. Since the sales of generics is at least in the thousands of patient days per quarter, the normalized experience variability (σ_δ/κ) will need to be fairly large if learning takes time. This implies that the marginal contribution of a single experience signal to the information set will be very small. Hence, including a patient's own signal should not make much difference in the updating process, but makes the state space much more complex.

[14] It might seem implausible that a patient can choose among generics in J_g, as this is largely a decision by pharmacists. But in a two-stage nested process, without loss of generality one can interpret that patients are

Alternatives are defined to be mutually exclusive, so that if $d_{ij}(t) = 1$ indicates that alternative j is chosen by patient i at time t and $d_{ij}(t) = 0$ indicates otherwise, then $\sum_{j \in J} d_{ij}(t) = 1$. It should be noted that the outside alternative includes receiving no treatment and other non-bioequivalence drugs, which could treat the same disease.

Let $I(t)$ denote the public information set at the beginning of time t. Patients gain access to $I(t)$ through physicians. Associated with each choice j at time t is a current period expected utility $E[U_{ij}(t)|I(t)]$, where $E[.]$ is the mathematical expectation operator. The expected utility is known to each patient at time t. When patient i makes his/her purchase decision, his/her objective is to maximize current period expected utility:

$$E[\sum_{j \in J} U_{ij}(t) d_{ij}(t) | I(t)]. \tag{9}$$

It is plausible that patients recognize that current choices may affect the public information set. As a result, they may have an incentive to experiment with new products to learn their true mean attributes. If such an incentive is strong, it may be more reasonable to model patients as maximizing their lifetime expected utility rather than their current expected utility. However, in the context of purchasing pharmaceuticals, some illnesses are very short-term and happen relatively infrequently during one's lifetime. In those cases, it seems plausible to assume that the incentive to experiment is small. Even for a long-term illness, an individual patient's incentive to try generic drugs will be significantly weakened if the normalized experience variability (σ_δ / κ) is large, again because the marginal contribution of a single experience signal to the information set will be very small.[15] As argued above, the slow diffusion of generic sales suggests that this would be the case for pharmaceutical markets. Hence, the assumption of maximizing expected current utility seems to be a reasonable approach.

I assume that the indirect utility of consuming a drug can be adequately approximated by an additive compensatory multi-attribute utility model (Lancaster[33]), and is given by the following

randomly assigned a generic drug j in the second stage. This captures the idea that patients do not know which generic drug they will receive when filling their prescriptions.

[15] It should be pointed out that there is an externality problem in the learning process. An individual patient does not take into account the spillover benefit of his/her experience signals to other patients. Since the total number of patients for any particular illness is typically very large (over a million), it may be socially optimal for the economy to experiment with generic drugs even though the normalized experience variability is large from an individual viewpoint.

expression:

$$U_{ijt} = -\alpha_i p_{jt} + \omega \tilde{A}_{ijt} - \omega r \tilde{A}_{ijt}^2 + \xi_{jt} + \zeta_{ilt} + e_{ijt}, \tag{10}$$

where U_{ijt} is the utility for patient i conditional on choice of product j at time t; p_{jt} is the price for product j at time t; ω is patients' attribute weight on $\tilde{A}$; r is the risk coefficient; α_i is the utility weight that patient i attaches to price; ξ_{jt} represents the mean valuation of product j's unobserved demand shock at time t; $(\zeta_{ilt} + e_{ijt})$ represents the distribution of consumer preferences about this mean. The parameters α_i, ζ_{ilt} and e_{ijt} are unobserved by the econometrician but observed by patients in the model when they make purchase decisions. It should be noted that $\tilde{A}_{ijt}$ is not observed by patients when they make their purchase decisions. It is observed by them only after they consume the drug, but remains unobserved by the econometrician. Therefore utility is a function of experienced attribute levels ($\tilde{A}_{ijt}$) and not the true mean attribute levels (A_j).

The diversity of insurance coverage of prescription drugs would likely translate to heterogeneity in price response when patients/physicians jointly decide between the brand-name drug or its generic counterparts. I therefore allow α_i to be heterogeneous in order to capture this institutional feature. It should be noted that ω and r are assumed to be homogeneous. I make this assumption because it is very difficult, if not impossible, to identify the parameters of the model if I allow all three coefficients, (α, ω, r), to be heterogeneous given the market share level data I have.

For each patient i, ζ_{ilt} is common to all products in group l. This introduces group correlation of utility levels. In the nested logit framework (Cardell[8]), e_{ijt} is distributed Extreme Value with variance $(\pi\mu_2)^2/3$, and $(\zeta_{ilt} + e_{ijt})$ is distributed Extreme Value with variance $(\pi\mu_1)^2/3$. One interpretation is that conditioning on choosing generics, e_{ijt} is an error term associated with generic drug j.

It is assumed that agents in the model can measure drug attributes according to a fixed scale, e.g., a patient can measure attributes such as how long his stomach pain would be suppressed after taking the drug.[16] Hence, one can represent patients' risk-averse behavior with respect to $\tilde{A}$ by using the concavity of the utility function. As I argued above, risk-averse behavior could play an important role in explaining the slow diffusion of generics observed in the data. I therefore

[16]Obviously, drug attributes are multi-dimensional. Implicitly, I assume patients are able to use a scoring rule to map all measurable attributes to a one-dimensional index. It is the value of this one-dimensional index that enters the utility function.

follow Erdem and Keane[22] and allow a quadratic term in $\tilde{A}$ to enter the utility function. Given a strictly positive ω, the patients are risk averse, risk neutral or risk seeking as $r > 0$, $r = 0$ or $r < 0$, respectively with respect to $\tilde{A}$.

It follows from Equation (10) that the expected utility associated with generic drug j is,

$$\begin{aligned} E[U_{ijt}|I(t)] &= -\alpha_i p_{jt} + \omega E[\tilde{A}_{ijt}|I(t)] - \omega r E[\tilde{A}_{ijt}|I(t)]^2 \\ &\quad -\omega r E[(\tilde{A}_{ijt} - E[\tilde{A}_{ijt}|I(t)])^2|I(t)] + \xi_{gt} + \zeta_{igt} + e_{ijt}. \end{aligned} \tag{11}$$

Patient i's expected utility of purchasing generic drug j at time t, given his/her perception at the beginning of time t, is a linear function of price, a concave ($r > 0$), linear ($r = 0$) or convex ($r < 0$) function of the expected levels of $\tilde{A}_{ijt}$, and a linear function of the perceived "variance" in $\tilde{A}_{ijt}$. Furthermore, the stochastic components of the utility function ($\xi_{jt}, \zeta_{igt}, e_{ijt}$) reappear in the expected utility equation because they are stochastic only from the econometrician's point of view.

Now note that in Equation (11), the term $E[(\tilde{A}_{ijt} - E[\tilde{A}_{ijt}|I(t)])^2|I(t)]$ can be decomposed into $\sigma_\delta^2 + \sigma_{A_g}^2(t)$ (see (1)), and $E[\tilde{A}_{ijt}|I(t)] = E[A_g|I(t)], \forall i, \forall j \in J_g$ because δ_{ijt} has zero mean (see (1)). I can rewrite Equation (11) as follows,

$$\begin{aligned} E[U_{ijt}|I(t)] &= -\alpha_i p_{jt} + \omega E[A_g|I(t)] - \omega r E[A_g|I(t)]^2 - \omega r(\sigma_\delta^2 \\ &\quad +\sigma_{A_g}^2(t)) + \xi_{gt} + \zeta_{igt} + e_{ijt}. \end{aligned} \tag{12}$$

Since I assume that the public has already learned perfectly about the true mean attribute level of the brand-name drug, A_b (i.e. $\sigma_{A_b}(t) = 0$ and $E[A_b|I(t)] = A_b, \forall t = 0, \cdots, T$), it follows from Equation (12) that the expected utility of purchasing a brand-name drug can be written as,

$$E[U_{ibt}|I(t)] = -\alpha_i p_{bt} + \omega A_b - \omega r A_b^2 - \omega r \sigma_\delta^2 + \xi_{bt} + \zeta_{ibt} + e_{ibt}. \tag{13}$$

Equations (10)-(11) apply only to the drugs under analysis. In each period, patients may also choose an outside alternative (i.e. other non-bioequivalent drugs or no treatment). I assume the expected utility associated with the outside alternative to be a linear function of time plus a stochastic error component,

$$E[U_{i0ts}|I(t)] = \phi_0 + \phi_t t + \zeta_{i0t} + e_{i0t}. \tag{14}$$

The time trend is meant to capture the possibility that the quality of alternative treatment may be improving over time (or their cost may be dropping) in a reduced form way.

As in Heckman and Singer[29], I specify the heterogeneity of the price response coefficient (α_i) follows a discrete multinomial distribution. Accordingly, we distinguish between K different "types" of individuals. The population proportion of each type is given by $\pi_k = Pr(\alpha_i = \alpha^k)$, where each type $k = 1, ..., K$. The expected demand for each product can be derived in a straightforward way.

As pointed out in Berry and Pakes[5] and Ackerberg and Rysman[1], the *i.i.d.* extreme value error terms (e_{ijt}'s) represent unobserved product differentiation that is symmetric across products.[17] The unobserved product differentiation could be due to the uncertainty about quality differences among individual generic drugs, which I do not model explicitly. This feature of the model has caused the price-cost margin to be strictly bounded away from zero even when the number of generics increases to infinity. The reason for this result is that each additional generic entrant creates one more dimension to the symmetric unobserved product differentiation (SUPD) space. Moreover, the higher the variance of e_{ijt}, the larger the bound as it increases market power of each product. Intuitively, μ_2, which measures the variance of e_{ijt}, represents the degree of SUPD. As I will discuss in Section 5.2.2, the price of generics consistently decreases over time even when the number of generic entrants becomes fixed. This suggests that the degree of SUPD may decrease over time. This could happen if the uncertainty about qualities of individual generic drugs is resolved over time. To capture this, I model μ_2 as a function of time since the first generic entry, g_t,

$$\mu_2(g_t) = \bar{\mu}_2 exp(-\iota g_t). \tag{15}$$

This approach is similar to Ackerberg and Rysman[1]. In this parameterization, I allow the possibility that μ_2 may decrease over time. As demonstrated in Ching[13], this feature has significantly improved the flexibility of a supply side model in generating pricing patterns that mimics the data.

[17] Note that $E[A_g|I(t)]$ is also an unobserved product characteristic but it is asymmetric across products.

4 Data and Estimation

4.1 Data

A drug is defined as a chemical or a combination of chemicals that is patented by its originator. It can be produced by either the originator or generic firms after patent expiration. My sample consists of 14 drugs with patents expired during the four year period from 1984 through 1987: 7 of them are heart disease drugs, 4 are anti-depressants, and 3 are anti-psychotic drugs.[18] Data sources for this study include: IMS America,[19] the pharmaceutical Manufacturers Association (PMA) the Food and Drug Administration (FDA), and the Statistical Abstract of the United States. Table 1 shows the summary statistics of my data set.

Data on sales revenue and quantities sold, and the number of distributors are obtained from the IMS U.S. Drugstore (USD) and U.S. Hospital (USH) database. For each drug, I observe quarterly revenue and quantity sold for both the brand-name original and the total sales of its generic counterparts from the quarter that the patent expired to the fourth quarter of 1990. Observations in this data set represent combined sales from drugstores and hospitals. Prices used in this study is the average wholesale price (AWP) which is obtained by dividing revenue by quantity sold. Certainly, retail transaction prices will be more accurate. Unfortunately, the transaction prices data is very difficult to obtain, and not available to me. One limitation of using AWP is that it does not take manufacturer rebate into account. A detailed discussion on IMS data collection process can be found in Berndt et al.[3]. Data on number of HMO enrollment is obtained from the United States Statistical Abstract.

The patent expiration dates are obtained from the FDA and the PMA's Report of Patents on Medical Products. The number of generic manufacturers and their approval dates for Abbreviated New Drug Applications (ANDA) for marketing generic drugs are obtained from the FDA's Orange Book. Daily Defined Dose (DDD) and Average Treatment Duration (ATD) are collected from the

[18]The data set described here is a subset of drugs used in Suh et al.[53]. The data on sales volume, revenue and patent expiration date were originally collected by Stephen Schondelmeyer on behalf of the U.S. Office of Technology Assessment. I also obtained a data set on patent expiration dates from Fiona Scott Morton. I used her data to cross check the information that I collected from other sources.

[19]IMS America is a marketing research company that specializes in collecting sales data for the pharmaceutical industry.

Medispan's Price-Trek database. DDD is used to standardize the unit to the number of patient days. ATD is used to obtain the number of patient days that on average each purchase decision would amount to.

The estimates of the number of patients who have been diagnosed with a particular condition are obtained from National Ambulatory Medical Care Survey and the National Hospital Discharge Survey. As in Stern[51], for each disease category, I use data from the National Ambulatory Medical Care Survey (NAMCS) and the National Hospital Discharge Survey (NHDS) to obtain an estimate of the total number of individuals who were diagnosed with a particular condition by a physician or a hospital in a particular year. I then obtain the mean total number of patients by averaging the total number of patients over years. The total size of the market (in number of patient days) is taken to be the ATD within the category multiplied by the mean total number of patients.

4.2 Estimation Method

A common concern of estimating this class of product differentiated market models is the potential endogeneity of price. If firms know unobserved product characteristics ($E[A_g|I(t)]$ and ξ_t) when they choose prices, it is likely that prices are correlated with them. If this correlation exists and the econometrician ignores it when estimating the model, not only will the price coefficient be biased, but so will the other preference parameters that determine the rate of learning.

Berry[6] and Berry et al.[7] (BLP) have developed a GMM based method to account for this endogeneity problem. To apply their procedure, one would first use a contraction mapping to recover each product's "aggregate" unobserved product characteristic (denoted as ε_{jt}) from the market shares and a given set of parameter values, and then use ε_{jt}'s to create the sample analog for the moment conditions. Unfortunately, ε_{jt}, being a function of $E[A_j|I(t)]$, $E[A_j|I(t)]^2$ and ξ_{jt}, is serially correlated and non-stationary in general.[20] Another complication is that when constructing the moment conditions, one needs to compute the mean of $E[A_g|I(t)]$ and $E[A_g|I(t)]^2$ conditioning on $\{q_{g\tau}\}_{\tau=0}^{t-1}$, which is quite computationally burdensome. These issues make it difficult to use GMM to estimate this model.

[20]To my knowledge, all the discrete choice product differentiation models, which are estimated using the BLP method, assume that there is only one unobserved product characteristic for each product (i.e., ξ_{jt}).

4.2.1 Maximum Likelihood: Approximation Approach

Instead of using the BLP procedure to estimate this model, I develop another estimation approach. To understand the contribution of my method, it would be useful to review the classical full information maximum likelihood approach (FIML). In FIML, the econometrician needs to model the oligopolistic supply side explicitly, and derive a pricing policy rule as a function of observed and unobserved product characteristics, and other state variables. The econometrician then forms the joint likelihood function of a sequence of prices and quantities, and consistent estimates of the parameters can be obtained by maximizing the likelihood function. FIML is an iterative process, which requires solving numerically the supply-side oligopoly model for a given set of parameter values, then evaluating the likelihood function, etc., until the likelihood is maximized. However, as the demand side involves learning and firms may be forward-looking, the full solution of the oligopoly model involves solving a multi-agent dynamic programming problem, which is very computationally demanding. For the dynamic oligopoly model of the pharmaceutical industry that is detailed in Ching[13], a single solution takes roughly 12 hours of cpu time on a Intel Pentium D 3.00GHz processor workstation. Hence, full information maximum likelihood is infeasible in this context. In addition, even if the econometrician has the computation power to apply FIML, biased estimates may still result if the equilibrium model is misspecified. In particular, a priori it is not clear whether firms choose price after observing $E[A_j|I(t)]$ and ξ_{jt}.

For these reasons, instead of generating a pricing policy function by solving a supply-side model explicitly, I approximate the pricing policy function. What state variables should enter the pricing policy function? As explained above, $E[A_g|I(t)]$ and ξ_t might be correlated with p_t, where $p_t = (p_{bt}, p_{gt})$. In addition, p_{jt} might also depend on $(\sigma^2_{A_g}(t), n_{gt}, t)$ through the dynamic oligopolistic equilibrium (recall that n_{gt} is the number of generic entrants at time t). The time trend, t, may affect equilibrium prices because it enters the utility function for the outside good. A time trend in the pricing policy function could also capture some systematic increase in production costs over time. Hence, the true pricing policy function, $\wp(.)$, should be a function of $(n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_{bt}, \xi_{gt})$. For $j \in \{b, g\}$,

$$p_{jt} = \wp_j(n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_{bt}, \xi_{gt})\nu_{jt}, \tag{16}$$

where ν is an error term, which captures productivity shocks, or "optimization" errors that prevent the firm from correctly implementing the optimal pricing policy function, $\wp_j(.)$. Implicitly, I assume that firms know that there are random factors that lead to ex post discrepancies between intended and realized decisions, and $\wp_j(.)$ has already taken these uncertainties into account.

Taking logs on both sides of Equation (16), I obtain,

$$log(p_{jt}) = log(\wp_j(n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_{bt}, \xi_{gt})) + log(\nu_{jt}). \tag{17}$$

To approximate $log(\wp_j(.))$, I propose to use a polynomial series estimator in Ching[12], i.e., projecting $log(p_{jt})$ onto a polynomial of $(n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_{bt}, \xi_{gt})$. Assuming that the error term, ν_{jt}, is distributed log normally, I obtain the conditional likelihood of observing p_t,

$$f_p(p_t|n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_l, \gamma), \tag{18}$$

where $\xi_t = (\xi_{bt}, \xi_{gt})$; γ is the vector of parameters that are associated with the state variables in $\wp_j(.)$; θ_l is a set of learning parameters that determines $\sigma^2_{A_g}(t)$ and $E[A_g|I(t)]$.[21]

The observed quantity demanded, q_{jt}, follows a multinomial distribution and therefore is subject to sampling errors, η_{jt}.[22] I incorporate these sampling errors explicitly into the estimation procedure. Given that the market sizes are always over one million, I assume the normal distribution approximates the multinomial distribution well. Let θ_d be the set of demand side parameters, which include θ_l and other preferences parameters. For $j \in \{b, g\}$, the quantity of output, q_{jt}, can be expressed as,

$$q_{jt} = MPr(j|p, n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_d) + \eta_{jt}, \tag{19}$$

where

$$Var(\eta_t) = M \begin{pmatrix} Pr(b|t)(1 - Pr(b|t)) & -Pr(b|t)Pr(g|t) \\ -Pr(b|t)Pr(g|t) & Pr(g|t)(1 - Pr(g|t)) \end{pmatrix}, \tag{20}$$

$$Pr(j|t) = Pr(j|p, n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_d). \tag{21}$$

[21] Note that since I approximate $log(\wp_j(.))$, ν_{jt} will also contain an approximation error, which should be a function of the state variables by construction. I assume that a polynomial series estimator is able to approximate $log(\wp_j(.))$ well, and hence the magnitude of the approximation error is very small, and can be ignored.

[22] BLP does not incorporate sampling errors into their estimation procedure. They consider the sample size, M, to be very large, and hence disregard sampling errors.

Notice that when the sample size is large (e.g. over one million in this context), $Var(\eta)$ may be so small that it alone is not sufficient to explain the discrepancies between the model and the data. Thus, it should be emphasized that the main sources of uncertainty for output are the structural disturbances: $E[A_g|I(t)]$ and ξ_t. I denote $f_q(q_t|p_t, n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_d)$ as the likelihood of observing q_t conditional on $(p_t, n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t)$, where ξ_{jt} is assumed to be *i.i.d.* normal for $j = b, g$.

The joint likelihood of observing (q_t, p_t) is simply the product of $f_q(q_t|p_t, .)$ and $f_p(p_t|.)$, i.e.,

$$l(q_t, p_t|n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_d, \gamma) = \tag{22}$$
$$f_q(q_t|p_t, n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_d) f_p(p_t|n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_l, \gamma).$$

Now note that $\sigma^2_{A_g}(t)$ is a function of $\{q_{g\tau}\}_{\tau=0}^{t-1}$ (see Equation (8)). Therefore, one can rewrite (22) as,

$$l(q_t, p_t|n_{gt}, t, \sigma^2_{A_g}(t), E[A_g|I(t)], \xi_t; \theta_d, \gamma) = \tag{23}$$
$$l(q_t, p_t|n_{gt}, t, \{q_{g\tau}\}_{\tau=0}^{t-1}, E[A_g|I(t)], \xi_t; \theta_d, \gamma).$$

For each market, the likelihood of observing $q = \{q_t\}_{t=0}^T$ and $p = \{p_t\}_{t=0}^T$ is,

$$L(q, p|\{n_{g\tau}, \tau, E[A_g|I(\tau)], \xi_\tau\}_{\tau=0}^T; \theta_d, \gamma) = \tag{24}$$
$$\prod_{t=0}^{T} l(q_t, p_t|n_{gt}, t, \{q_{g\tau}\}_{\tau=0}^{t-1}, E[A_g|I(t)], \xi_t; \theta_d, \gamma).$$

But $(\xi_t, E[A_g|I(t)])$ are unobserved to the analyst and therefore must be integrated over to form the unconditional sample likelihood for (q_t, p_t), that is,

$$L(q, p|\{n_{g\tau}\}_{\tau=0}^T, \{\tau\}_{\tau=0}^T; \theta_d, \gamma) = \tag{25}$$
$$\int\int \prod_{t}^{T} l(q_t, p_t|n_{gt}, t, \{q_{g\tau}\}_{\tau=0}^{t-1}, E[A_g|I(t)], \xi_t; \theta_d, \gamma) dF(\{\xi_\tau\}_{\tau=0}^T) dF(\{E[A_g|I(\tau)]\}_{\tau=0}^T).$$

If ξ_t is *i.i.d.*, the above integrals can be rewritten as,

$$L(q, p|\{n_{g\tau}\}_{\tau=1}^T, \{\tau\}_{\tau=0}^T; \theta_d, \gamma) = \tag{26}$$
$$\int \{\prod_{t}^{T} [\int l(q_t, p_t|n_{gt}, t, \{q_{g\tau}\}_{\tau=0}^{t-1}, E[A_g|I(t)], \xi_t; \theta_d, \gamma) dF(\xi_t)]\} dF(\{E[A_g|I(\tau)]\}_{\tau=0}^T).$$

Evaluating (26) numerically is very difficult. It involves high order integrals because $E[A_g|I(t)]$ is autocorrelated. I resolve this problem by using the method of simulated maximum likelihood. In the appendix, I detail how I carry out the simulation.

The estimation approach described here is in the spirit of Olley and Pakes[44], Geweke and Keane[25] and Erdem, Keane and Sun[23], who use polynomials to approximate the decision rule for selection and investment, an agent's future payoffs, and manufacturers' coupons offering policy, respectively, in order to correct the selection/endogeneity problem. This method has the advantage of correcting the simultaneity problem for the demand model without imposing a particular supply-side model. As a result, using the parameter estimates, one can simulate various supply-side models and see which one generates the pricing function that is closest to the data. The parameter estimates for the pseudo-pricing policy function also allow one to learn the structure of the true pricing policy function. Given this framework, one can easily carry out likelihood ratio tests to see if the pricing policy function depends on unobserved product characteristics. However, this method comes with a cost: it increases the number of parameters and hence the sample size requirement. In particular, if the sample size is not large enough, it may be difficult to identify the effects of the unobserved state variables in the pricing policy functions.

4.2.2 Identification

Now I discuss the identification of the structural parameters in details. Notice that the likelihood function requires the entry time of generic entrants to be exogenous, i.e., uncorrelated with ξ_{jt} and $E[A_g|I(t)]$. If this assumption fails, then the price coefficients would likely be biased. For instance, if entry is more likely when ξ_{jt} and $E[A_g|I(t)]$ are high, it might cause prices and these unobservables to be correlated because prices would likely be functions of the number of generic entrants. But as I discussed in section 2.4, the entry time of generic firms depends on when they receive approvals from the FDA. Usually, the regulatory approval time does not depend on market conditions, i.e., ξ_{jt} and $E[A_g|I(t)]$. Instead, it depends on the quality of the application and the total workload at the Office of Generic Drugs. This institutional feature introduces exogenous randomness to the entry time of generic entrants. This generates exogenous variation in prices and choice sets, which help identify the price coefficients and the extent of consumer heterogeneity (π_0) (Petrin[47]), respectively.[23]

[23]Recall that n_{gt} and t enter the pricing equation only because they both affect the equilibrium prices via the oligopolistic equilibrium. Although n_{gt} enters the demand model, it affects demand only through the denominator of the logit formula. Also, note that the time trend only appears in the utility of choosing the outside good.

When there are two unobserved product characteristics (ξ_t and $E[A_g|I(t)]$), it might seem difficult to identify their standard deviations. But note that the standard deviation of $E[A_g|I(t)]$ will become arbitrarily small in the long run. Hence, the standard deviation of ξ_t can be identified by the steady state fluctuation in market shares.[24] Before reaching the steady state, the empirical fluctuation of market shares net of those due to ξ_t must be contributed by $E[A_g|I(t)]$. The rate at which the fluctuation of market shares converges to the steady state basically depends on the rate of learning, and therefore identifies $\sigma^2_{A_g}(0)$ and σ_δ/κ (Equation (7) and (8)). The initial market shares help identify the initial prior mean (A). The long-run steady state market shares identify the true mean attribute level of generics (A_g). The evolution of the market shares, $\sigma^2_{A_g}(t)$ and $E[A_g|I(t)]$ identify the utility weight for attribute (ω), and the risk aversion parameter (r).

It should be noted that there are a number of markets in which market shares of brand-name drugs stay at a surprisingly high level, even though these markets seem to have reached long run equilibrium with several years having elapsed since patent expiration. Given the large price differentials between brand-name and generic drugs, the estimates of the true mean attribute of the brand-name drug is necessarily higher than that of the generics, i.e., $A_b > A_g$. However, since the FDA has certified the equivalence of generics, some people may be uncomfortable with this result. In fact, it is likely that physicians and patients may value the reputation or the image of the brand-name drug. Hence, when interpreting A_b, one should think of it as the mean attribute level of the brand-name drug plus some psychological benefit of consuming it.

The identification of the coefficients of the unobserved state variables in the pricing policy function partly hinges on the functional form assumptions. This is why I propose to use a flexible functional form to approximate the pricing policy function in Ching[12]. Ideally, if there are no data limitations, one should experiment with different orders of polynomial and select the specification that best fits the data. Given a particular functional form, the coefficients for ξ_{bt} and ξ_{gt} can be identified by the steady state correlations of prices and market shares. The non-steady state correlation between prices and market shares, controlling for ξ_{bt} and ξ_{gt}, identify the coefficient for $E[A_g|I(t)]$.

As for $\bar{\mu}_2$ and ι in Equation (15), recall that these two parameters determine the degree of symmetric unobserved product differentiation among generics over time, which in turn determine

[24] By steady state, I mean $E[A_g|I(t)]$ becomes a constant.

the level of prices and the generic pricing trend generated by a supply side model. This feature has significantly improved the flexibility of a supply side model in generating pricing patterns. I have tried to estimate $\bar{\mu}_2$ and ι jointly with other parameters, but find that the parameter estimates generate unsatisfactory equilibrium pricing patterns that are mostly far too high and flat. This is not surprising as a major source of identification for $\bar{\mu}_i$ and ι should come from the supply side, which is avoided by the estimation method used here. Unfortunately, as discussed above, explicitly incorporating the supply side into the estimation procedure has proved to be computationally infeasible. Therefore, I have decided to calibrate the initial guess of all the parameters by informally matching predicted equilibrium market shares and pricing patterns with the observed data. Then I fix $\bar{\mu}_i$ and ι at the calibrated values and estimate other parameters. The calibrated values for $\bar{\mu}_i$ and ι are given in Table 2.

5 Results

5.1 Econometric Specification

The estimation results presented here are based on data for 14 drugs (i.e., 14 markets). They are grouped into two categories by therapeutic class: (i) heart disease drugs, (ii) anti-depressants and anti-psychotic drugs. I first discuss the baseline specification, which I refer to Model 1a. I assume that there are two types of patients ($k = 0, 1$), and the coefficients for price are type specific. In the baseline case, the following parameters are allowed to differ across markets: the coefficients for the utility of the outside good (ϕ_0, ϕ_t), the mean attribute levels of generics (A_g), and the fraction of experience signals that is revealed to physicians in each period (κ). The rest of the parameters are common across markets within a category, but are allowed to differ across categories. These common parameters include the price coefficients (α^k), the weight attached to the attribute (ω), the risk coefficient (r), the initial prior variance ($\sigma^2_{A_g}(0)$), the experience variability (σ^2_δ), the proportion of each patient type (π_k), and the standard deviation of the unobserved product characteristics (σ_ξ).

For heart disease drugs, I further divide them into two subgroups according to their main uses: (i) high blood pressure (amiloride, clonidine, methyldopa, hydrochlorothiazide methyldopa, and propranolol); and (ii) irregular heart beat or chest pain (disoypramide and verapamil). I expect that patients/physicians may have different initial priors or different degree of heterogeneity

in terms of price sensitivity for these two groups of drugs. Patients may be more tolerant about the absolute differences in terms of purity or coating for drugs that mainly treat high blood pressure – the slight differences may still be able to lower their blood pressure satisfactorily. However, for the later group, the same absolute differences in purity/coating may result in more easily noticeable differences in health outcomes. This may in turn translate to "larger" differences in attribute level measured in patients' utility function.[25] For similar reasons, I expect to see that the proportion of price-sensitive consumers (i.e., π_0) may be smaller for the later group of drugs.

For anti-depressants and anti-psychotic drugs, it is not obvious how to separate our sample drugs a priori according to their uses because all of them treat mental diseases with health outcomes that are easily observed. I therefore consider another way to incorporate heterogeneity for the initial prior and π_0. As I discuss in section 2, it is plausible that the public may gain more confidence about generics over time. This could result in a less diffuse and pessimistic initial prior. Moreover, it is possible that a manufacturer slowly builds up its sales forces. If these two factors matter in explaining the slow diffusion, this should result in more favorable initial prior for generics in markets with patents expired later. The proportion of price-sensitive consumers (π_0) could also change over time because of the penetration of HMOs during the 80s. Based on these intuitions, I divide my sample of anti-depressants and anti-psychotic drugs into two groups: (i) early markets (desipramine, trazodone, perphenazine and thiothixene) refer to the drugs that have their patents expired during December 1984 - June 1985; (ii) late markets (doxepin, maprotiline and haloperidol) refer to the drugs that have their patents expired during July 1985 - December 1987. Table 1 shows the patent expiry dates.[26] I allow the initial prior and π_0 to differ across these two groups.

I also estimate two other specifications for robustness check. It is possible that the initial slow diffusion of generics is due to factors other than learning. As mentioned in section 2.1, it might take time for physicians to learn that generics are available after gaining approval. It might also take time for pharmacies to receive the initial shipments. Unfortunately, I do not have data to directly control for these factors. So Model 1a could overestimate the importance of learning.

[25] In particular, although disopyramide is used to treat irregular heartbeat, it actually could produce new irregular heartbeat. Therefore, it is used in carefully selected patients (MedicineNet[42]).

[26] For heart disease drugs, after I separate the drugs according to their uses, I no longer have enough markets to allow for early and late markets.

These factors, if present, should be relatively more important during the first few periods immediately after the first generic has entered the market. Therefore, to address this potential problem, I re-estimate the model by withholding the first four periods of observations since the first generic enters the market. In other words, I do not require the model to fit the first year of the generic diffusion paths. I refer to this specification as Model 1b. I am hoping that this specification will help alleviate the potential bias due to other factors that mainly influence the initial stage of the diffusion process. Another advantage of estimating Model 1b is that I can use its results to test whether learning is a primary factor in explaining diffusion. If the estimates of Model 1b turn out to be similar to those of Model 1a, then I cannot reject the hypothesis that learning plays a primary role in explaining the slow diffusion. If the estimates of Model 1a and 1b turn out to be quite different, this will suggest that other factors play non-negligible roles in explaining the slow diffusion.

Another factor that could explain the slow diffusion is the penetration of HMO. The number of enrollment in HMO increases from 16.7 millions in 1984 to 33.1 millions in 1990. In Model 2, I control this factor by modeling the proportion of price-sensitive consumers as a function of HMO enrollment levels.

$$\pi_t = \pi_0 + \pi_{hmo} * HMO_t, \tag{27}$$

where HMO_t is the number of patients enrolled in an HMO at time t.[27]

For the pricing policy function, I use the same functional form for all three specifications and for all categories. Ideally one should experiment different order of polynomials when approximating the pricing policy function. However, the number of observations for this study is not sufficiently large to support an extensive specification search. As an initial step, I use a first-order polynomial to approximate the pricing policy function. For $j \in \{b, g\}$,

$$log(p_{jt}) \quad = \quad \gamma_{j0} + \gamma_{j1}t + \gamma_{j2}n_{gt} + \gamma_{j3}E[U^g|I(t)] + \gamma_{j4}(\xi_{jt} - \xi_{-jt}) + \tilde{\nu}_{jt}, \tag{28}$$

where $E[U^g|I(t)] = \omega E[A_g|I(t)] - \omega r E[A_g|I(t)]^2 - \omega r \sigma^2_{A_g}(t)$. Some coefficients for the pseudo-pricing policy functions are common for markets within a category. These common coefficients

[27]Note that the U.S. Statistical Abstract reports the number of HMO enrollment once a year. I extrapolate the HMO enrollment data to obtain quarterly observations when estimating the model. Admittedly, this introduces measurement error, but this also simplifies the estimation significantly.

include those on $E[U^g|I(t)]$, and the difference between the brand-name demand shock and the generic demand shock. Other parameters, which include the intercept, the coefficients for the time trend, the number of generic entrants and the variances of the prediction errors, are allowed to differ across drugs, but with restrictions. This simple specification might first seem inadequate in approximating the pricing policy function. But as I will demonstrate in section 5.2.3, it does quite well in fitting the observed pricing pattern.

For heart disease drugs, the total number of parameters for the demand model (Model 1a) and the pricing policy function are 44 and 36, respectively; for anti-depressants and anti-psychotic drugs, they are 39 and 37, respectively. Treating product/quarter as one observation, the number of observations are 300 for heart disease drugs, and 256 for anti-depressants and anti-psychotic drugs. In all of these specifications, a generic entrant is a manufacturer with an ANDA approval.

5.2 Parameter Estimates

For simplicity, ξ_t's are assumed to be *i.i.d.* normal. For identification reasons, in each category the fraction of experience signals revealed (κ) for one market must be fixed. I fix $\kappa(amiloride)$ for heart disease drugs, $\kappa(desipramine)$ for anti-depressants and anti-psychotic drugs. Similarly, the mean attribute level of one product must be fixed, because the absolute levels have no meaning. A natural choice is to fix the mean attribute level for the brand-name drug, which is set to zero for all categories. The parameter estimates and standard errors are shown in Tables 3 - 9. The number of draws that I use is 100 for $\{E[A_g|I(t)]^r\}_{t=0}^T$ and for ξ_t^s.

5.2.1 Utility Function

All estimates for the structural parameters are statistically significant and share similar qualitative features across categories and specifications. I first discuss the estimates for the learning parameters, which are reported in Table 3. Overall, the utility weight on attribute (ω) and the risk coefficient (r) are positive for all categories, indicating risk-averse behavior. In other words, an increase in the perceived variance of the generic attribute will lower the expected utility of choosing generics. The initial prior variances ($\sigma^2_{A_g}(0)$) are large, indicating that the patients are uncertain about the mean attribute levels of generics initially. The initial prior means (A) are negative, suggesting that the initial expectation about the mean attribute of generics is lower than the value of A_b, which is fixed

at zero. The estimates for consumer heterogeneity are reported in Table 4. The price coefficient of type 0 consumers (i.e., α^0) is much higher than that of type 1 consumers (i.e., α^1) for all categories and all specifications. Their ratios are 3 for heart disease drugs, and 4 for anti-depressants and anti-psychotic drugs. The proportion of price-sensitive patients is roughly 0.28 for drugs that mainly treat irregular heartbeat or chest pain, and 0.35 for the rest of the drugs.

For heart disease drugs, the proportions of the price-sensitive consumers (i.e., π_0) are higher for high blood pressure drugs compared with other drugs that mainly treat irregular heartbeat or chest pain: they are 0.344 vs. 0.283, 0.357 vs. 0.267, and 0.333 vs. 0.273 for Model 1a, 1b, and 2, respectively (see Table 4). The initial prior variances for drugs that mainly treat irregular heartbeat and chest pain (i.e., disoypramide and verapamil) are more diffuse (e.g., 26.55 vs. 23.44 for Model 1a), and their initial prior means are also more negative, compared with other drugs that mainly treat high blood pressure (see Table 3). A likelihood ratio test rejects the hypothesis that the initial prior and π_0 are the same across these two groups of drugs. This confirms the hypothesis that patients/physicians feel that the absolute differences of qualities have stronger impact on the efficacy or side-effects for drugs that mainly treat irregular heartbeats and chest pain. Consequently, the initial priors for this group of drugs are more diffuse and the patients are also less price-sensitive on average, compared with the group of high blood pressure drugs.

For anti-depressants and anti-psychotic drugs, the proportion of the price-sensitive consumers remains fairly stable at around 0.36 in both early and late markets for Model 1a and 1b (see Table 4). According to the point estimates, the initial prior variances and the initial prior means become slightly less diffuse (e.g., 37.18 vs. 38.19 for Model 1a) and slightly less negative (e.g., -24.33 vs. -25.02 for Model 1a), respectively, in the late market (see Table 3). The results do not confirm that consumers gain more confidence about generics over time for anti-depressants and anti-psychotic drugs, or the slow expansion of sales forces matters in explaining the slow diffusion. This could be because (i) the media coverage about the quality of generic drugs was mixed, partly due to the negative campaign organized by brand-name companies to persuade consumers to stay with the branded originals (Lewin[35]); (ii) selling generic drugs, unlike the brand-name originals, need not rely on large sales forces (James[31]), as I discussed earlier in section 2.2.

In Model 2, the coefficient for HMO enrollment (π_{hmo}) is negative and significant for heart disease drugs (see Table 4). But its magnitude is very small, implying that π_0 stays roughly constant

at around 0.33 (high blood pressure drugs) and 0.27 (irregular heart beast/chest pain drugs) over time. For anti-depressants and anti-psychotic drugs, π_{hmo} is positive and marginally significant. But again, its magnitude is very small, implying that π_0 merely increases from 0.360 (when generics just enter the market) to 0.367 (at the end of the sample period). This might be because (i) some patients did not have insurance coverage for prescription drugs before joining HMOs; (ii) many HMOs' formularies did not structure to induce patients to use generics during the 80s.[28] As a result, the overall effects of HMO enrollments on π_0 could be canceled out.

Table 5 presents the estimates for mean attribute levels, and the fractions of experience signals revealed. All of the estimates are significant. Recall that it is the ratio between the experience variability (σ_δ^2) and the fraction of experience signals revealed (κ) that determines the rate of learning. The absolute values of σ_δ^2 and κ by themselves have no meaning. The estimated true mean attribute levels of generics (A_g's) for all markets are significantly different from zero. They are negative and thus lower than the mean attribute level of brand-name originals, which are fixed at zero for all markets. As discussed in section 4.2.2, A_b includes some psychological benefit of choosing a brand-name drug that cannot be sorted out from its actual mean attribute level in the current framework. Hence, one should not conclude that the quality of the brand-name drug is better than the quality of generic drugs simply from the higher value of A_b. There is only one out of 14 true mean attribute levels (verapamil) that is lower than the initial prior. This suggests that patients generally have pessimistic initial priors about generic qualities. This seems plausible for the period of 1984-90 because there were much fewer generic drugs available on the market before 1984. Pessimistic priors could result from lack of actual experiences with generic drugs, and the general perception that generic products have inferior qualities compared with brand-name products. Moreover, this result is also consistent with some survey studies in the 80s (e.g., Carroll et al.[9], Carroll[10]), which find that physicians, pharmacists and patients expressed concerns about the quality of generics.

Table 6 presents the estimates for the utility of the outside goods. All of the parameter estimates are significant. Most of the time trends, except for verapamil, are positive. This indicates

[28] According to a survey conducted by Doering et al.[19] in the late 80s, about 32.5% of the HMOs required a patient to pay the difference in the cost of the medication, and merely 4.8% required a patient to pay higher copayment if he/she wanted the brand-name drug dispensed.

that in general the total demand for the chemicals tend to decrease over time. This is consistent with the previous findings by Caves et al.[11], who argue that the decrease in the total demand for a chemical is due to the dramatic reduction of detailing efforts by brand-name firms after patent expiration.

It should be emphasized that the estimates of the three models are very similar. Even though we try to control for the slow increase in awareness of the existence of generic alternatives (Model 1b), and the slow penetration of HMO (Model 2), the estimates remain virtually unchanged compared with the base model (Model 1a). This provides evidence to support the hypothesis that learning plays a primary role in explaining the slow diffusion of generics.

5.2.2 Pricing policy functions

Now I turn to discuss the results of the pricing policy function. In searching for a specification I first regressed log prices on the observed explanatory variables, (i.e., the time trend and the number of generic entrants), allowing the coefficients to be market specific and unrestricted. But with limited data available, I am forced to impose restrictions on the coefficients when jointly estimating the pricing policy function and the demand model. To gain degrees of freedom, I restrict coefficients on the observed variables that were "close" for different markets to be equal. Restrictions on coefficients are self-explanatory from Table 7 and 8. I also restrict coefficients on unobserved variables to be equal across markets within a category, but allow them to differ across categories. The unobserved variables include the perceived variance of the generic attribute, the expected attribute level of generics, the brand-name demand shock and the generic demand shock. I focus my discussion on the time trend, the number of generic entrants, and the unobserved variables. Since the intercepts do not have decisive structural interpretations, their estimates are not reported here.

The time trend and the number of generics

Generic drugs: Now I discuss the estimates of the time trends and the number of generics for generic drug pricing policy functions. The results are reported in Table 7. For heart disease drugs, the time trends are not significant. But the number of generics is negative and significant for six out of seven markets for all three models. For anti-depressants and anti-psychotic drugs, the point estimates of the time trend are similar across all models. Five out of seven time trends are negative and significant for Model 1a; but it is insignificant for maprotiline in Model 1b and 2, and

insignificant for doxepin for Model 2. The coefficients for the number of generic drugs are negative and significant for all markets. Overall the results suggest that the number of generics plays a role in lowering the generic prices.

Brand-name drugs: Table 8 reports the estimates of the time trend and the number of generics for brand-name drug pricing policy function. For heart disease drugs, the time trends, which are restricted to be the same, are positive and significant at 10% level for Model 1a and 2. But none of the number of generics is significant for all models. For anti-depressants and anti-psychotic drugs, the time trends are significant in five out of seven markets; among them, three are positive and one is negative for all three models. The number of generics are positive and significant for two markets: perphenazine and thiothixene. For them, the entry of generic drugs have caused the brand-name firms to raise their prices. For markets which show the number of generics is not significant, this may reflect that the composition of consumers who keep choosing brand-name drugs is mainly determined by consumer learning, instead of the number of generic alternatives available.

Unobserved variables

Table 9 reports the estimates of unobserved variables: the partial expected utility ($E[U_{ig}^s|I(t)]$), and the difference between the brand-name demand shock and the generic demand shock ($\xi_j - \xi_{-j}$). It turns out that none of the unobserved variables are significant here. This could be due to several reasons. First, $E[U_{ig}^s|I(t)]$ may not be a sufficient statistics for $E[A_g|I(t)]$ and $\sigma^2_{A_g}(t)$ from firms' viewpoints. It might be more appropriate to model $E[A_g|I(t)]$ and $\sigma^2_{A_g}(t)$ to enter the pricing policy function separately. However, this would increase the number of parameters, which is a constraint for this study because of the data sample size. Second, it is possible that firms make their pricing decisions before $E[U_{ig}^s|I(t)]$ and ξ_j are realized. If this is the case, the endogeneity problem of price might be less of a concern here. Overall, given these limitations, I am reluctant to draw a definite conclusion on whether the pricing decisions depend on the unobservables: $E[A_g|I(t)]$, $\sigma^2_{A_g}(t)$ and ξ_j.

5.2.3 Goodness-of-fit

Overall, I find that the results are robust across all specifications. To illustrate the goodness-of-fit, I simulate 100 sequences of price and quantity pairs for both the brand-name drug and the generic drug, from the demand model and the pseudo-pricing policy function using the parameter estimates.

The number of generic firms is taken as exogenous. I then compute the average predicted price and quantity for each period by averaging simulated prices and quantities. I only show the results based on Model 1a. The simulation results from the other two specifications are similar.

Figure 4 and 5 plot the average predicted demand vs. the actual demand for brand-name drugs and generic drugs, respectively. In general, the model is able to fit the demand pattern quite well, particularly for brand-name drugs. For generic drugs, some of the predicted demand exhibits inverted U-shapes. This is partly because I restrict the standard error of the logit taste shocks for generics to decline over time. Such a restriction implies that the utility of choosing generics drops over time, holding everything else fixed. Another factor that causes the demand for generics finally drops is that the time trend for the utility of outside goods is positive. The initial increase in the utility of choosing generics is due to the decrease in generic prices and uncertainty. Figure 6 and 7 plots the average predicted price and the actual price for brand-name drugs and generic drugs, respectively. Despite the simple functional form for the pricing policy functions, they do well in fitting the pricing pattern.

I also calculate the mean absolute percentage error (MAPE). Table 10 shows the results. Overall, the MAPE for both the brand-name prices and generic prices are always lower than 10%. For the brand-name quantity sold, it is always smaller than 17%. However, it is relatively higher for the generic quantity sold: one market has 14.6%; seven markets have below 30%; two markets have between 30% and 40%; two markets have between 40% and 50%; two markets have more than 70%.

To further illustrate the goodness-of-fit for this model, I take the average of the predicted demand for generics across all markets. The results are reported in the fifth column of Table 11. The average actual generic demand for all markets is reported in the fourth column of Table 11. The second column of Table 11 lists the number of markets used to calculate average generic demands. The number of markets available drops over time because some markets have their patents expired later than others, and my data end in the fourth quarter of 1990. Period zero refers to the quarter in which generics just entered the market. The initial average predicted demand is 2.02 million patient days, which is about 50 percent of the average actual demand. In the 4th quarter, it reaches the level of 11.55 million patient days, which is very similar to the observed level. From the fifth quarter to the 13th quarter, the average predicted demand remains slightly higher than the average

actual demand. From the 14th quarter on, the average predicted demand becomes slightly lower than the average actual demand.[29]

Overall, the simulation results demonstrate that the model fits the pricing patterns and the brand-name demand pattern quite well. Although the model is able to generate a diffusion path for generics that mimic the data on average, it has some difficulties in fitting the exact shape of the observed diffusion paths for generics. It is common that structural econometric models have difficulties in fitting some features of the data because of the tight restrictions imposed by the theory. My model is no exception here. However, it should be emphasized that my primary goal of estimating this structural model is to obtain more insights about the prescription drug market by counterfactual exercises. I will demonstrate this in the next three subsections using Model 1a.

5.3 Change in Demand Composition

It is likely that the diffusion rate of generics varies across different segments of consumers. Intuitively, I expect that the diffusion rate of generics is faster in the price-sensitive segment of the market. However, with only aggregate level data, I do not observe the demand patterns of the price-sensitive patients and the price-insensitive patients separately. Having estimated π_0, α^0 and α^1 (along with other structural parameters), I am able to simulate the demand patterns of each patient type. The simulated demand patterns are used to illustrate how the diffusion rate of generics varies across patient types.

Figure 8 shows the demand patterns for the price-sensitive patients. The demand for generics surpasses that for brand-name drugs in 11 out of 14 markets over time – the exceptions are amiloride, disopyramide and verapamil. Moreover, in most cases, it takes less than ten quarters for generic drugs to outperform brand-name drugs in these markets. Figure 9 shows the demand patterns for price-insensitive patients. For these patients, there are only six markets (clonidine, desipramine, doxepin, haloperidol, perphenazine, and thiothixene), in which the demand for generic drugs outperforms that for brand-name drugs. Brand-name drugs remain dominant throughout the period in five markets (amiloride, propranolol, disopyramide, verapamil and maprotiline).

[29] I do not focus my discussion on the period beyond the 19th quarter because the number of markets has become quite small.

A comparison of Figure 8 and 9 reveals that on average price-sensitive patients switch to generics quicker than price-insensitive patients as the uncertainty about the generic quality has been resolved over time. In particular, the proportion of brand-name drugs consumers who are price-insensitive keeps increasing over time even after the number of generic entrants has become fixed. This suggests that the demand faced by the brand-name firms becomes less price sensitive over time. This provides a possible explanation for why the brand-name firms keep increasing their prices even after the number of generic entrants has been stabilized.

5.4 Branded Price Elasticity of Demand over Time

The previous literature (e.g., Grabowski and Vernon[26], Frank and Salkever[24]) has conjectured that the reason why the brand-name firms raise their prices after facing generic entry is due to market segmentation. They argued that when generics become available, price-sensitive consumers switch to generics, causing the demand for brand-name drugs to become more inelastic after generic entry. However, as I mentioned earlier, consumer heterogeneity alone is not sufficient to explain why brand-name prices keep increasing over time after the number of generic entrants has become stabilized.

To further investigate what drives the branded pricing pattern, I compute the static branded price elasticities of demand over time. I find that the brand-name price elasticities, evaluated at the observed prices, are often less than one and increase over time. To illustrate the pattern, I report the average brand-name price elasticity across markets in the last column of Table 11. It shows that the average static brand-name price elasticity increases from the 0th to the 11th quarter and is less than one. The pattern becomes less clear beyond the 11th quarter because the number of markets has dropped. The results appear to be inconsistent with the hypothesis that firms choose prices to maximize their current profits. However, they are consistent with the hypothesis that firms are forward-looking and physicians/patients learn about the qualities of generics via consumption experiences. With uncertainty about the quality of generic drugs, brand-name firms may set prices lower than what they would do if they were myopic, in order to slow down the learning process. This could explain why the brand-name price elasticities are less than one. Note that the dynamic incentive to slow down the learning process diminishes over time when the uncertainty about the generic quality has slowly resolved. Consequently, brand-name firms gradually price more like

myopic firms over time. This could also be another contributing factor that explains why the brand-name prices gradually increase over time. The results not only provide further support for the learning hypothesis, but also suggest that brand-name firms' forward-looking behavior may play a role in explaining the branded pricing pattern.

5.5 Rate of Learning and Diffusion

One way to measure the rate of learning is to use the rate at which the perceived variance of the generic attribute diminishes over time. The third column of Table 11 shows the change in the average perceived variance (taken across all markets) over time. Again, the simulation and the average are obtained in the same fashion as the previous section. Notice that the average perceived variance quickly decreases by about 50 percent in the first four quarters after the first generic entered the market (from 31.04 to 15.71). Then it decreases by approximately another 50 percent in the next three quarters (from 15.71 to 8.88). The rate of learning keeps diminishing as the average perceived variance becomes smaller. This is consistent with the Bayesian updating formula for the perceived variance (see Equation (8)).[30]

How much of the slow diffusion of generics is due to learning? This question is not trivial to answer. Other than learning, the decline in generic prices and the change in the value of the outside good can also affect the diffusion rate. One advantage of estimating the structural parameters of the learning model is that it allows me to disentangle the effect of learning from other factors, by simply changing the parameters values for the initial prior. To investigate the effect of learning, I set the initial perceived variance ($\sigma^2_{A_g}(0)$) to zero, and the initial prior mean attribute level (A) to the true mean attribute level (A_g). Keeping everything else at the estimated parameter values, I then re-simulate the model and compute the average predicted quantities and prices. In this counterfactual situation, the simulated data represents outcomes from markets where patients are certain about the mean attribute level right from the beginning. By examining the difference between the simulated data from the model without uncertainty and that from the original model with uncertainty, I can conclude how much of the slow diffusion is due to learning.

[30] One may notice that the average perceived variance is fluctuating, instead of diminishing, from quarter 13 to the quarter 18. This is due to the decline in the number of markets that are available for computing the average perceived variance (See column two of Table 11).

Again, I use the demand model and the pricing policy function to conduct this exercise. The results are summarized by reporting the average predicted demand for generics. The sixth column of Table 11 reports the average predicted generic demand for the model without uncertainty. It appears that without learning the demand for generics is much too high initially, and stays too high for many quarters compared with the average actual demand (the fourth column of Table 11), or the average predicted demand from the original model with uncertainty (the fifth column of Table 11). In sum, the predicted diffusion rate in the model without uncertainty is much quicker than the actual diffusion rate.

It should be emphasized that the change in generic demand in the model without uncertainty has filtered out the learning effect, and therefore is mainly due to (i) the slow decrease in generic prices and (ii) the slow increase in the number of generic entrants. A comparison about the predictions between these two version of the model shows that these two factors are not sufficient to generate the generic diffusion rate observed in the data. In other words, the uncertainty about generic attribute appears to be crucial in explaining the observed diffusion patterns.

The results also suggest that it may be worthwhile to launch an advertising campaign to educate the public about the safety and bioequivalence of generic drugs. Such a campaign could reduce the consumer uncertainty about the generic quality, and hence increase generic substitution and lower the prescription drug expenditures, as predicted by this experiment. Recently, an insurance company, Blue Cross Blue Shield of Michigan, has adopted this approach to promote the use of generic drugs. Their one million dollars statewide advertising campaign for generics in 2001, which emphasizes generic drugs are safe and effective, has resulted in about 30 million dollars savings this year according to the calculation by Blue Cross Blue Shield (Sherrid[50]).

5.6 More Robustness Checks

I have tested whether the generic scandal may create a structural change on the prior mean and prior variance in the 3rd quarter of 1989, when the generic scandal broke out. I do not find any significant effects for all markets in my sample. This could be because there are only a few generic firms involved in the scandal. Moreover, some generic drugs have already been on the market for a few years when the scandal broke out in the late 80s. Those drugs should have accumulated a

lot of consumption experiences when the scandal surfaced. According to the Bayesian updating theory, their priors had become quite precise and should be less susceptible to change.

In order to control for the possibility of the slow expansion of distribution channels, I have also re-estimated the model by replacing the number of generic manufacturers with the number of distributors.[31] The results on the structural demand model parameter estimates remain robust for this specification. But the coefficient for the number of generic firms in the pricing policy function becomes much smaller. This is not surprising because the number of distributors is typically much larger than the number of manufacturers. Due to the space constraint, I do not report the details of the results here. They are available upon request.

6 Conclusion

Motivated by the slow diffusion of generics and the counterintuitive pricing behavior of brand-name firms, I start out with a demand model of prescription drugs, which incorporates consumer learning and consumer heterogeneity. I estimate the model using data from 14 markets, and find evidence that consumer learning plays a role in explaining the slow diffusion of generics during the 80s. The estimates of the model suggest that the public was risk-averse, uncertain about generic qualities, and had a pessimistic initial prior. Using the estimated model, I demonstrate that the diffusion rate of generics for price-sensitive patients is much faster than that for price-insensitive patients. In contrast to what the previous literature believes, I also find that the brand-name price elasticity of demand increases over time, and is often less than one. The results suggest that brand-name firms might be forward-looking and set prices lower than what they would do if they were myopic, in order to slow down the learning process for generics. But such an incentive diminishes over time as the uncertainty slowly resolves, leading the brand-name firms to gradually price higher over time.

This paper also develops a new procedure to take the endogeneity problem of prices into account when estimating discrete choice product differentiation models using aggregate market share data. Here I do not find evidence that prices are correlated with the unobservables in the model. In another application, Ching and Ishihara[14] use this method to estimate the effect of

[31] A generic manufacturer may use multiple distributors to sell their generic drugs. See James[31] or Caves et al.[11] for a discussion.

detailing, and find evidence that the method corrects the bias due to the endogeneity problem of detailing. While in principle the procedure is able to identify the parameters of the utility function and the pseudo-pricing policy function, it may require more data relative to the standard procedure, in order to achieve satisfactory results. More evidence on the properties of this estimator remains to be gathered in the future.

One limitation of this study is that the counterfactual exercise used to determine the importance of learning assumes the reduced form pricing policy function remains a good proxy for the true equilibrium pricing policy function after changing the initial priors. To address this issue, I generate the pricing policy function by solving a stochastic dynamic oligopoly structural model numerically in Ching[13], and examine the extent to which this pricing policy function can explain the data empirically. One important reason for estimating the structural preference parameters and solving a dynamic oligopoly structural model is that they are crucial for evaluating various policy experiments. For example, in Ching[13] I use the equilibrium model to study the welfare impact of a policy, which reduces the FDA average approval time for marketing generic drugs.

Another limitation is that this paper does not study how detailing may affect the demand for brand-name drugs and generic drugs. In particular, I do not address the possibility that the depreciation of detailing goodwill stock may also explain the slow diffusion of generics. If this effect is strong, I would overestimate the importance of learning here. However, Caves et al.[11] find that detailing by brand-name firms typically drops significantly after patent expiration. Moreover, generic firms do not use detailing to market their products. The evidence seems to suggest that detailing might not be of first order importance in affecting the choice between brand-name drugs and generics.[32] Certainly, one cannot draw a definite conclusion without further investigation. Disentangling between depreciation of detailing goodwill stock and learning will be an interesting problem for future research.

[32]When it comes to the choice between brand-name originals or their generic versions, the literature finds evidence that the effect of detailing is mainly chemical specific instead of brand specific. Ellison and Ellison[20] make this assumption and derive a hypothesis that brand-name firms have incentives to reduce their detailing/advertising efforts prior to patent expiration, in order to make the market less attractive for generics to enter. They find evidence to support this hypothesis. In addition, Scott Morton[49] does *not* find evidence that the detailing efforts prior to patent expiration deter generic entry.

Table 1: Summary statistics

	patent expiry date(mm/yy)	entry delay*	total # entrants	max $\frac{p_g}{p_b}$	min $\frac{p_g}{p_b}$	max $\frac{q_g}{q_b+q_g}$	min $\frac{q_g}{q_b+q_g}$
Heart disease drugs							
High blood pressure drugs							
Amiloride	12/84	4	1	0.543	0.442	0.153	0.028
Clonidine	7/86	0	12	0.387	0.085	0.786	0.163
Methyldopa	12/84	0	17	0.696	0.248	0.633	0.035
Hydrochlorothiazide Methyldopa	12/84	5	12	0.630	0.348	0.439	0.018
Propranolol	12/84	3	21	0.686	0.128	0.521	0.129
Irregular heart beats and chest pain drugs							
Disopyramide	1/86	2	11	0.815	0.378	0.261	0.036
Verapamil	10/86	0	12	0.677	0.273	0.159	0.101
Anti-depressants and Anti-psychotic drugs							
Early Markets							
Desipramine	2/85	4	4	0.775	0.191	0.739	0.051
Trazodone	4/85	6	11	0.815	0.204	0.675	0.057
Perphenazine	5/85	9	3	0.769	0.568	0.636	0.032
Thiothixene	12/84	11	5	0.682	0.377	0.716	0.041
Late Markets							
Doxepin	7/86	2	11	0.434	0.195	0.767	0.158
Maprotiline	8/85	9	4	0.717	0.543	0.353	0.033
Haloperidol	1/86	1	13	0.748	0.320	0.733	0.206

*entry delay (measured in quarter) - the time lag between the first generic entry and the patent expiry date.

Table 2: Calibrated parameter values determining the variance of Logit errors

	μ_1^0	μ_1^1	μ_2^0	μ_2^1	ι	ψ
HEART DISEASES						
amiloride	0.5	0.5	0.35	0.35	0.00	1
clonidine	1.0	1.0	0.70	0.70	0.10	5
methyldopa	1.00	1.00	0.70	0.70	0.025	8
hydrochlorothiazide methyldopa	0.90	0.90	0.63	0.63	0.03	4
propranolol	0.75	0.75	0.53	0.53	0.08	50
disopyramide	1.75	1.75	1.23	1.23	0.01	10
verapamil	1.25	1.25	0.88	0.88	0.01	4
ANTI-DEPRESSANTS and ANTI-PSYCHOTIC						
Early Markets						
desipramine	2.50	1.50	2.50	1.50	0.04	1
trazodone	2.50	1.50	2.50	1.50	0.05	1
perphenazine	3.00	1.80	3.00	1.80	0.035	1
thiothixene	1.50	0.90	1.50	0.90	0.035	3
Late Markets						
doxepin	1.00	0.60	1.00	0.60	0.05	10
maprotiline	2.20	1.32	2.20	1.32	0.035	15
haloperidol	2.20	1.32	2.20	1.32	0.035	1

The superscript of μ indexes patient type;
the subscript of μ indexes level for the nested logit demand model;
ψ is the constant term for increasing the sampling errors.

Table 3: Estimated parameters for learning

	Heart Diseases			Anti-depressants Anti-psychotics (anti-dppsy)		
Model	1a	1b	2	1a	1b	2
risk coefficient (r)	0.576*	0.584*	0.570*	0.207*	0.215*	0.192*
	(0.009)	(0.011)	(0.009)	(0.003)	(0.007)	(0.004)
utility weight for attribute (ω)	0.014*	0.014*	0.015*	0.038*	0.038*	0.040*
	(2.2e-4)	(2.0e-4)	(2.3e-4)	(4.0e-4)	(0.001)	(0.001)
experience variability (σ_δ^2)	0.195*	0.193*	0.192*	0.782*	0.716*	0.816*
	(0.005)	(0.008)	(0.005)	(0.015)	(0.066)	(0.031)
initial prior variance ($\sigma_{A_g}^2(0)$)						
early markets for anti-dppsy/	23.44*	22.45*	23.33*	38.19*	38.11*	37.98*
high blood pressure drugs	(0.984)	(1.960)	(1.226)	(0.756)	(1.320)	(1.885)
late markets for anti-dppsy/	26.55*	26.55*	26.55*	37.18*	37.53*	36.95*
heart beat or chest pain drugs	(0.921)	(1.002)	(1.011)	(0.913)	(1.00)	(2.196)
initial prior means for anti-dppsy						
early markets (A^e)				-25.02*	-24.78*	-24.97*
				(0.142)	(0.579)	(0.664)
late markets (A^l)				-24.33*	-24.32*	-24.21*
				(0.151)	(0.401)	(0.678)
initial prior means for heart disease drugs						
High blood pressure						
Amiloride	-15.21*	-15.19*	-15.21*			
	(0.222)	(1.000)	(0.147)			
Clonidine	-19.70*	-19.55*	-19.69*			
	(0.226)	(0.318)	(0.315)			
Methyldopa	-17.03*	-17.26*	-17.08*			
	(0.189)	(0.826)	(0.151)			
Hydrochlorothiazide Methyldopa	-20.85*	-20.82*	-20.86*			
	(0.256)	(0.575)	(0.129)			
Propranolol	-17.48*	-17.58*	-17.42*			
	(0.234)	(0.168)	(0.194)			
Irregular heart beats or chest pain						
Disopyramide	-30.38*	-30.21*	-30.34*			
	(0.459)	(0.448)	(0.369)			
Verapamil	-24.85*	-24.82*	-24.90*			
	(0.334)	(1.001)	(0.325)			

Standard errors are reported in parenthesis.
Number of draws for demand shocks = 100.
Number of draws for $E[A_g|I(t)] = 100$.
Notes:
* - t-statistic > 1.96
anti-dppsy - anti-depressants and anti-psychotic drugs

Table 4: Estimated parameters for consumer heterogeneity

	Heart Diseases			Anti-depressants Anti-psychotics (anti-dppsy)		
Model	1a	1b	2	1a	1b	2
Consumer heterogeneity parameters						
type 0 price coefficient (α^0)	0.029*	0.031*	0.031*	0.053*	0.051*	0.053*
	(9.7e-4)	(4.0e-4)	(2.6e-4)	(6.1e-4)	(0.001)	(6.3e-4)
type 1 price coefficient (α^1)	0.008*	0.008*	0.008*	0.012*	0.012*	0.011*
	(3.0e-4)	(2.0e-4)	(3.0e-4)	(1.2e-4)	(3.7e-4)	(1.1e-4)
proportion of type 0 (π_0)						
high blood pressure drugs	0.344*	0.357*	0.333*			
	(0.037)	(0.011)	(0.014)			
heart beat/chest pain drugs	0.283*	0.267*	0.273*			
	(0.015)	(0.008)	(0.015)			
early markets for anti-depressants/psychotics				0.357*	0.363*	0.352*
				(0.008)	(0.023)	(0.012)
late markets for anti-depressants/psychotics				0.360*	0.375*	0.352
				(0.008)	(0.013)	
π_{hmo}			-7.6e-3*			4.3e-3**
			(3.0e-3)			(2.4e-3)
standard deviation of unobserved product characteristic (σ_ξ)	0.287*	0.281*	0.271*	0.478*	0.468*	0.475*
	(0.001)	(0.003)	(0.014)	(0.002)	(0.006)	(0.003)
Log Likelihood	-3347	-1754	-3188	-3356	-1290	-3104

Standard errors are reported in parenthesis.
Number of draws for demand shocks = 100.
Number of draws for $E[A_g|I(t)] = 100$.
Notes:
* - t-statistic > 1.96.
** - t-statistic > 1.65.

Table 5: Estimated Mean Attribute levels and Fraction of experience signals revealed

	True mean attributes (A_g)			Fraction of experience signals revealed (κ)		
	1a	1b	2	1a	1b	2
Heart Diseases Drugs:						
amiloride	-12.61*	-12.56*	-12.61*	6.7e-6	6.7e-6	6.7e-6
	(0.155)	(0.150)	(0.147)			
clonidine	-0.945*	-0.815*	-0.979*	1.0e-10*	1.0-e10*	1.1e-6*
	(0.375)	(0.332)	(0.315)	(4.8e-12)	(5.1e-12)	(1.9e-12)
methyldopa	-15.00*	-14.91*	-15.00*	2.4e-9*	2.0e-9*	1.8e-9*
	(0.171)	(0.156)	(0.151)	(1.2e-10)	(1.1e-10)	(3.8e-11)
hydrochlorothiazide methyldopa	-13.34*	-13.35*	-13.37*	4.3e-10*	4.7e-10*	4.3e-10*
	(0.164)	(0.163)	(0.129)	(2.2e-12)	(3.0e-11)	(9.9e-12)
propranolol	-8.94*	-8.64*	-8.91*	2.2e-12*	3.0e-11*	9.9e-12*
	(0.527)	(0.312)	(0.194)	(9.6e-12)	(9.6e-12)	(1.1e-11)
disopyramide	-25.14*	-24.96*	-25.07*	3.3e-10*	2.7e-10*	7.0e-10*
	(0.560)	(0.537)	(0.369)	(3.8e-11)	(4.4e-11)	(4.8e-11)
verapamil	-25.83*	-25.79*	-25.86*	1.4e-12*	4.2e-13*	6.1e-13*
	(0.334)	(0.328)	(0.325)	(8.0e-10)	(1.7e-9)	(3.7e-10)
Anti-depressants and Anti-psychotic drugs:						
Early Markets						
desipramine	-16.21*	-16.19*	-15.77*	1.0e-8	1.0e-8	1.0e-8
	(0.112)	(0.285)	(0.175)			
trazodone	-17.17*	-17.15*	-17.05*	1.4e-9*	1.4e-9*	1.3e-9*
	(0.123)	(0.364)	(0.250)	(2.0e-11)	(5.5e-11)	(5.4e-11)
perphenazine	-12.81*	-12.70*	-12.57*	8.0e-9*	7.8e-9*	8.5e-9*
	(0.105)	(0.310)	(0.152)	(1.5e-10)	(5.7e-10)	(2.8e-10)
thiothixene	-12.06*	-11.95	-11.98*	3.6e-8*	3.5e-8*	4.0e-8*
	(0.076)	(0.262)	(0.159)	(5.3e-10)	(4.1e-9)	(1.1e-9)
Late Markets						
doxepin	-10.24*	-10.26*	-10.23*	1.0e-8*	9.6e-9*	1.0e-8*
	(0.072)	(0.158)	(0.143)	(2.3e-10)	(4.2e-10)	(6.3e-10)
maprotiline	-18.26*	-18.13*	-18.12*	3.3e-10*	8.5e-10*	6.5e-10*
	(0.143)	(0.418)	(0.242)	(1.4e-9)	(6.1e-10)	(1.3e-9)
haloperidol	-14.22*	-14.16*	-14.30*	9.2e-10*	8.6e-10*	7.7e-10*
	(0.181)	(0.261)	(0.188)	(4.0e-11)	(5.1e-11)	(2.8e-11)

Notes:
* - t-statistic > 1.96

Table 6: Estimated Coefficients for the outside good

	intercept (ϕ_0)			time trend (ϕ_t)		
Model	1a	1b	2	1a	1b	2
Heart Diseases Drugs:						
Amiloride	1.673*	1.690*	1.668*	3.8e-4**	-4.9e-4**	0.001*
	(0.015)	(0.010)	(0.009)	(2.0e-4)	(2.7e-4)	(1.9e-4)
Clonidine	0.063	0.017	0.074*	0.073*	0.077*	0.074*
	(0.044)	(0.018)	(0.018)	(8.0e-4)	(8.7e-4)	(7.5e-4)
Methyldopa	-1.140*	-1.187*	-1.127*	0.085*	0.087*	0.085*
	(0.067)	(0.006)	(0.025)	(0.001)	(0.025)	(6.4e-4)
Hydrochlorothiazide Methyldopa	0.093*	0.167*	0.090*	0.071*	0.068*	0.073*
	(0.039)	(0.020)	(0.019)	(5.0e-4)	(4.0e-4)	(5.6e-4)
Propranolol	-1.760*	-1.708*	-1.731*	0.059*	0.056*	0.059*
	(0.063)	(0.020)	(0.019)	(0.001)	(8.0e-4)	(8.1e-4)
Disopyramide	-5.773*	-5.488*	-5.781*	0.186*	0.169*	0.188*
	(0.052)	(0.071)	(0.052)	(0.003)	(0.003)	(0.002)
Verapamil	0.101*	0.191*	0.111*	-0.268*	-0.278*	-0.268*
	(0.025)	(0.015)	(0.011)	(0.002)	(0.002)	(0.002)
Anti-depressants and Anti-psychotic drugs:						
Early Markets						
desipramine	0.383*	0.368*	0.363*	0.010*	0.008*	0.013*
	(0.016)	(0.036)	(0.017)	(6.0e-4)	(0.002)	(4.8e-4)
trazodone	-0.112*	-0.099*	0.087*	0.0046*	2.6e-3	0.006*
	(0.016)	(0.033)	(0.023)	(5.6e-4)	(2.6e-3)	(6.1e-4)
perphenazine	1.223*	1.177*	1.293*	0.0198*	0.022*	0.0198*
	(0.019)	(0.081)	(0.023)	(5.2e-4)	(2.7e-3)	(6.0e-4)
thiothixene	-1.216*	-1.217*	-1.084*	0.043*	0.042*	0.043*
	(0.014)	(0.036)	(0.018)	(3.0e-4)	(1.3e-3)	(4.8e-4)
Late Markets						
doxepin	-1.094*	-1.132*	-1.060*	0.016*	0.016*	0.019*
	(0.012)	(0.023)	(0.020)	(4.9e-4)	(9.3e-4)	(4.9e-4)
maprotiline	1.367*	1.342*	1.395*	0.093*	0.095*	0.092*
	(0.014)	(0.052)	(0.015)	(5.6e-4)	(0.002)	(5.9e-4)
haloperidol	-3.243*	-3.309*	-3.182*	0.187*	0.185*	0.201*
	(0.039)	(0.052)	(0.055)	(0.003)	(0.003)	(0.002)

Notes:
* - t-statistic > 1.96

Table 7: Pricing policy function for generic drugs: time trend and number of generics

Model	1a	1b	2
Heart disease drugs:			
time trend (γ_{g1}):			
amiloride,			
hydrochlorothiazide methyldopa,			
verapamil	0.008	0.008	0.008
	(0.023)	(0.020)	(0.013)
clonidine	-0.037	-0.035	-0.035
	(0.033)	(0.048)	(0.034)
methyldopa	-0.033	-0.030	-0.035
	(0.021)	(0.035)	(0.025)
propranolol	0.005	0.010	0.006
	(0.005)	(0.027)	(0.005)
disopyramide	-0.002	-0.007	-0.003
	(0.006)	(0.018)	(0.025)
no. of generics (γ_{g2}):			
amiloride, clonidine,			
disopyramide,			
hydrochlorothiazide methyldopa	-0.046*	-0.046*	-0.034*
	(0.020)	(0.023)	(0.019)
methyldopa	-0.014	-0.019	3.3e-4
	(0.022)	(0.025)	(5.7e-4)
propranolol	-0.085*	-0.093*	-0.090*
	(0.019)	(0.036)	(0.019)
verapamil	-0.113*	-0.117*	-0.114*
	(0.040)	(0.065)	(0.061)
Anti-depressants and Anti-psychotic drugs:			
time trend (γ_{g1}):			
desipramine, haloperidol	-0.077*	-0.070*	-0.074*
	(0.009)	(0.015)	(0.021)
trazodone	-0.030	-0.024	-0.027
	(0.035)	(0.055)	(0.062)
perphenazine	-0.013**	-0.009	-0.012
	(8.5e-3)	(0.012)	(0.014)
thiothixene	-0.046*	-0.041*	-0.043*
	(0.009)	(0.009)	(0.015)
doxepin	-0.031*	-0.028*	-0.033
	(0.018)	(0.013)	(0.034)
maprotiline	-0.021*	-0.017	-0.018
	(0.009)	(0.013)	(0.016)
no. of generics (γ_{g2}):			
desipramine, doxepin, maprotiline,			
perphenazine, thiothixene	-0.032*	-0.031*	-0.036*
	(0.011)	(0.013)	(0.016)
trazodone	-0.074*	-0.039*	-0.074*
	(0.014)	(0.017)	(0.020)
haloperidol	-0.014*	-0.021*	-0.013
	(0.007)	(0.011)	(0.015)

* - t-statistic > 1.96; ** - t-statistic > 1.65.

Table 8: Pricing policy function for brand-name drugs: time trend and number of generics

Model	1a	1b	2
Heart disease drugs:			
time trend (γ_{b1}):	0.014**	0.014	0.014**
	(9.3e-3)	(9.6e-3)	(8.0e-3)
no. of generics (γ_{b2}):			
amiloride	0.027	0.036	0.029
	(0.035)	(0.077)	(0.060)
clonidine	0.017	0.018	0.018
	(0.024)	(0.027)	(0.015)
methyldopa	0.002	0.003	0.003
	(0.006)	(0.011)	(0.006)
hydrochlorothiazide methyldopa	0.002	0.005	0.003
	(0.009)	(0.010)	(0.010)
propranolol	0.002	0.003	0.003
	(0.004)	(0.009)	(0.013)
disopyramide	-0.022	-0.017	-0.020
	(0.032)	(0.041)	(0.022)
verapamil	-0.008	-0.004	-0.007
	(0.024)	(0.021)	(0.019)
Anti-depressants and Anti-psychotic drugs:			
time trend (γ_{b1}):			
desipramine, trazodone	0.041*	0.041*	0.045*
	(0.018)	(0.007)	(0.009)
haloperidol	-0.055*	-0.050*	-0.051*
	(0.025)	(0.011)	(0.015)
perphenazine	0.006	-0.007	0.007
	(0.008)	(0.010)	(0.012)
thiothixene	0.012*	0.018*	0.016*
	(0.006)	(0.008)	(0.009)
doxepin	0.025*	0.022*	0.028*
	(0.009)	(0.010)	(0.013)
maprotiline	0.010	0.002	0.009
	(0.011)	(0.011)	(0.013)
no. of generics (γ_{b2}):			
desipramine, trazodone	-0.034	-0.019	-0.036
	(0.033)	(0.013)	(0.014)
haloperidol	0.025	0.016	0.003
	(0.022)	(0.010)	(0.014)
perphenazine	0.020*	0.070	0.014
	(0.006)	(0.060)	(0.039)
thiothixene	1.5e-3*	-0.012	1.3e-3
	(6.0e-4)	(0.034)	(0.023)
doxepin	-1.5e-3	-6.3e-3	-1.3e-3
	(0.006)	(0.005)	(0.013)
maprotiline	-0.008	0.016	-0.009
	(0.037)	(0.027)	(0.016)

* - t-statistic > 1.96; ** - t-statistic > 1.65.

Table 9: Pricing policy function: unobserved variables

	Heart Diseases			Anti-depressants Anti-psychotics		
Model	1a	1b	2	1a	1b	2
Brand-name:						
$E[U_{ig}^s \mid I(t)]$ (γ_{b3})	-0.006	0.002	-0.004	-2.7e-3	-7.0e-4	-0.005
	(0.009)	(0.002)	(0.011)	(0.010)	(0.019)	(0.011)
$\Delta\xi_j$ (γ_{b4})	5.9e-5	-2.0e-4	5.0e-4	5.8e-5	2.6e-4	5.6e-4
	(6.0e-4)	(0.001)	(0.004)	(0.002)	(0.003)	(0.002)
Generic:						
$E[U_{ig}^s \mid I(t)]$ (γ_{b3})	-0.106	-0.111	-0.110	-0.011	-0.010	0.026
	(0.061)	(0.090)	(0.076)	(0.008)	(0.026)	(0.025)
$\Delta\xi_j$ (γ_{b4})	8.4e-4	5.0e-4	8.4e-4	9.0e-4	2.7e-4	8.3e-4
	(0.006)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)
Brand-name and generic:						
variance of prediction error:	2.10	2.62	2.09	0.239	0.238	0.411
	(0.31)	(0.45)	(0.30)	(0.10)	(0.096)	(0.221)

Standard errors are reported in parenthesis
Notes:
* - t-statistic > 1.96
** - t-statistic > 1.65

Table 10: Goodness-of-fit (Mean Absolute Percentage Error)

	Demand		Price	
	brand-name	generics	brand-name	generics
Heart disease drugs				
High blood pressure drugs				
Amiloride	0.155	0.384	0.025	0.030
Clonidine	0.089	0.364	0.025	0.029
Methyldopa	0.169	0.257	0.015	0.019
Hydrochlorothiazide Methyldopa	0.113	0.146	0.047	0.052
Propranolol	0.146	0.221	0.032	0.037
Irregular heart beats and chest pain drugs				
Disopyramide	0.058	0.276	0.021	0.026
Verapamil	0.124	0.237	0.035	0.046
Anti-depressants and Anti-psychotic drugs				
Early Markets				
Desipramine	0.114	0.295	0.018	0.022
Trazodone	0.158	0.293	0.040	0.051
Perphenazine	0.122	1.130	0.035	0.042
Thiothixene	0.113	0.435	0.029	0.035
Late Markets				
Doxepin	0.086	0.468	0.013	0.015
Maprotiline	0.133	0.735	0.023	0.030
Haloperidol	0.164	0.237	0.070	0.089

Table 11: Average generic sales, predicted generic sales and predicted perceived variance (number of patient days, million), mean price elasticity of demand

Time (quarter)	Number of markets	$E[\sigma^2_{A_g}(t)]$	Data*	Model with* uncertainty	Model without* uncertainty	Elasticity brand-name
0	14	31.04	4.02	2.02	19.06	0.74
1	14	24.01	7.04	4.02	26.27	0.77
2	14	19.64	8.43	6.48	29.05	0.80
3	14	15.71	10.41	9.10	31.33	0.83
4	14	12.75	11.21	11.55	32.34	0.84
5	14	10.56	12.72	13.75	32.93	0.86
6	14	8.88	13.44	16.11	33.89	0.87
7	14	7.63	14.14	16.83	32.95	0.88
8	14	6.73	15.58	17.77	32.64	0.90
9	14	6.03	15.33	18.35	32.21	0.92
10	14	5.49	16.96	18.51	31.12	0.92
11	14	5.05	17.35	18.68	30.50	0.94
12	13	4.65	18.94	19.75	31.70	0.93
13	12	4.54	20.13	20.88	32.67	0.92
14	10	5.07	22.51	22.12	35.39	0.84
15	10	4.82	22.43	21.17	33.40	0.85
16	10	4.61	23.32	20.21	31.27	0.86
17	8	5.18	23.11	18.79	31.70	0.87
18	5	5.96	21.99	17.76	33.68	0.76
19	5	5.79	19.16	17.27	31.85	0.78
20	3	7.80	32.35	25.12	47.98	0.73
21	3	7.62	32.32	23.81	45.16	0.75
22	2	5.89	11.86	9.76	10.37	0.75
23	1	0.13	21.24	17.01	17.13	0.92
24	1	0.13	20.52	15.81	15.93	0.94

Notes:
quarter 0 refers to the quarter in which generics just entered the market.
*-column 4,5,6 are generic sales (million number of patient days).

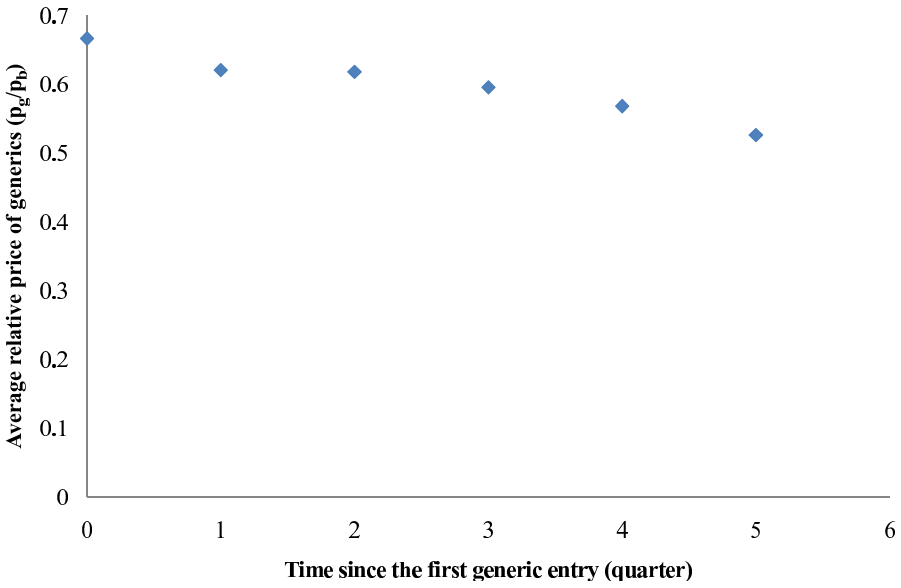

Figure 1: Average relative prices of generics vs time

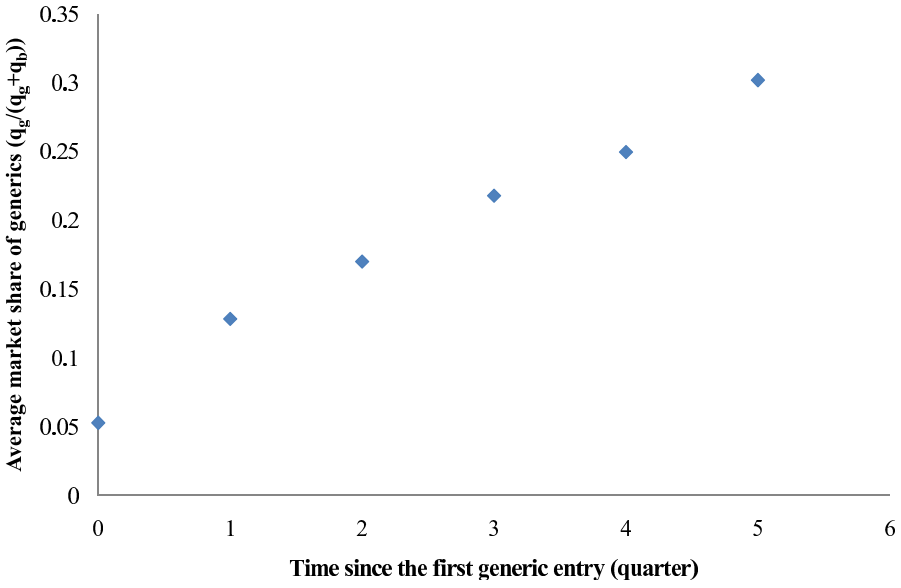

Figure 2: Average market shares of generics vs time

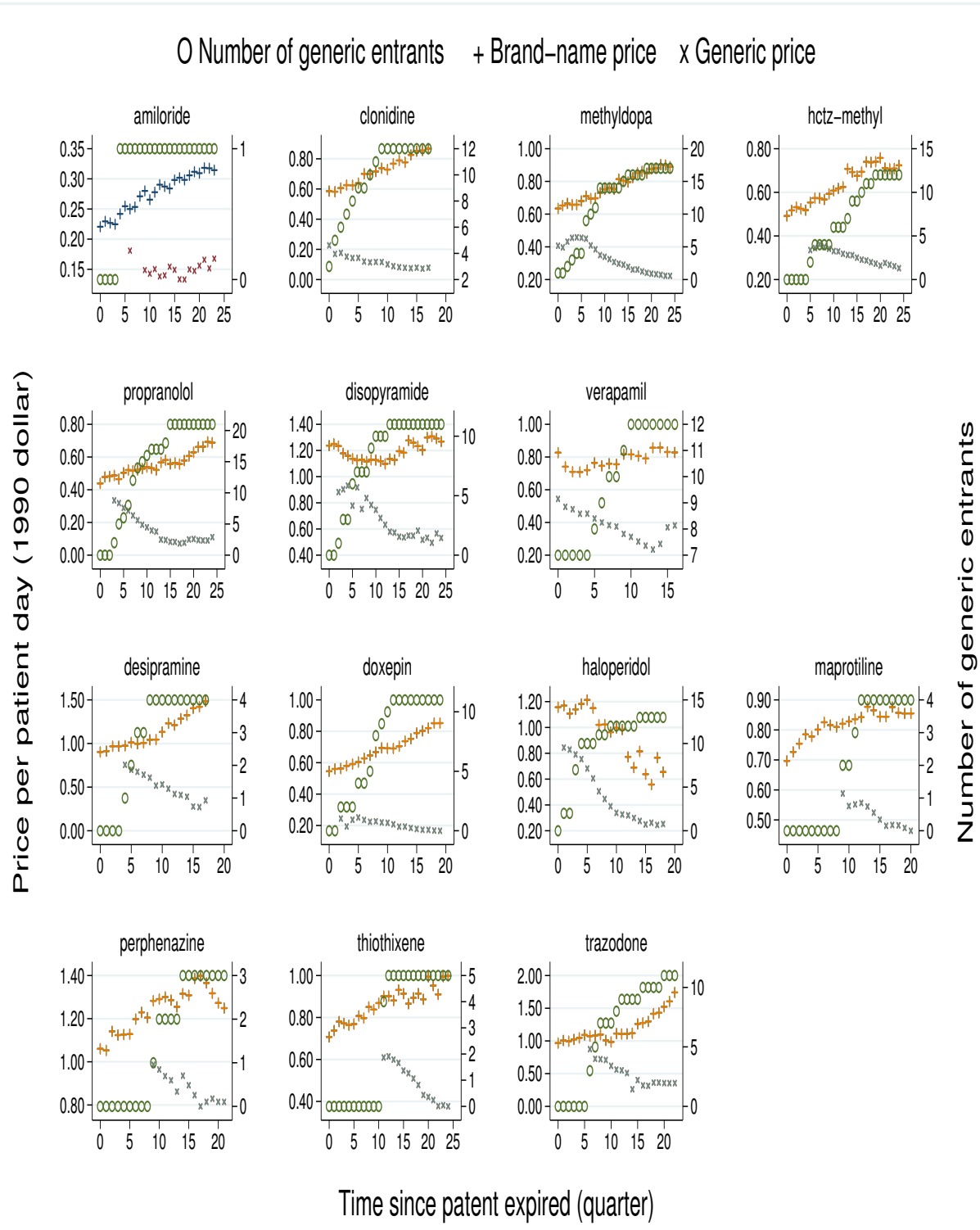

Figure 3: Number of generic entrants vs. time

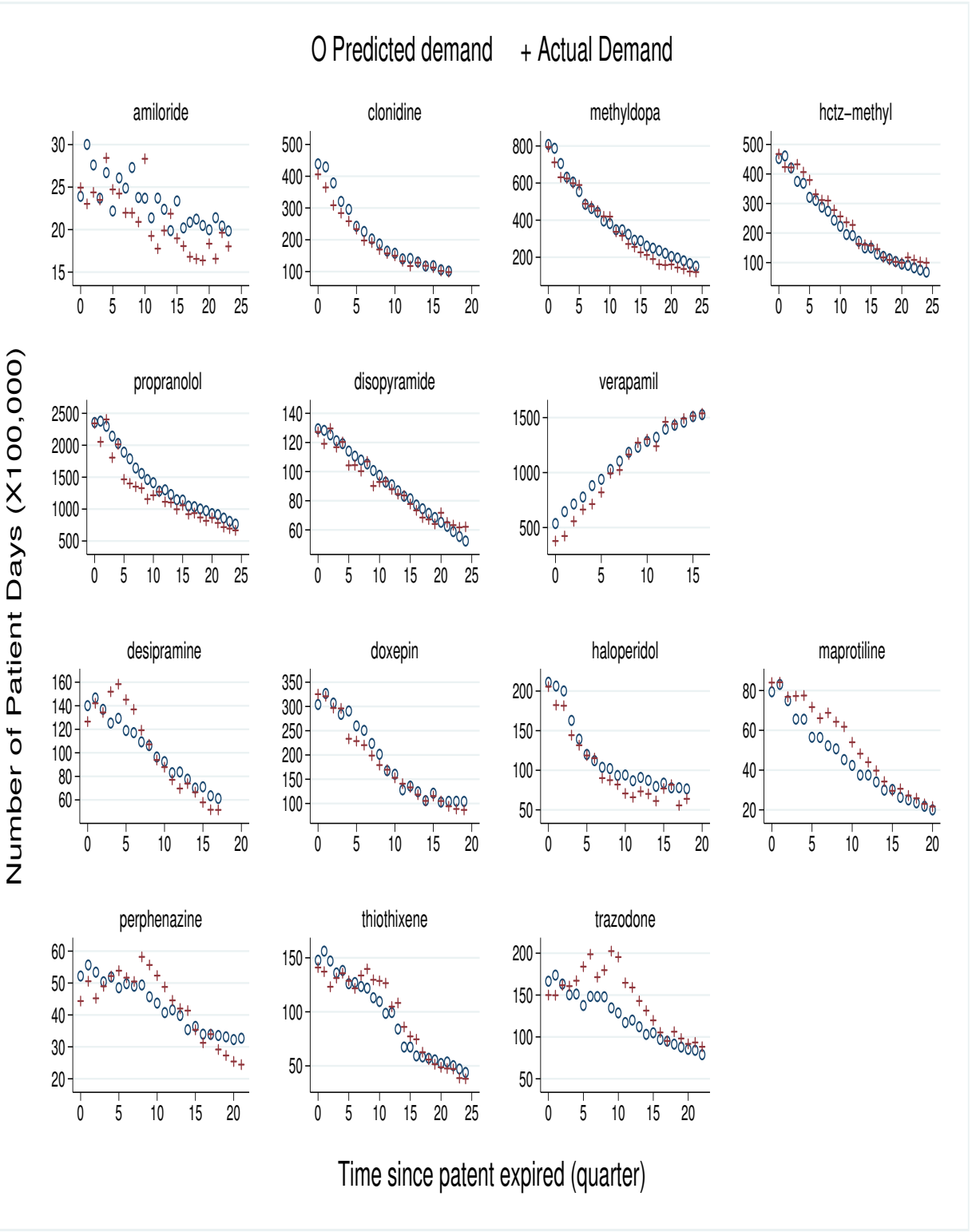

Figure 4: Predicted and actual demand for brand-name drugs (prices generated by the reduced form pricing functions)

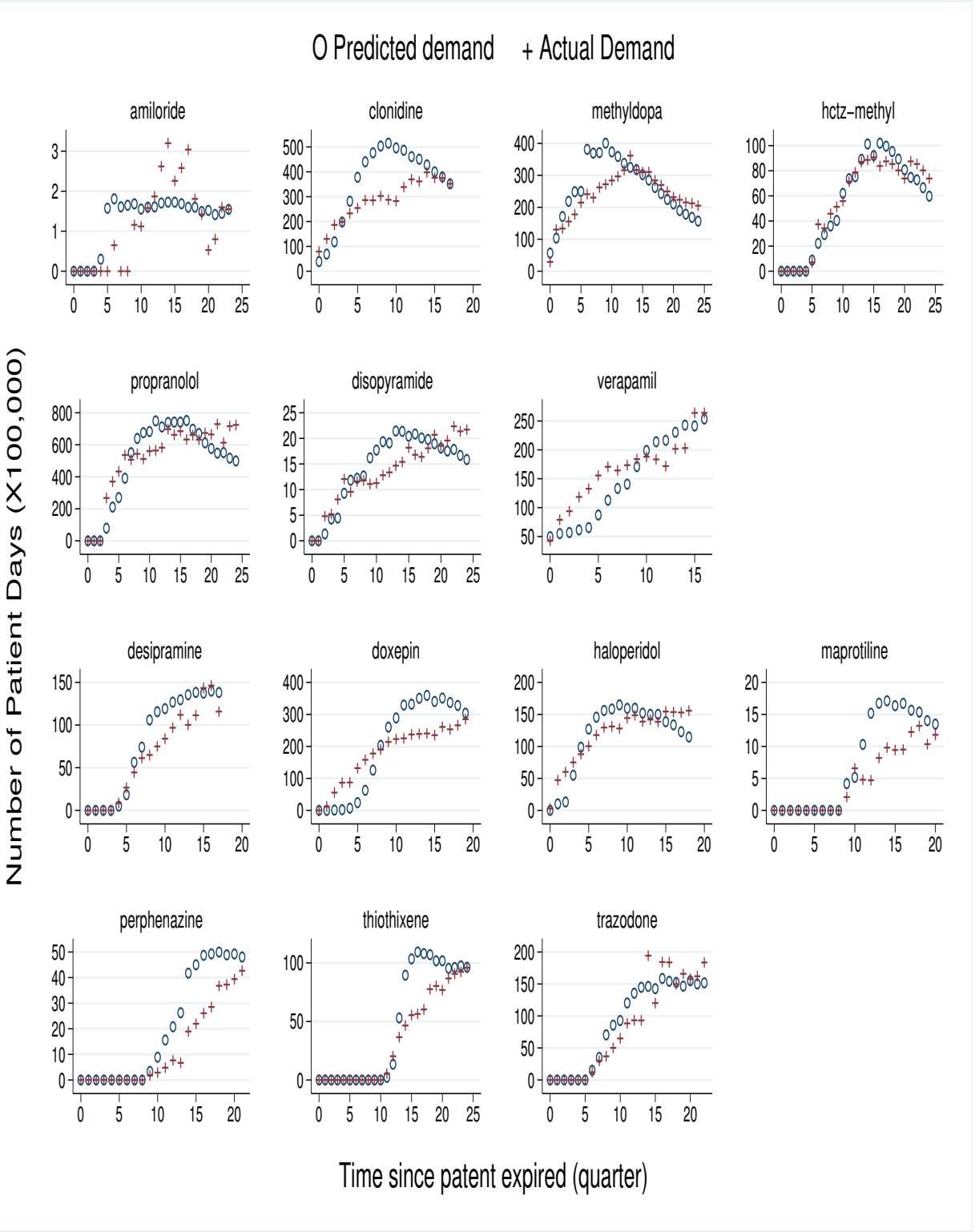

Figure 5: Predicted and actual demand for generic drugs (prices generated by the reduced form pricing functions)

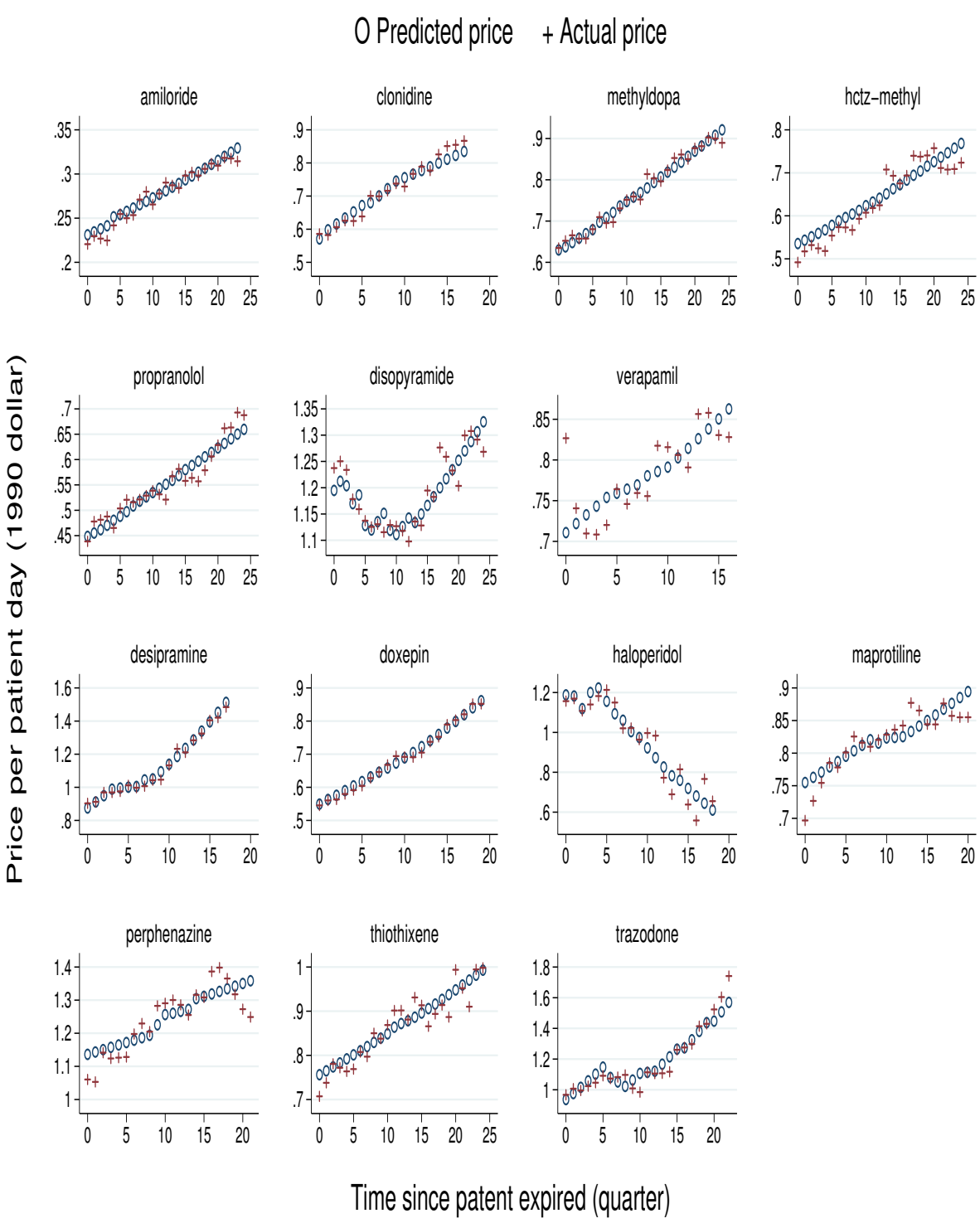

Figure 6: Predicted and actual price for brand-name drugs

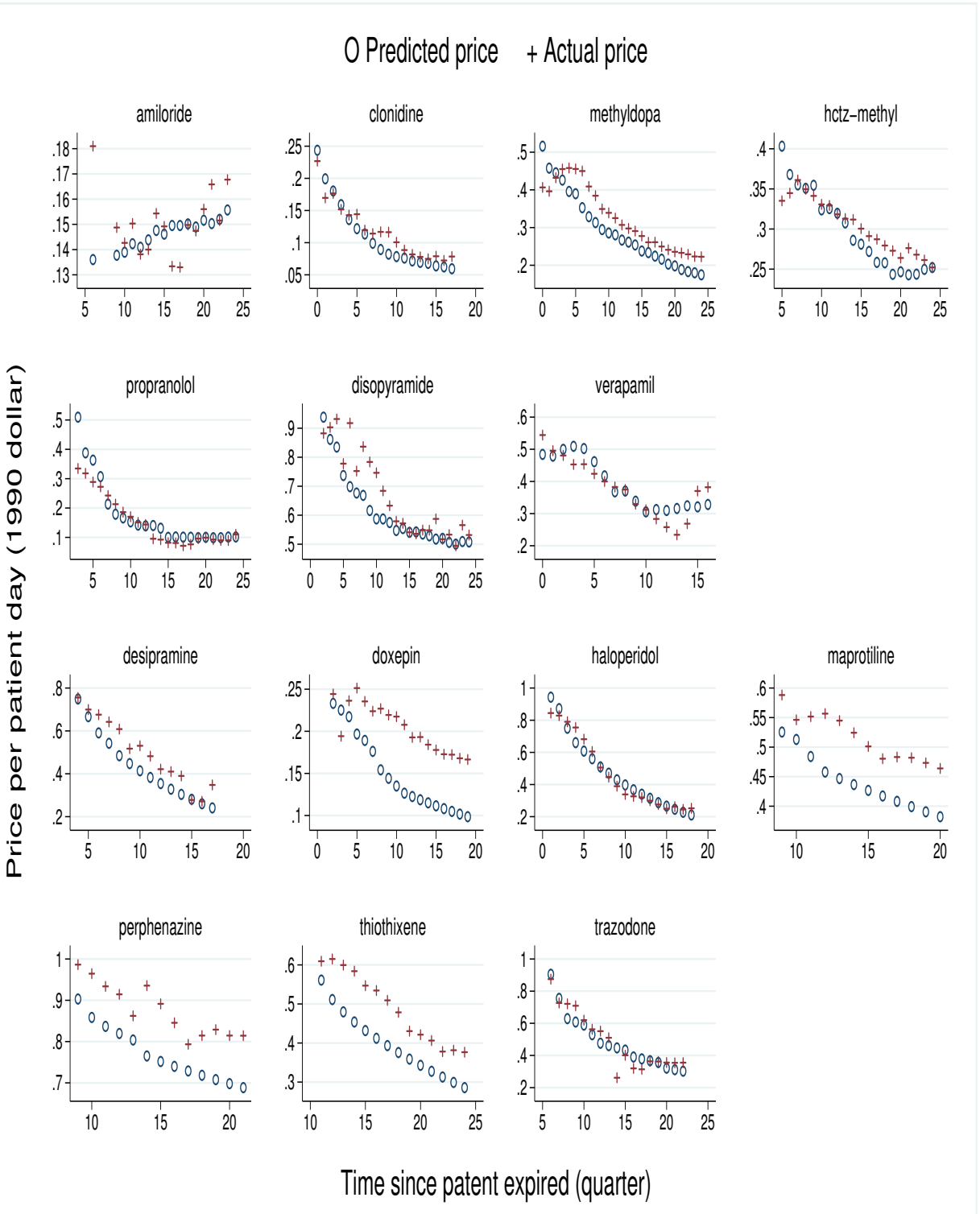

Figure 7: Predicted and actual price for generic drugs

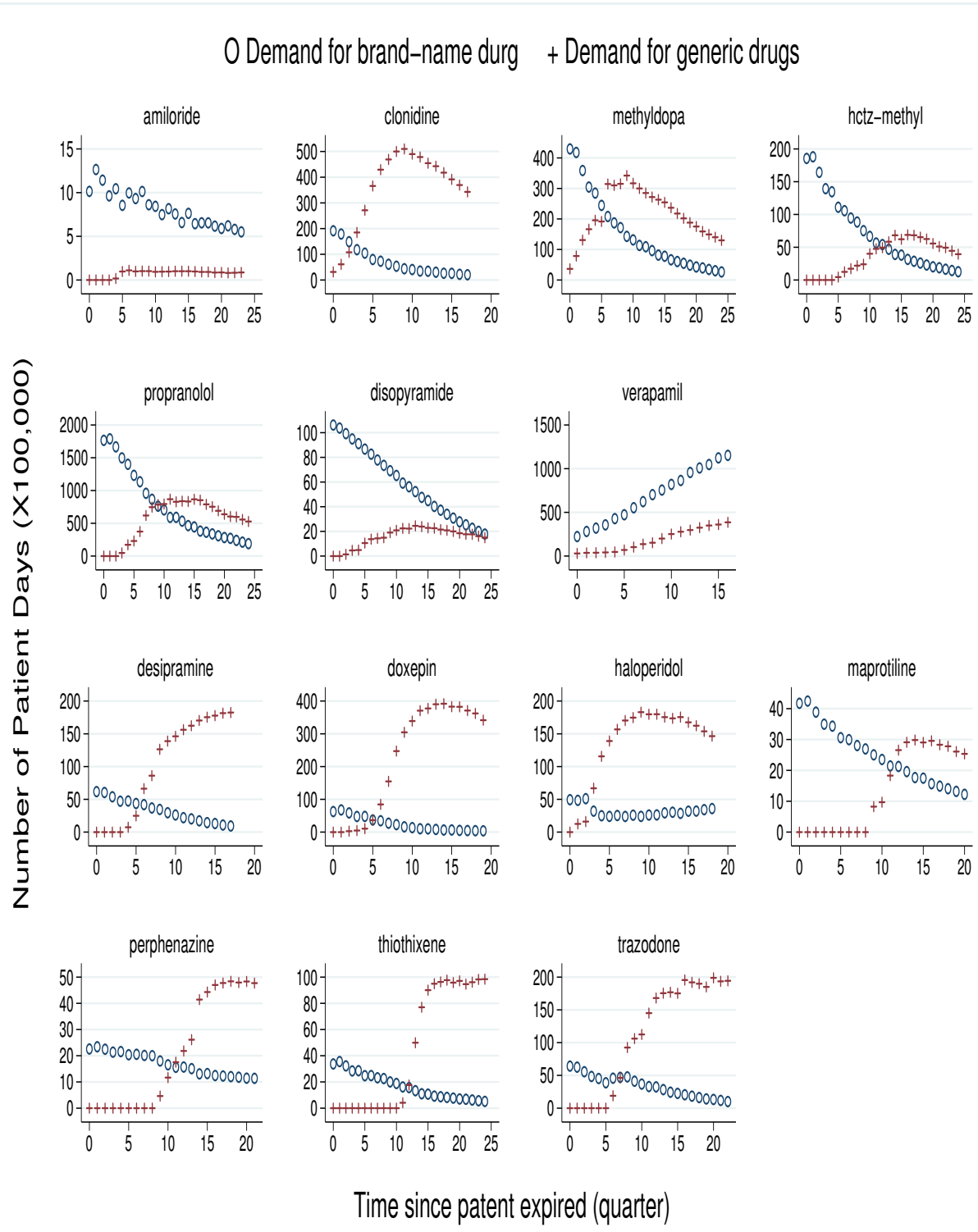

Figure 8: Predicted demand of price-sensitive patients

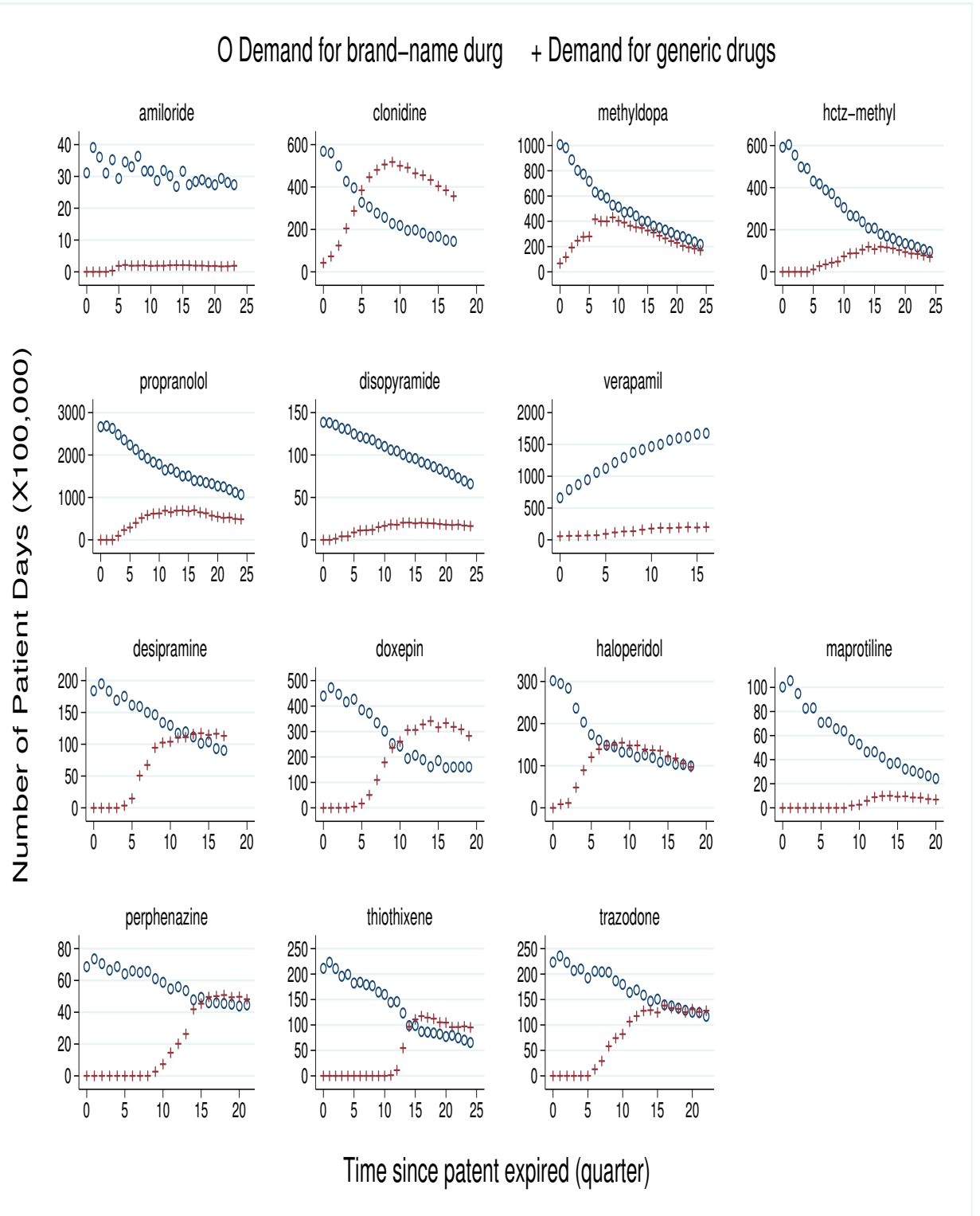

Figure 9: Predicted demand of price-insensitive patients

APPENDIX

In the simulation approach, one uses Monte Carlo methods to simulate the high order integrals that enter the likelihood function rather than evaluating them numerically (Pakes[45], Lerman and Manski[34], McFadden[40], Pakes and Pollard[46], Keane[32]). To obtain the simulated likelihood for (q_t, p_t), I first make D_ξ draws of (ξ_t^s) and D_A draws of $\{E[A_g|I(t)]^r\}_{t=0}^T$ from their distributions $F(\xi_t)$ and $F(\{E[A_g|I(t)]\}_{t=0}^T)$, respectively, where the superscript s and r distinguish the simulated values from the actual values. Then the simulated likelihood can be obtained by averaging the conditional likelihood over all of the simulated sets of unobservables,

$$L(q,p|\{n_{g\tau}\}_{\tau=1}^T;\theta_d,\gamma) \simeq \frac{1}{D_A}\sum_{r=1}^{D_A}\{\prod_t^T[\frac{1}{D_\xi}\sum_{s=1}^{D_\xi} l(q_t,p_t|n_{gt},t,\{q_{g\tau}\}_{\tau=0}^{t-1},E[A_g|I(t)]^r,\xi_t^s;\theta_d,\gamma)]\}. \tag{29}$$

It is worth emphasizing how to obtain a sequence $\{E[A_g|I(t)]^r\}_{t=0}^T$. Recall Equation (5),

$$E[A_g|I(t+1)] = E[A_g|I(t)] + \beta_g(t)(\bar{A}_{gt} - E[A_g|I(t)]),$$

where

$$\begin{aligned} \bar{A}_{gt} &= A_g + \epsilon_t \frac{\sigma_\delta}{\sqrt{\kappa q_{gt}}}, \\ \epsilon_t &\overset{iid}{\sim} N(0,1). \end{aligned} \tag{30}$$

Hence, $E[A_g|I(t)]$ is a function of $\{q_{g\tau}\}_{\tau=0}^{t-1}$ and $\{\epsilon_\tau\}_{\tau=0}^{t-1}$. To draw a sequence $\{E[A_g|I(t)]^r\}_{t=1}^T$, I first draw a sequence $\{\epsilon_t^r\}_{t=0}^{T-1}$, which in turn can generate a sequence of sample means of experience signals, $\{\bar{A}_{gt}^r\}_{t=1}^T$ (Equation (30)). Using this sequence $\{\bar{A}_{gt}^r\}_{t=1}^T$ and the Bayesian updating formula for $E[A_g|I(t)]$ (Equation (5)), one can recursively generate a sequence $\{E[A_g|I(t)]^r\}_{t=1}^T$.

It should be noted that the sampling errors for quantities demanded (η) and the prediction errors for prices (ν) serve the function of kernel smoothers in forming the simulated likelihood function. For each draw of the unobservables $(\xi_t^s, E[A_g|I(t)]^r)$, the conditional likelihood $l(q_t,p_t|.)$ in Equation (29) is differentiable and assigns positive density to any value of quantity demanded and price. They also smooth the likelihood function so that it is differentiable with respect to the parameters. Consequently, it is possible to maximize the likelihood using optimization techniques based on derivative information.

However, when the variance of sampling errors for quantity demanded is very small, the simulated likelihood function will closely approximate a step function of the parameters. This could create problems when applying a derivative-based optimization algorithm. In estimating the model, I found that the sampling errors were indeed too small for a derivative-based search algorithm to work well. So I have increased the variances of the sampling errors by multiplying them by constants in order to smooth the likelihood function.[33] This approach is similar to the kernel-smoothed frequency simulator for a discrete response model suggested by McFadden[40]. The values of the constant terms are reported in the last column of Table 2.

As in all kernel smoothing, there is a tradeoff. A higher degree of smoothing makes the search algorithm performs better, but also induces bias. Hence the degree of smoothing should be as small as possible while permitting the search algorithm to perform well. The reported scaling factors are the smallest that I can achieve after trials and errors.

[33]This is equivalent to adding one normal error term to each pricing equation, and assuming that error term to be independent across equations.

References

[1] Daniel Ackerberg and Marc Rysman. Unobserved Product Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects. *Rand Journal of Economics*, 36(4):771–788, 2005.

[2] Frank M. Bass. A New Product Growth Model for Consumer Durables. *Management Science*, 15(5):215–227, 1969.

[3] Ernst Berndt, Linda T. Bui, David H. Reiley, and Glen L. Urban. The Role of Marketing, Product Quality and Price Competition in the Growth and Composition of the U.S. Anti-Ulcer Drug Industry. In *The economics of new goods, NBER Studies in income and wealth*, volume 58, pages 277–322. University of Chicago Press, Chicago and London, 1997.

[4] Ernst Berndt, Robert Pindyck, and Pierre Azoulay. Consumption Externalities and Diffusion in Pharmaceutical Markets: Antiulcer Drugs. *The Jounral of Industrial Economics*, 51(2), 2003.

[5] Steven Berry and Ariel Pakes. The Pure Hedonic Discrete Choice Model. International Economics Review, forthcoming, 2007.

[6] Steven T. Berry. Estimating discrete-choice models of product differentiation. *Rand Journal of Economics*, 25(2):242–262, Summer 1994.

[7] Steven T. Berry, James Levinshohn, and Ariel Pakes. Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–890, July 1995.

[8] N. Scott Cardell. Variance Components Structures for the Extreme-value and Logistic Distribution with application to Models of Heterogeneity. *Econometric Theory*, 13:185–213, 1997.

[9] Norman V. Carroll, Chanaporn Siridhara, and Jack E. Fincham. Perceived Risks and Pharmacists' Generic Substitution Behavior. *Journal of Consumer Affairs*, 20(1):36–47, 1986.

[10] Norman V. Carroll and Alan P. Wolfgang. Risks, Benefits, and Generic Substitution. *Journal of Consumer Affairs*, 25(1):110–121, 1991.

[11] Richard E. Caves, Mihchael D. Whinston, and Mark A. Hurwitz. Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry. *Brookings Papers on Economic Activity, Microeconomics*, 1:1–48, 1991.

[12] Andrew Ching. *Dynamic Equilibrium in U.S. Prescription Drug Market After Patent Expiration*. PhD thesis, University of Minnesota, August 2000.

[13] Andrew Ching. A Dynamic Oligopoly Structural Model for the Prescription Drug Market After Patent Expiration. mimeo, Department of Economics, The Ohio State University, 2002.

[14] Andrew Ching and Masakazu Ishihara. The Effects of Detailing on Prescribing Decisions Under Two-sided Learning. working paper, Rotman School of Management, University of Toronto, 2007.

[15] Andrea Coscelli and Matthew Shum. An empirical model of learning and patient spillover in new drug entry. *Journal of Econometrics*, 122:213–246, 2004.

[16] Gregory S. Crawford and Matthew Shum. Uncertainty and Experimentation in Pharmacetical Demand. *Econometrica*, 73(4):1135–1174, July 2005.

[17] Gillian R. Currie and Sangin Park. Estimating the Effects of Marketing and Consumption Experience on Demand for Antidepressant Drugs. mimeo, Department of Economics, Yale University, 1996.

[18] Morris H. DeGroot. *Optimal Statistical Decisions*. McGraw-Hill, New York, 1970.

[19] P.L. Doering, W.L. Russell, W.C. McCormick, and D.L. Klapp. Therapeutic substitution in the health maintenance organization outpatient environment. *Drug Intelligence and Clinical Pharmacy*, 22:125–130, 1988.

[20] Glenn Ellison and Sara Fisher Ellison. Strategic Entry Deterrence and the Behavior of Pharmaceutial Incumbents Prior to Patent Expiration. mimeo, Department of Economics, MIT, 2007.

[21] Sara F. Ellison, Iain Cockburn, Zvi Griliches, and Jerrry Hausman. Characteristics of Demand for Pharmaceutical Products: An Examination of Four Cephalosporins. *Rand Journal of Economics*, 28(3):426–446, Autumn 1997.

[22] Tülin Erdem and Michael P. Keane. Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent consumer Goods Markets. *Marketing Science*, 15(1):1–20, 1996.

[23] Tülin Erdem, Michael P. Keane, and Baohong Sun. Missing Prices and Coupon Availability Data in Scanner Panels: Correcting for the Self-selection bias in Choice Models Parameters. *Journal of Econometrics*, 89:177–196, 1999.

[24] Richard G. Frank and David S. Salkever. Generic Entry and the Pricing of Pharmaceutical. *Journal of Economics and Management Strategy*, 6(1):75–90, 1997.

[25] John F. Geweke and Michael P. Keane. Bayesian Inference for Dynamic Discrete Choice Models without the Need for Dynamic Programming. In Mariano, Schuermann, and Weeks, editors, *Simulation Based Inference and Econometrics: Methods and Applications*, pages 100–131. Cambridge University Press, 2000.

[26] Henry Grabowski and John Vernon. Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act. *Journal of Law and Economics*, 35:331–350, 1992.

[27] Zvi Griliches and Iain Cockburn. Generics and New Goods in Pharmaceutical Price Indexes. *American Economic Review*, pages 1213–1232, December 1994.

[28] James Heckman. Heterogeneity and State Dependence. In *Studies in Labor Markets*, pages 91–139. University of Chicago Press, Chicago and London, 1981.

[29] James Heckman and B. Singer. A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, 52:271–320, 1984.

[30] Judith Hellerstein. The importance of the Physician in the Generic Versus Trade-name Prescription Decision. *Rand Journal of Economics*, 28(3):108–136, Spring 1998.

[31] Barrie H. James. *The Marketing of Generic Drugs*. Associated Business Press, London, 1981.

[32] Michael Keane. A Computationally Practical Simulation Estimator for Panel Data. *Econometrica*, 62(1):95–116, 1994.

[33] K. J. Lancaster. *Consumer Demand: A New Approach*. Columbia University Press, New York, 1971.

[34] Steven Lerman and Charles Manski. On the Use of Simulated Frequencies to Approximate Choice Probabilities. In C. Manski and D. McFadden, editors, *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge, 1981.

[35] Tamar Lewin. Drug Makers Fighting Back Against Advance of Generics. *New York Times*. July 28, 1987.

[36] Robert Lucas. Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy*, 1:19–46, 1976.

[37] Vijay Mahajan, Eitan Muller, and Frank M. Bass. New Product Diffusion Models in Marketing: A Review and Directions for Research. *Journal of Marketing*, 54(1):1–26, 1990.

[38] J. Barry Mason and William O. Bearden. Generic Drugs: Consumer, Pharmacist and Physician Perception of the Issues. *Journal of Consumer Affairs*, 14(1):192–205, 1980.

[39] Alison Masson and Robert L. Steiner. Generic Substitution and Prescription Drug Prices: Economic Effects of State Drug Product Selection Laws, 1985.

[40] Daniel McFadden. A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. *Econometrica*, 57:995–1026, 1989.

[41] Michelle Meadows. Greater Access to Generic Drugs: New FDA initiatives to improve drug reviews and reduce legal loopholes. http://www.pueblo.gsa.gov/cic_text/health/access-gen-drugs/access-gen-drugs.htm.

[42] MedicineNet.com. Medications. http://www.medicinenet.com.

[43] Bureau of Labor Statistics. Pharmaceutical and Medicine Manufacturing. http://www.bls.gov/oco/cg/cgs009.htm.

[44] G. Steven Olley and Ariel Pakes. The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64:1263–1297, 1996.

[45] Ariel Pakes. Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica*, 54:755–784, 1987.

[46] Ariel Pakes and David Pollard. Simulation and the Asymptotics of Optimization Estimates. *Econometrica*, 57:1027–1058, 1989.

[47] Amil Petrin. Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy*, 110:705–729, 2002.

[48] Fiona M. Scott Morton. Entry Decisions in the Generic Pharmaceutical Industry. *Rand Journal of Economics*, 30(3):421–440, 1999.

[49] Fiona M. Scott Morton. Barriers to Entry, Brand Advertising, and Generic Entry in the U.S. Pharmaceutical Industry. *International Journal of Industrial Organization*, 18:1085–1104, 2000.

[50] Pamela Sherrid. The New Drug War. *US News and World Report*, December 2002.

[51] Scott Stern. Innovation and Market Definition: Substitution Patterns in Pharmaceutical Markets. mimeo, Sloan School of Management, MIT, 1996.

[52] H. David Strutton, James R. Lumpkin, and Scott J. Vitell. The Elderly's Perceptions of the Risk of Generic Over-the-Counter Medication. *Journal of Research in Pharmaceutical Economics*, 4(3):25–39, 1992.

[53] Dong-Churl Suh, Stephen W. Schondelmeyer, Willard G. Manning, Ronald S. Hadsall, and John A. Nyman. Price Trends Before and After Patent Expiration in the Pharmaceutical Industry. *Journal of Research in Pharmaceutical Economics*, 9(2):17–31, 1998.

[54] Office of Technology Assessment U.S. Congress. Pharmaceutical R & D: Cost, Risks and Rewards. Washington, 1993.