Innovation-and-Learning-Adjusted Price Indexes for Prescription Drugs[**]

Sangin Park[*]

Abstract : In general, prescription drugs have an unobservable quality prior to consumption, which is called an 'experience characteristic'. Consumers learn these experience characteristics from both consumption experience and advertising exposure. Based on the Bayesian learning process of experience characteristics and the characteristics approach to demand functions, this paper proposes innovation-and-learning-adjusted price indexes for prescription drugs. This structural approach not only resolves the quality adjustment of new molecules but also avoids arbitrary assumptions on the link-in of generic drugs to the originator branded drug. The suggested price indexes are applied to the data for antidepressant drugs during the years 1980-1995. We have found: (i) the average annual growth rate of the focal price index is about -9.5%, which suggests that the existing price indexes for prescription drugs may seriously overstate the rate of inflation in a rapidly growing market with the entry of innovative products; and (ii) consumers' learning about experience characteristics were substantial especially after active generic entry in 1986 and the entry of Prozac in 1988.

Keywords : price indexes; experience characteristics; link-in problem; learning; antidepressant drugs.

JEL Classification: C43; L65.

INTRODUCTION

Currently price indexes for prescription drugs are a major issue in the U.S. health care reform. There are several reasons that the construction of price indexes is a challenging task in the case of prescription drugs. In the pharmaceutical industry, as in many other important industries, technological progress takes the form of new products: There are usually significant entries of new molecules and generic entry after patent expiration in a certain therapeutic class. In general, it is not easy to construct price indexes which truthfully reflect the changes in consumer welfare that result from the entry of new products (see Hausman (1997)). The conventional Laspeyres price index fails to reflect the entry of new products, while the conventional Paasche price index requires measuring reservation prices of new products before market introduction. Measuring these reservation prices of new products is a difficult task. One possibility is to predict these reservation prices based on the estimates of hedonic price indexes (see Berndt, Griliches

and Rappaport (1995); and Berndt, Cockburn and Griliches (1996)).[1] Although the hedonic price indexes can lead to a lower bound for consumer welfare gains from new products as shown in Pakes (2001), hedonic price indexes do not fully reflect the quality adjustment of new products in imperfect competition (see Trajtenberg (1990)). A characteristics approach to the demand system is needed to go further.

In addition to this general problem, in the case of prescription drugs, the entry of generics makes things more complicated. As noted in Griliches and Cockburn (1994), the estimates of hedonic price indexes cannot be used as a prediction of the reservation price of a generic drug since a generic drug is 'therapeutically equivalent' to the originator branded drug and thus has the same observable characteristics even with a substantially lower price than its originator branded drug.[2] Hence, how to link in a generic drug to its originator branded drug is an important and challenging issue in the case of prescription drugs.[3] Until recently, the Bureau of Labor Statistics (BLS) calculated the Producer Price Indexes (PPIs), treating generic drugs as entirely distinct and non-substitutable products. The Food and Drug Administration (FDA), on the other hand, argues the other extreme: "a pill, is a pill, is a pill." In this case, the relevant price for a molecule will be the weighted average price of all generic drugs and the branded drug within the molecule. Griliches and Cockburn (1994) tackled the link-in issue, assuming a linear utility framework and a uniform distribution function on consumers' different tastes for 'brandedness'. They also noticed a diffusion problem: Without much change in the price differential, the share of generics significantly increased between six months and a year after generic entry. Combining the assumption of consumers' tastes for brandedness with the observation of the diffusion problem, they calculated the reservation price of a generic drug before the market introduction as the average six-month-later price of the generic drug and the originator branded drug.

In this paper, we propose structural-model-based price indexes for prescription drugs, which incorporate the link-in of a generic drug to the originator branded drug and the entry of a new molecule into a characteristics approach to the consumer's rational choice problem of experience goods. In general, prescription drugs, branded or generic drugs, have an unobservable quality prior to consumption, which is called an 'experience characteristic' (see Currie and Park (2000)). The consumer (the physician-patient pair, with the patient making the choice upon the advice of his/her physician) will learn the experience characteristics of drugs from both consumption experience and advertising exposure. Prescription drugs are usually advertised in two forms: printed advertisements in medical journals and detailing (face-to-face visits to physicians by pharmaceutical company representatives).[4] Both types of advertising serve to inform physicians of the existence of a new product and provide the information on the product's attributes.

---

1) An alternative approach is based on a constant-elasticity-of-substitution utility function for the variety of differentiated products (see Feenstra (1996)). This approach takes the reservation price as infinite, and calculates a finite consumer surplus from the introduction of a new variety.

2) The Food and Drug Administration requires generic drugs to have the exact same active ingredients in the same form and concentration as the originator branded drug.

3) For example, Griliches and Cockburn (1994) found significant differences in price indexes for anti-infective drugs across different assumptions on the link-in.

4) Very recently, advertisements directed at patients, such as TV commercials has begun.

Moreover, a firm's advertising level could be a signal to consumers that the product has high experience characteristic quality. As Nelson (1970) argued, a high-quality product is more likely to attract repeat purchases. Hence, an initial sale is more valuable to a high-quality producer who would be willing to advertise more to attract an initial sale. The Nelson insight is very appealing in the case of prescription drugs since the pharmaceutical producers typically put most of their marketing effort on the introduction of a new molecule, and then steadily decreases their marketing to nothing. By adopting a Bayesian learning process of experience goods as in Erdem and Keane (1996) and Ackerberg (1997), we will model a mechanism in which consumers update expectations for experience characteristics through consumption experience and advertising exposure.

Hence, in the paper, the link-in of a generic drug to the originator branded drug is treated in the context of informational product differentiation. The branded drug usually has an advantage over generics in both consumption experience and advertising, and thus rational consumer behavior can give pioneering brands advantages. Indeed, this informational product differentiation between generic drugs and the originator branded drug has been recognized in the literature (see Schmalensee (1982)). Although generics are therapeutically equivalent to the originator branded drug, there is some possible difference in the inactive ingredients, shelf life, etc. that can affect the quality of the generics. However, as generics accumulate more sales and consumption experiences, the informational advantage of the branded drug diminishes. In this sense, our model articulates and integrates the consumers' different tastes for brandedness, and the diffusion problem in Griliches and Cockburn (1994) in the context of the informational product differentiation and a learning process.

To reflect that prescription drugs are differentiated by observable product characteristics such as side effects as well as by (expected) experience characteristics, we model the consumer's rational choice for prescription drugs, using the nested logistic assumptions as in Berry (1994) and Stern (1996). Then we will be able to derive a closed-form expenditure function and to calculate ideal price indexes. These ideal price indexes are called *innovation-and-learning-adjusted price indexes* in the paper. Since we obtain the expenditure function reflecting entry of new products and consumers' learning about experience characteristics, we can avoid arbitrary assumptions both on the link-in of generic drugs to the branded drug and on the reservation prices of new products prior to the market introduction. Hence, the innovation-and-learning-adjusted price indexes truthfully quantify changes in consumer welfare that result from the entry of new products as well as consumers' learning about experience characteristics in both the framework of the characteristics approach to the demand system and a Bayesian learning process of experience characteristics.

We will apply the innovation-and-learning-adjusted price indexes to the data for antidepressant drugs during the years 1980-1995. During these years, the market for antidepressants experienced 'exceptional and remarkable' innovations in terms of entry of both new molecules and generics. In 1980 and 1981, the second generation of antidepressant drugs called tricyclic antidepressants (TCAs) entered the market, replacing the first generation of TCAs, most of which were introduced in 1960's. Beginning in 1986, there was an active entry of generic drugs induced by the passage of the 1984 Waxman-Hatch Act. Most importantly, the breakthrough drug, Prozac, was introduced

to the market in 1988, and subsequently four more drugs in the same therapeutic subclass entered by the year 1995. The calculated innovation-and-learning-adjusted price indexes confirm the occurrence of exceptional and remarkable innovations during these years: the Average Annual Growth Rate (AAGR) of our focal price index is about -9.5 percent. It will be also shown that consumers' learning about experience characteristics were substantial especially after active generic entry in 1986 and the entry of Prozac in 1988. We will further compare our innovation-and-learning-adjusted price indexes with the other existing indexes such as the old BLS PPI, the new BLS PPI and the Griliches-Cockburn adjusted Paasche Diffusion (GCPD) method. This comparison will suggest that the existing price indexes for prescription drugs may seriously overstate the rate of inflation in a rapidly growing market with the entry of innovative products.

In the sense that the innovation-and-learning-adjusted price indexes quantify consumers' learning about experience characteristics, these indexes do not depend solely on the principle of commodity substitution and materialize the idea of the outlet substitution problem recognized by Reinsdorf (1993). The outlet substitution problem in consumer price index estimation is induced by the rapid growth of low-price outlets and is now considered one of the most important issues in the research on price index measurement (see Berndt (1999)). The significance of consumers' learning effects in the case of antidepressant drugs supports the concern of the outlet substitution problem and indicates that the correct formula for price indexes may not be determined solely by the principle of commodity substitution.

The remainder of the paper is organized as follows. Section 2 will describe the model for consumers' decision-making and the learning process of experience characteristics. Section 3 will construct the innovation-and-learning-adjusted price indexes based on our structural approach. Section 4 will calculate these price indexes for antidepressant drugs during the years 1980-1995, which will be further compared with other indexes discussed in Berndt, Cockburn, and Griliches (1996). Section 5 will conclude the paper.

THE MODEL

A. Nested Logistic Assumptions

In general, prescription drugs have an unobservable quality prior to consumption, which is called an experience characteristic in Currie and Park (2000). Usually, prescription drugs of a certain therapeutic class are differentiated by observable product characteristics such as side effect profiles as well as by experience characteristics. For example, the various antidepressant drugs have almost the same efficacy rates but different side effect profiles (See Depression Guideline Panel, 1993). The indexes of severity of side effects, however, are infrequently updated and too unrefined to provide information that is more distinctive. For example, Prozac is considered as a break-through drug because it is not fatal in overdose and has fewer side effects, but the (official) indexes of severity of side effects do not give Prozac a significant edge over its main competitors (See Berndt, Cockburn, and Griliches (1996), Table 1). In addition,

generics and the originator branded drug have the same indexes of severity of side effects although there are possible differences in inactive ingredients, shelf life, etc., which may distinguish the quality of the generics from that of the originator branded drug. Hence, the insufficient information of the (official) indexes of severity of side effects can be considered as a main reason that consumers learn and anticipate experience characteristics based on available information.[5]

In this paper, a "product" means a branded version or a generic version of a molecule, and the "consumer" means the physician-patient pair, with the patient making the choice upon the physician's advice. Consequently, we do not explicitly consider the principle-agent problem between physician and patient. Rather, we will pay attentions to the physician's role in pooling the relevant information on experience characteristics in the learning process.

Taking account of product differentiation by observable product characteristics and experience characteristics, we follow the nested logistic assumptions to specify a consumer's utility level for a prescription drug. We begin with some notations. Let p_{jt} denote the price of product j in period t, and X_{jt} denote a vector of its observable product characteristics. Let ψ_{ijt} denote consumer i's utility of product j's experience characteristic in period t, and $\varepsilon_{it}(g, m, j)$ denote consumer i's idiosyncratic taste for product j of molecule m in therapeutic subclass g. This consumer's idiosyncratic taste reflects the anticipated patient-drug specific interactions and other patient-specific characteristics. We assume that the consumer's experience utility ψ_{ijt} is distributed as: $\psi_{ijt} = \delta_j + \eta_{ijt}$, where δ_j is the experience characteristic of product j, and η_{ijt} is consumer i's unexpected idiosyncratic experience in period t and has the zero mean conditioned on his/her information set, I_{it}. We treat the experience characteristic δ_j as a random variable, taking into account that this quality may vary across individual products and the consumers can randomly get "lemons" or "windfalls." Then, in period t, the consumer's expected (prior-to-consumption) utility for product j of molecule m in therapeutic subclass g is specified by the nested logistic assumptions as follows:

$$E[U_{ijt}|I_{it}] = y_t + \theta_1 - \theta_2 P_{jt} + X_{jt}\theta_3 + E[\delta_j|I_{it}] + \varepsilon_{it}(g, m, j)$$
with
$$\varepsilon_{it}(g, m, j) = \zeta_{igt} + (1- \sigma_g)\,\varsigma_{imt} + (1- \sigma_m)\,\varepsilon_{ijt},$$

(1)

where y_t is the average income, ζ_{igt}, ς_{imt} and ε_{ijt} denote consumer i's idiosyncratic tastes for therapeutic subclass g, molecule m, and product j, respectively, and σ_m and σ_g are parameters which have values greater than or equal to zero and less than one. If σ_m (σ_g) has a value closer to one, then the products (molecules) within a molecule (subclass) are considered more homogenous. Note that the consumer's utility is additively separable in y_t, and thus income effects are assumed away in the nested logit model. Note also that the coefficient of y_t is normalized to be one (instead of 'βy_t' ($\beta > 0$) in (1)). Since income effects are assumed away in the model, we cannot estimate the coefficient of y_t, but this normalization will not affect the calculation of the innovation-and-learning-adjusted

---

5) Crawford and Shum (2000), based on patient-level panel data on antiulcer drug prescriptions, found that patients quickly learn the effectiveness of drugs.

price indexes defined in section 3. An individual consumer's deviation from the average income y_t and his/her health insurance status are understood to be reflected in the consumer's idiosyncratic tastes $\varepsilon_{it}(g, m, j)$.[6]

This nested logistic specification very well fits the consumer's choice of prescription drugs for several reasons. First, there is well-defined daily dosage for prescription drugs, and thus the idea of discrete choice can be comfortably applied. Second, the consumer's choice of prescription drugs has a natural nesting procedure. As implied by the employed nested logistic assumptions, consumers first choose a relevant therapeutic subclass (or the outside alternative), then a molecule within the chosen subclass, and then a branded version or a generic version of the chosen molecule. Lastly, the consumer's idiosyncratic tastes, $\varepsilon_{it}(g, m, j)$, reflect anticipated patient-drug specific interactions (as well as patients' deviation from the average income and their health insurance status), and this interpretation is consistent with the idea of 'tastes for products' in Berry and Pakes (1999). The (additively separable) tastes for products in the typical logit model insure that some consumers will like a new product infinitely more than the existing products. Thus, as pointed out in Berry and Pakes (1999), the dimension of the product space increases with the number of products, and thus consumer welfare gains from product introductions may be over-evaluated in the typical discrete choice model. In the case of prescription drugs, however, the consumer's tastes for products with infinite supports make a sense since there are usually some consumers who prefer a certain drug due to their idiosyncratic (anticipated) patient-drug specific interactions.

B. *Learning Mechanism*

We now discuss a mechanism in which consumers update and form expectations of an experience characteristic, $E[\delta_j | I_{it}]$, in each period. The consumer will learn and anticipate an experience characteristic based on the history of the patient's consumption experience and the physician's prescription experience and advertising exposure. Note that a firm's advertising level may be a signal to consumers for its experience characteristic as discussed in literature of experience goods (See, for example, Nelson (1970); Milgrom and Roberts (1986); Ackerberg (1997)). We will take into account that the physicians can pool each other's information through journal articles, professional conferences, and informal communications. Berndt, Pindyck and Azoulay (2000) found evidence on this kind of informational diffusion at the brand level in the antiulcer drug market. Therefore, we assume that all consumers in a period receive the same history of consumption signals, say $\{\mu_{js}\}_{s \leq t}$, and the advertising signals, say $\{A_{js}\}_{s \leq t}$, for each product. In other words, under this information pooling assumption, each consumer's information set I_{it} contains the same history of consumption and advertising signals for all products. Hence, let $m^{\delta}_{jt} = E[\delta_j | I_{it}]$ for all consumers.

In our model, the idea that the advertising signal may provide indirect information about an experience characteristic is captured by the positive correlation between

---

6) The consumer-level information can be incorporated by extending the nested logistic specification to a random-coefficient model, as in Berry, Levinsohn and Pakes (1995).

experience characteristic, δ_j, and the mean level of advertising, say A_j. Like the experience characteristic δ_j, we will treat the mean level of advertising A_j as a random variable. We will assume that consumers learn and update m^{δ}_{jt} in a Bayesian fashion. Initially, consumers have the prior distribution of $(\delta_j, A_j)'$ for each product. In the first period, consumers have no consumption signal but receive the advertising signal from the current advertising level. However, in each following period, consumers will obtain a consumption signal, μ_{jt}, from the previous consumption as well as an advertising signal, A_{jt}. Consumers believe that μ_{jt} and A_{jt} indicate, on average, the experience characteristic, δ_j, and the mean level of the advertising, A_j, respectively. In other words, $\mu_{jt} = \delta_j + \nu_{jt}$, and $A_{jt} = A_j + \xi_{jt}$, where ν_{jt} and ξ_{jt} are interpreted as errors in the process of information pooling. Due to the correlation between δ_j and A_j, consumers will utilize advertising signals, $\{A_{js}\}_{s \le t}$, as well as consumption signals, $\{\mu_{js}\}_{s \le t}$, to update and anticipate m^{δ}_{jt}.

To facilitate the construction of a Bayesian learning process, we make the following two assumptions as in Ackerberg (1997). First, suppose that the errors in the information pooling process, ν_{jt} and ξ_{jt}, are *i.i.d.* normal random variables with zero means and variances of σ^2_{ν} and σ^2_{ξ}, respectively. Hence,

$$\begin{bmatrix} \mu_{jt} \\ A_{jt} \end{bmatrix} \sim i.i.d. N\left(\begin{bmatrix} \delta_j \\ A_j \end{bmatrix}, \begin{bmatrix} \sigma^2_{\nu} & 0 \\ 0 & \sigma^2_{\xi} \end{bmatrix} \right). \tag{2}$$

Second, suppose that all consumers have the same prior distribution of $(\delta_j, A_j)'$ for each product as follows:

$$\begin{bmatrix} \delta_j \\ A_j \end{bmatrix} \sim i.i.d. N\left(\begin{bmatrix} m^{\delta}_{j0} \\ m^{A}_{j0} \end{bmatrix}, \Sigma_0 \right) \text{with } \Sigma_0 = \begin{bmatrix} \sigma^2_1 & \sigma_{12} \\ \sigma_{12} & \sigma^2_2 \end{bmatrix}, \tag{3}$$

where Σ_0 is the initial covariance matrix of $(\delta_j, A_j)'$. Hence, if σ_{12} is positive, the advertising signal provides indirect information on δ_j via A_j.

Then using the theory of conjugate distributions, the posterior distribution of $(\delta_j, A_j)'$ is given by a normal distribution with a mean vector $(m^{\delta}_{jt}, m^{A}_{jt})'$ and a covariance matrix Σ_{jt} as follows:

$$\begin{bmatrix} m^{\delta}_{jt} \\ m^{A}_{jt} \end{bmatrix} = \Sigma_{jt}\left(\Sigma_0^{-1} \begin{bmatrix} m^{\delta}_{j0} \\ m^{A}_{j0} \end{bmatrix} + \begin{bmatrix} ((t_j-1)/\sigma^2_{\nu})\,\overline{\mu}_{jt} \\ (t_j/\sigma^2_{\xi})\,\overline{A}_{jt} \end{bmatrix} \right), \text{and} \tag{4}$$

$$\Sigma_{jt} = \left(\Sigma_0^{-1} + \begin{bmatrix} (t_j-1)/\sigma^2_{\nu} & 0 \\ 0 & t_j/\sigma^2_{\xi} \end{bmatrix} \right)^{-1},$$

$\overline{\mu}_{jt} = \Sigma^t_{s=t_{j0}+1}\mu_{js}/(t-t_{j0})$, where $\overline{A}_{jt} = \Sigma^t_{s=t_{j0}}A_{js}/(t-t_{j0}+1)$, and t_{j0} denotes the year when product j was introduced. Note that if δ_j and A_j are perfectly correlated, then we cannot obtain the inverse of Σ_0 in (4). Hence, the formula in (4) will hold only if $\sigma^2_{12}/(\sigma^2_1\sigma^2_2) \ne 1$. Note also that σ^2_{ν} and σ^2_{ξ} are assumed known to consumers in the conjugate distribution theory although this assumption can be avoided at a cost of complexity of

an updating rule.[7]

We now discuss how the Bayesian learning process that was specified above fits the typical sales and advertising pattern of prescription drugs. Typically, upon the introduction of a new branded drug, there are intensive marketing efforts, which diminish thereafter to zero (See Currie and Park (2000)).[8] On the other hand, the sales of the branded drug continue to rise and then begin to fall upon the entry of generics. Following Erdem and Keane (1996) and Ackerberg (1997), we assume that the consumption signal μ_{jt} is proportional to the sales in the previous period, say q_{jt-1} (that is, $\mu_{jt} = \theta_4 q_{jt-1}$, where θ_4 is called the parameter of consumption experiences in the paper). Let $t_j = t-t_{j0} + 1$. For simplicity, set the mean values of the prior distribution of δ_j and A_j to be zero, i.e., $m_{j0}^\delta = m_{j0}^A = 0$. Then equation (4) leads to:

$$\begin{bmatrix} m_{jt}^\sigma \\ m_{jt}^A \end{bmatrix} = \frac{t_j(t_j-1)(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)}{\left(\sigma_1^2 + \frac{t_j-1}{\sigma_\nu^2}\right)\left(\sigma_2^2 + \frac{t_j}{\sigma_\xi^2}\right) - \sigma_{12}^2} \begin{bmatrix} \left(\frac{\sigma_1^2}{t_j} + \frac{1}{\sigma_\xi^2}\right)\frac{\theta_4 \overline{q}_{jt-1}}{\sigma_\nu^2} + \frac{\sigma_{12}}{t_j-1}\frac{\overline{A}_{jt}}{\sigma_\xi^2} \\ \left(\frac{\sigma_2^2}{t_j-1} + \frac{1}{\sigma_\nu^2}\right)\frac{\overline{A}_{jt}}{\sigma_\xi^2} + \frac{\sigma_{12}}{t_j}\frac{\theta_4 \overline{q}_{jt-1}}{\sigma_\nu^2} \end{bmatrix}, \quad (5)$$

where $\overline{q}_{jt-1} = \Sigma_{s=t_{j0}}^{t-1} q_{js}/(t_j-1)$. Note that $\sigma_1^2\sigma_2^2 - \sigma_{12}^2 = \sigma_1^2\sigma_2^2(1 - \sigma_{12}^2/(\sigma_1^2\sigma_2^2)) > 0$. Suppose that producers have a certain targeting level (or range) of the experience characteristic that they want to signal.[9] Then equation (5) indicates that producers will put high initial advertising efforts to compensate for low initial consumption experiences, and then reduce their advertising efforts due to increased cumulative consumption. This interpretation is consistent with the insight of Nelson (1970), who argued that a high-quality product is more likely to attract repeat purchases; consequently, an initial sale is more valuable to a high-quality producer who would be willing to advertise more to attract an initial sale.[10] Equation (5) also indicates that the influence of advertising signal on the expected experience characteristic will phase out over time. Advertising signal will have persistent influences if δ_j and A_j are perfectly correlated in (3).[11]

To directly reflect that advertising intensities may compensate for low initial consumption experiences, we also consider an alternative specification of experience

---

7) In the actual of calculation of Section 4, we will use the sample variances of sales and advertisements as the estimates of σ_ν^2 and σ_ξ^2, respectively.

8) In contrast, generic entrants advertise sporadically if at all.

9) Producers as well as consumers may learn and adjust the expectations for the experience characteristics of their products in the first few years. Currie and Park (2000) showed that in the case of antidepressant drugs, the expected experience characteristic of a product, m_{jt}^δ, typically leveled off a few years after market introduction.

10) Price as well as advertising may signal product quality of experience goods. Milgrom and Roberts (1986) showed that advertising is used as a signal only if price dose not by itself achieve the necessary differentiation. However, it is an empirical question whether price serves as a signal for product quality in the case of prescription drugs.

11) Suppose that $A_j = \beta_0 + \beta_1 \delta_j$. Then instead of (5), we obtain:

$$m_{jt}^\delta = \frac{\theta_4 \overline{q}_{jt-1}}{((t_j-1)\sigma_\xi^2 + t_j\beta_1^2\sigma_\nu^2)/t_j\sigma_\xi^2 + \sigma_\nu^2/((t_j-1)\sigma_1^2)} + \frac{\beta_1(\overline{A}_{jt} - \beta_0)}{((t_j-1)\sigma_\xi^2 + t_j\beta_1^2\sigma_\nu^2)/t_j\sigma_\nu^2 + \sigma_\xi^2/(t_j\sigma_1^2)}.$$

signals in which advertising directly signals an experience characteristic. Specifically, we assume that the consumer receives an experience signal in period t as a linear function of the previous sales and the current advertising level: i.e., $\mu_{jt} = \beta_2(q_{jt-1} + \beta_3 A_{jt})$, where β_2 and β_3 are parameters. This specification allows the consumption level, q_{jt-1}, and the advertising level, A_{jt}, to substitute directly for each other in an experience signal. We also assume that in each period, the experience signal indicates an experience characteristic on average, and apply similar normal distribution assumptions in (2) and (3). Then the theory of conjugate distributions leads to the posterior distribution of δ_j given by a normal distribution with a mean of m^δ_{jt} and a variance of $\hat{\sigma}^2_{jt}$ as follows:

$$m^\delta_{jt} = \frac{\sigma^2_\xi m^\delta_{j0} + t_j \sigma^2_1 \beta_2 (\overline{q}_{jt-1} + \beta_3 \overline{A}_{jt})}{\sigma^2_\xi + t_j \sigma^2_1} \text{ and } \hat{\sigma}^2_{jt} = \frac{\sigma^2_\xi \sigma^2_1}{\sigma^2_\xi + t_j \sigma^2_1}. \tag{6}$$

Equation (6) indicates that the typical sales and advertising pattern of prescription drugs is also consistent with this alternative specification. In this direct signaling case (as well as the case of perfectly correlated indirect signals), however, the influence of advertising signal will not phase out. In Section 4, we will apply both signal specifications in (4) and (6) to calculate the price indexes for antidepressant drugs.

C. Market Share Functions

After the information pooling and the learning process discussed above, an individual consumer will choose one of the products or the outside alternative to maximize the expected (prior-to-consumption) utility in (1). Let for $\rho_{jt} = \theta_1 - \theta_2 p_{jt} + X_{jt}\theta_3 + m^\delta_{jt}$ for $j = 1, ..., J$, and call it the mean level of the utility of product j in period t. A composite good, say $j = 0$, other than the J products under consideration is called the outside alternative. Typically, the mean level of the utility of the outside alternative is normalized to be zero in the logit model. Then the nested logistic assumptions on $\varepsilon_{it}(g, m, j)$ in Cardell (1997) leads to a market share function for each product as follows:

$$S_{jt}(\theta_0) = \frac{e^{\rho_{jt}/(1-\sigma_g)(1-\sigma_m)} E^{-\sigma m}_{mt} D^{-\sigma_g}_{gt}}{1 + \Sigma_g D^{1-\sigma_g}_{gt}}, j = 1, 2, .., J, \tag{7}$$

where $\theta_0' = (\theta_1, \theta_2\theta_3', \theta_4, \sigma_m, \sigma_g, m^\delta_{j0}, m^A_{j0}, \sigma^2_\nu, \sigma^2_\xi, \Sigma_0)$, $E_{mt} = \Sigma^{}_{j \in m} e^{P_{jt}/(1-\sigma_g)(1-\sigma_m)}$, and $D_{gt} = \Sigma_{m \in g} E^{1-\sigma_m}$.

INNOVATION-AND-LEARNING-ADJUSTED PRICE INDEXES

In this section, we will construct ideal price indexes based on the structural model of Section 2. These structural-model-based price indexes will quantify consumers' welfare changes resulting from quality improvements or price changes of existing products, the introduction of new products, or consumers' learning about experience characteristics. Note that consumer i's (expected) gain from the optimal choice is: $\max_j E[U_{ijt} \mid I_{it}]$.

Using the nested logistic assumptions applied to the derivation of the market share function in (7), we can aggregate individual consumers' gains from the consumption of one of the products and obtain a (aggregate) consumer surplus function as follows:[12]

$$\gamma_t = e_t + W(p_t, X_t, m_t^{\delta}) \text{ with}, \quad W(p_t, X_t, m_t^{\delta}) = \ln\left[\Sigma_g D_{gt}^{1-\sigma_g}\right]/\theta_2, \tag{8}$$

where e_t is the (average) income spent on the products under consideration, and (p_t, X_t, m_t^{δ}) denotes the vector of $\{(p_{jt}, X_{jt}, m_{jt}^{\delta})\}_{j=1,\ldots,J}$.[13] Note that e_t and p_t are *real* in the sense that they are deflated by the price index for the outside alternative. Usually, the Consumer Price Index (CPI) is used as the price index for the outside alternative (see Trajtenberg (1990)). Consequently, the price indexes developed in this section indicate the change of innovation-and-learning-adjusted real prices and are called *real* price indexes. For the comparison with other available indexes, the real price indexes will be multiplied by the CPI and converted into (nominal) price indexes.

Inverting the consumer surplus function in (8), we can derive an (aggregate) expenditure function as follows:

$$e(\gamma_t, \ p_t, \ X_t, \ m_t^{\delta}) = \gamma_t - W(p_t, \ X_t, m_t^{\delta}). \tag{9}$$

The expenditure function in (9) assigns the minimum expenditure required to achieve the consumer surplus, γ_t, for given prices, product characteristics and experience characteristics, (p_t, X_t, m_t^{δ}). Then, depending on whether we employ γ_t or γ_{t-1} as the reference level, we can calculate ideal price indexes between period t and period $t-1$ as follows:

$$I^0 = \frac{e(\gamma_{t-1}, p_t, X_t, m_t^{\delta})}{e(\gamma_{t-1}, p_{t-1}, X_{t-1}, m_{t-1}^{\delta})} = \frac{\gamma_{t-1} - W(p_t, X_t, m_t^{\delta})}{\gamma_{t-1} - W(p_{t-1}, X_{t-1}, m_{t-1}^{\delta})}, \tag{10}$$

and

$$I^1 = \frac{e(\gamma_t, p_t, X_t, m_t^{\delta})}{e(\gamma_t, p_{t-1}, X_{t-1}, m_{t-1}^{\delta})} = \frac{\gamma_t - W(p_t, X_t, m_t^{\delta})}{\gamma_t W(p_{t-1}, X_{t-1}, m_{t-1}^{\delta})}. \tag{11}$$

Note that I^0 will have a negative sign if there is a drastic innovation and learning such that $(W(p_{t-1}, X_{t-1}, m_{t-1}^{\delta}) <) \gamma_{t-1} < W(p_t, X_t, m_t^{\delta})$.

We can also construct the price indexes suggested in Trajtenberg (1990), which do not depend on the reference level of consumer surplus. Since the expenditure function is

---

12) In this aggregation procedure, the market size is normalized to be 1, and thus the consumer surplus obtained in (8) can be interpreted as the per capita surplus. Hence, the price indexes based on this consumer surplus function are not affected by the change of the market size (i.e., the number of potential buyers) over time.

13) Since the mean level of the utility of the outside alternative is normalized to be 0, the consumer surplus from the outside alternative will be: $y_t - e_t + 1$.

additively separable in γ_t, the compensating and equivalent variations are the same. Hence we have:

$$e(\gamma, p_{t-1}, X_{t-1}, m^{\delta}_{t-1}) - e(\gamma, p_t, X_t, m^{\delta}_t) = W(p_t, X_t, m^{\delta}_t) - W(p_{t-1}, X_{t-1}, m^{\delta}_{t-1}) \equiv D\_W_t.$$

Let ϕ_t be the hypothetical average price reduction that would have had the same welfare consequences as the innovations and consumers' learning that actually took place. Then

$$D\_W_t = W((1-\phi_t)p_{t-1}, X_{t-1}, m^{\delta}_{t-1}) - W(p_{t-1}, X_{t-1}, m^{\delta}_{t-1}). \tag{12}$$

Hence a price index can be computed simply as $I^0 = 1 - \phi_t$. This price index, like I^0, will have a negative sign when innovations and learning effects are drastic. The other price index, which, like I^1, still has a positive sign in the occurrence of drastic innovations and learning, can be obtained by solving for φ_t from

$$D\_W_t = W(p_t, X_t, m^{\delta}_t) - W((1+\varphi_t)p_t, X_t, m^{\delta}_t). \tag{13}$$

That is, if the prices of the improved products had been $(1+\varphi_t)$ times higher than actual prices, then the implied percentage price reduction of $\varphi_t/(1+\varphi_t)$ would be equivalent to the value of the innovations and consumers' learning that took place. Hence the last price index is computed simply as $I^1 = 1/(1+\varphi_t)$. Note that the price of each product at time t can always be written as: $p_{jt} = p_t^* + \Delta_t p_{jt}^*$, where p_t^* and Δ_t are the average price and the standard deviation of prices across the products under consideration, respectively. Trajtenberg (1990) showed that ϕ_t in (12) and φ_t in (13) can be easily calculated if the distribution of prices moves by a factor of $(1-\phi_t)$ while the standard deviation remains unchanged over time (i.e., $\Delta_t = \Delta$). Then $p_{jt} = (1-\phi_t)p_{t-1}^* + \Delta p_{jt-1}^*$, and thus $\phi_t = D\_W_t / p_{t-1}^*$. Similarly, $1+\varphi_t = (D\_W_t + p_t^*) / p_t^*$. This assumption, however, will not be valid in the case of prescription drugs. A stylized fact in the pharmaceutical industry is that the prices of branded drugs rise over time while generic prices fall (See Currie and Park (2000)). Therefore, in the following empirical example, we will focus on the previous two innovation-and-learning-adjusted price indexes, I^0 and I^1.

AN EMPIRICAL APPLICATION TO ANTIDEPRESSANT DRUGS

A. *The Market for Antidepressant Drugs*

We now discuss an empirical application of our price indexes to the data for antidepressant drugs during the years 1980-1995. During these years, the market for antidepressant drugs was one of the fastest growing industries: the Average Annual Growth Rates (AAGRs) of daily dosage units sold and revenues were 11.44% and 24.19%, respectively. For more detailed discussions of this market, refer to Berndt, Cockburn and Griliches (1996), and Currie and Park (2000). To treat chronic depression, the first antidepressant drug was introduced in 1958.[14] As of 1995,

---

14) Depression symptoms that last for 24 or more months are referred to as chronic depression.

antidepressants are categorized into four therapeutic subclasses: tricyclic antidepressants (TCAs), monoamine oxidase inhibitors (MAOIs), selective serotonin reuptake inhibitors (SSRIs), and the other antidepressant drugs ("Others"). The first of MAOIs was introduced to the market in 1959, but since then MAOIs have maintained only a small market share (less than 1.5% for the years 1980-1995). In the 1960s, a considerable number of (the first generational) TCAs entered the market. In fact, until the introduction of the SSRI subclass in 1988, TCAs dominated the market (See Figure 1).

Antidepressant drugs are mainly differentiated by side effects[15] such as anticholinergic (AC),[16] drowsiness (DR), insomnia/agitation (IA), orthostatic hypotension (OH), cardiac arrythmia (CA), gasrointestinal distress (GID), weight gain (WTG), and fatal in overdose (Fatal). Other product characteristics include the daily frequency of taking the pill (Freq) and half-life (Half).[17]

The data used in this paper are the same annual data in Currie and Park (2000), which contain the prices, sales, advertising and observable product characteristics of all the antidepressant drugs for the years 1980-1995: i.e., 21 antidepressant molecules and 35 products (branded or generic versions of a molecule).[18]
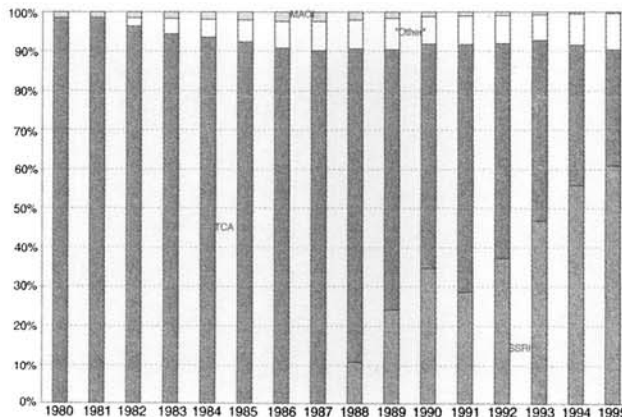

Figure 1. Market Shares.

During the years 1980-1995, the market for antidepressants experienced 'exceptional and remarkable' innovations in terms of entry of new products (both new molecules and generics). Table 1 lists the entry of new molecules and generics for the entire years in question. There have been few exits during these years, and all exits were secondary brands in the TCA subclass. In 1980 and 1981, there was the branded entry of the second-generational TCAs along with the entry of a branded drug in the "Other" subclass. These new second-generational TCAs rapidly replaced the first generation of

---

15) There is no strong clinical evidence that various antidepressants have different rates of efficacy.

16) 'AC' includes dry mouth, blurred vision, urinary hesitancy and constipation.

17) 'Half life' is the number of hours for the drug to leave a patient's bloodstream.

18) The data of prices, sales and advertising come from IMS America. Based on actual invoices, IMS tracks more than 99% of manufacturers and wholesale transactions to drugstores or their purchasing agents.

TCAs, most of which had been introduced in 1960s and had dominated the antidepressant market. Beginning in 1986, generic entry in the TCA subclass became significant. This active entry of generic drugs might have been due to the reduced costs of generic entry facilitated by the passage of the 1984 Waxman-Hatch Act. Most importantly, Prozac was introduced to the market in 1988, and subsequently four more drugs in the same therapeutic subclass followed suit. Prozac, the first of the SSRIs, was a breakthrough: it, and others in the SSRI subclass, has biologically more specific actions and thus fewer side effects. Since the introduction of its first molecule, the market share of the SSRI subclass rose steadily, overtook the TCA subclass in 1993 and occupied 63% of the market by 1995 (See Figure 1).

Table 1. Entry in the Antidepressant Drug Market, 1980 - 1995

year	molecules
1980	amoxapine (branded version, TCA)
1981	maprotiline (branded version; TCA); trazodone (branded version; Others)
1982	
1983	
1984	
1985	
1986	doxepin (generic version; TCA); trazodone (generic version, Others)
1987	desipramine (generic version; TCA)
1988	fluoxetine (branded version, SSRI); maprotiline (generic version, TCA); trimipramine (generic version, TCA)
1989	bupropion (branded version; Others); amoxapine (generic version, TCA)
1990	clomipramine (branded version; TCA)
1991	
1992	sertraline (branded version, SSRI); nortriptyline (generic version, TCA); nortriptyline (generic version, TCA)
1993	paroxetine (branded version, SSRI)
1994	fluvoxamine (branded version, SSRI); venlafaxine (branded version, Others)
1995	nefazodone (branded version, SSRI)

The growth of the market for antidepressant drugs has accelerated since the introduction of the SSRI subclass in 1988. The AAGR of daily dosage units was about 5.3% from 1980 to 1987 but more than tripled to 18.3% from 1987 to 1995. Figure 2 illustrates the dramatic increases in the usage of antidepressant drugs to treat chronic depression from 1988, although the market size (i.e., the number of people who suffer from chronic depression) has grown steadily since 1980.[19] The difference between the

---

19) The yearly market size is computed from the estimate of the prevalence of depression in the U.S. population multiplied by the population, assuming six-month treatment duration. (See Depression Guideline Panel, 1993).

market size and the usage of antidepressants indicates the share of the outside alternative, which includes no treatment at all, or only non-drug treatments such as psychotherapy. Hence, Figure 2 implies that more and more people who suffered from chronic depression began to use antidepressant drugs since the introduction of the SSRI subclass.[20]
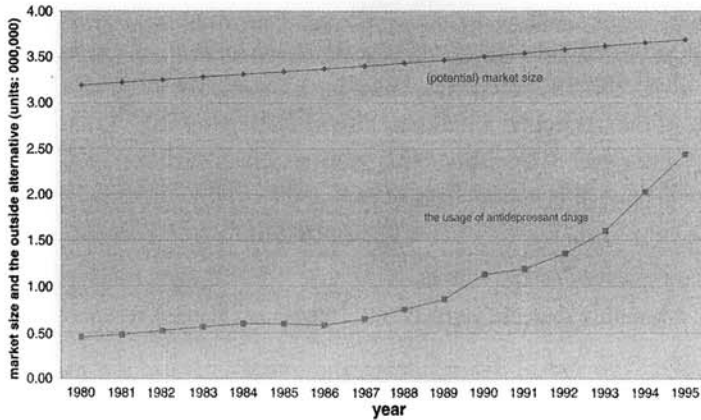

Figure 2. Market Size and Drug Usage.

B. *Estimation Results*

In order to calculate the innovation-and-learning-adjusted price indexes in (10)-(11), we first have to estimate the parameters of the structural model in (1) and (4) (or in (1) and (6)). We will mainly use the estimation results in Currie and Park (2000) to calculate these price indexes in the next subsection. In this subsection, we will briefly review the estimation procedure and results in Currie and Park (2000).

To estimate the parameters, θ_0, of the structural model in (1) and (4), Currie and Park (2000) explicitly considered measurement errors in the data of reported market shares (or sales) and applied a non-linear least squares estimation procedure. Specifically, assume that $S_{jt} = S_{jt}(\theta_0)e^{u_{jt}}$, where S_{jt} is the reported market share of product j in period t, and $e^{u_{jt}}$ is a measurement error with $E[u_{jt}|p_t, X_t, A_{t0},..., A_t, q_{t0},...,q_{t-1}]= 0$. Then,

$$\ln S_{jt} = \ln S_{jt}(\theta_0) + u_{jt}. \tag{14}$$

Based on (14), we can apply a non-linear least squares estimation procedure to obtain the estimate of θ_0. For details, refer to Currie and Park (2000). Note that an empirical study with logistic demand functions for differentiated products and market-level data usually encounters an unobservable (to economists) product characteristic, which can be treated as an error term in an estimation procedure (See Berry (1994)). In the case of

---

20) In practice, the combination of psychotherapy and antidepressant drugs, in varying proportions, constitute a common treatment strategy for chronic depression symptoms (See Berndt, Bir, Busch, Frank, and Normand (2000)). However, we are unaware of any evidence that advances in psychotherapy accelerated the use of antidepressant drugs especially since 1988.

prescription drugs, the experience characteristic can be treated as this unobservable characteristic since all the other product characteristics such as side effects and efficacy rates are usually available information. However, if the experience characteristic, which is a function of all the histories of sales and advertising, is treated as an error term in the estimation procedure, we may not have any valid instrumental variables. Hence, it is desirable to estimate the (expected) experience characteristic, instead of treating it as an error term.

Based on the distance metric statistic in Newey and McFadden (1994), Currie and Park (2000) also conducted several hypothesis tests. Currie and Park (2000) found that there was a positive correlation between the mean level of advertising and the experience characteristic in the market for antidepressants during the years 1980 -1995. Hence, as discussed in Section 2, this positive correlation implies that advertising provides indirect information on the experience characteristic. In addition, Currie and Park (2000) found: (i) there was no image effect of advertising; (ii) there were no significant spillover effects of advertising and consumption experiences within the molecule; (iii) the SSRI subclass had a different value of the parameter of consumption experience, θ_4; (iv) the experience characteristics of the first generation of TCAs and MIAOs were equal to zero; and (v) the initial priors of δ_j and A_j had zero means (i.e., $m_{j0}^\delta = m_{j0}^A = 0$). In addition, in this paper, we test whether consumers are risk-averse about experience characteristics. For this statistical test, we add the standard deviation of expected experience characteristic in (4) to the mean level of the utility of a product, that is, $\rho_{jt} = \theta_1 - \theta_2 p_{jt} + X_{jt}\theta_3 + m_{jt}^\delta + \theta_5 \sigma_{jt}$, where σ_{jt} is the squared root of the (1,1)-th element of Σ_{jt} in (4). Then with the null hypothesis that $\theta_5 = 0$ and the alternative hypothesis that $\theta_5 < 0$, we apply the distance metric statistic, which turns out to have the value of almost zero with the chi-squared distribution of the degree of freedom one. Hence we statically reject the argument that consumers are risk-averse about experience characteristics in the case of antidepressant drugs.

Table 2 reports the estimates of the parameters of the structural model in (1) and (4) applied to the antidepressant drugs for the years 1980-1995. The estimated within-molecule coefficient, σ_m, is quite low, which implies that the consumer has a strong idiosyncratic taste for a branded or a generic version of the same molecule even after we take account of differences in experience characteristics. The estimated within-subclass coefficients, σ_g, are not very high, (that of the TCA subclass is especially low), which may reflect the noticeable differences between the first and the second generations of TCAs. All the estimated within-group coefficients are reasonably significant. The observable product characteristic profiles did not change over time in our data set. The estimated coefficients of product characteristics such as IA, DR, and OH are reasonably significant. However, we have positive signs of the estimated coefficients of AC, Fatal, GID and CA, which we would expect negatively valued by the consumer, although they are imprecisely estimated. The fact that the coefficient of Fatal has a positive sign is particularly disturbing since the most significant novel characteristic of SSRIs is that they are not fatal in overdoses. However, the index of severity of side effect profiles does not indicate any significant differences in this characteristic between the second-generation TCAs and SSRIs. Recall that the difference, which is not reported in the official measurement, is reflected in consumers' experience characteristics. Currie and

Table 2. Results of Regression

variable	estimate	st. error	t ratio
constant	0.235	1.726	0.14
price	-0.324	0.294	-1.10
generic	0.427	0.187	2.28
Within-group Correlation			
sigma_tca	0.063	0.037	1.71
sigma_maoi	0.535	0.271	1.98
sigma_ssri	0.495	0.265	1.87
sigma_other	0.356	0.291	1.22
sigma_molecule	0.262	0.116	2.25
Product Characteristics			
Insomnia/Agitation	-0.826	0.303	-2.72
Drowsiness	-0.426	0.331	-1.29
Anticholinergic	0.466	0.526	0.89
Frequency	-0.263	0.298	-0.88
Fatal	0.305	0.483	0.63
Gastrointenstinal	0.110	0.136	0.81
Weight Gain	-0.157	0.345	-0.45
Cardiac	0.186	0.401	0.46
Orthostatic Hypotension	-0.186	0.153	-1.22
Half-life	0.004	0.006	0.68
Consumption Experience			
SSRI	0.257	1.114	0.23
all the others	0.673	0.575	1.17
Prior Covariance Martix			
Var[delta]	0.576	1.414	0.41
Var[A_detail]	0.270	1.550	0.17
Var[A_journal]	1.229	7.316	0.17
Cov[delta, A_detail]	0.117	0.343	0.34
Cov[delta, A_journal]	0.329	0.719	0.46
Cov[A_detail, A_journal]	0.604	0.011	53.64

Park (2000) confirmed that the average expected experience characteristic of SSRIs was much higher than that of the second-generational TCAs although the coefficient of the consumption experience, θ_4, of SSRIs is estimated to be smaller than that of the others. The coefficient of the dummy variable for generic drugs is significant and positive, which may reflect the existence of generic substitution law. The initial variances and covariances of experience characteristic, δ_j, the mean level of detailing, say A^D_j, and the mean level of journal advertising, say A^J_j, are not precisely estimated except the covariance between A^D_j and A^J_j. In the data set, detailing and journal advertising are measured by the total minutes of detailing and the total costs of journal advertising, respectively. The correlation coefficient between δ_j and A^D_j is 0.3 and the correlation coefficient between δ_j and A^J_j is 0.39.

Overall, the estimates reported in Table 2 are imprecise, but these are more precise estimates than those from other alternative estimations. First, as an alternative, we also

estimate the utility function in (1), using the specification of experience signals in (6).[21] The estimates under the alternative specification in (6), however, turn out to be less precise overall. Second, we apply the same data to a typical nested-logit-based estimating equation as in Stern (1996), in which cumulative advertising is added directly to the consumer's mean level of utility and an unobservable characteristic such as an experience characteristic is treated as an error term. This typical estimation, however, leads to not only less precise estimates but also the incorrect sign of the coefficient of 'price'. Hence all these estimations of the structural model for prescription drugs, based on the product-level aggregate data, end up with imprecise estimates. A reason for this may be the sample size of product-level aggregate data. Because of patent protection and FDA regulation, there are few different prescription drugs in the market for a certain therapeutic class. In our case, there are only 404 observations for 35 different antide-pressant products over 16 years. A possible remedy for this imprecision issue, therefore, is exploiting consumer-level data. Indeed, when we presume that the data are generated from the observations of each individual consumer's choices and estimate the same structural model based on the maximum likelihood estimation procedure, the estimates become very (actually, too) precise. With reasonable consumer-level variations incorpo-rated to the model, we may expect to obtain sufficiently precise estimates of the struc-tural model.

C. Price Indexes

We now discuss our innovation-and-learning-adjusted price indexes for antidepressant drugs based on the estimates reported in Table 2. Figure 3 illustrates the weighted average (nominal and real) prices of antidepressants during the years 1980-1995. Since 1988 in which Prozac was introduced to the market, the weighted average prices, both real and nominal, began to increase more rapidly, with the exception of 1991.[22] The AAGR of the weighted average nominal prices was 9.45% from 1980 to 1987 and 11.21% from 1987 to 1995, while the AAGR of the weighted average real prices was 4.02% from 1980 to 1987 and 7.71% from 1987 to 1995. Except in 1981 and 1991, the weighted average real prices have continued to rise. As discussed above, however, during the same time, the market for antidepressant drugs experienced a series of exceptional and remarkable innovations resulting from the entry of new products (both new molecules and generics). Hence, if the value of these innovations and consumers' learning about experience characteristics dominated the effects of price increases, the consumer surplus might rise during the time. Recall that the observable product characteristics did not change over time in the case of antidepressant drugs.

Based on the estimates reported in Table 2, we first calculate changes of consumer surpluses from the consumption of antidepressant drugs for the years in question. Figure 4 illustrates the changes in consumer surplus, $D\_W = W(p_t, X_t, m^\delta_t) - W(p_{t-1}, X_{t-1}, m^\delta_{t-1})$, and the changes in consumer surplus when the expected experience characteristics are set to

---

21) We are not aware of any test statistic for the non-nested test between these two specifications of experience signals in (4) and (6).

22) Real prices mean the prices deflated by the CPI to convert into 1980 dollars.
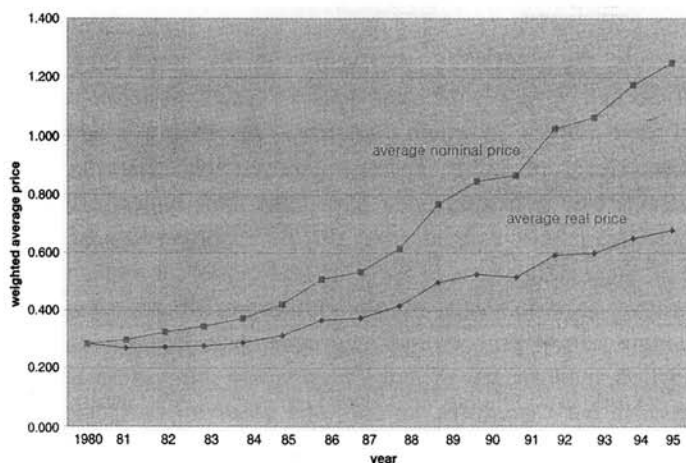
Figure 3. Average Price.

be zero, say $D\_W^* = W(p_t, X_t, 0)-W(p_{t-1}, X_{t-1}, 0)$. The difference between D_W and D_W^* represents the (aggregate) contribution of experience characteristics to the increase of consumer surplus.[23] Over the entire years in question, the aggregate learning effects accounted for 11.17% of the increase in consumer surplus, D_W. Figure 4 indicates that big welfare gains from consumers' learning were followed by the active generic entry (facilitated by the 1984 Waxman-Hatch Act) in 1986, and the entry of Prozac in 1988. For 1989-1991, there was a setback following these huge learning effects. However, the aggregate learning effects accounted for 17.3% of the average increase in consumer surplus for 1987-1993. The pattern of learning effects over time illustrated in Figure 4 implies that consumers' expectations for experience characteristics quickly level off.
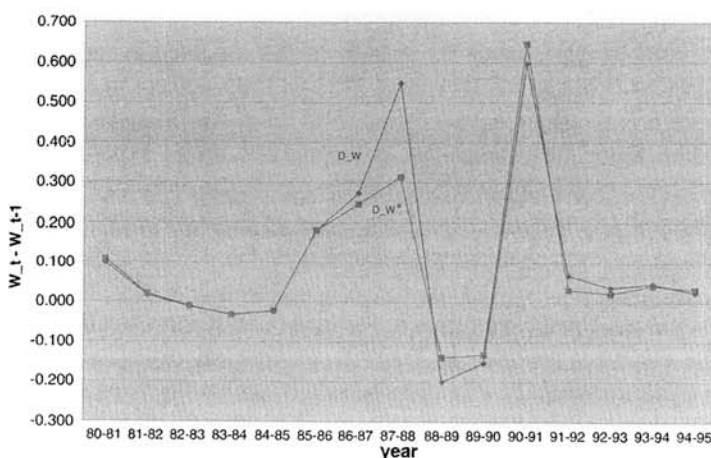

Figure 4. Changes in Consumer Surplus.

---

23) Here we do not explicitly consider the possible changes in optimal prices. With positive consumer learning effects, producers may set prices higher (due to higher mean levels of utility for the products) or lower (in order to build up more consumption experiences).

Figure 4 shows two big jumps of D_W in 1987-1988 and 1990-1991. The big jump in 1987-1988 indicates the impact of the introduction of Prozac, which proved that it was a breakthrough in the antidepressant drug industry. The jump in 1990-1991 was due primarily to the decrease in weighted average real price. Except in 1991, as illustrated in Figure 3, the weighted average prices increased rapidly from 1988. This is mainly because the SSRIs were priced substantially higher than others were,[24] but their market share continued to rise (except in 1991). Therefore, the sharp drop of the weighted average real price in 1991 reflects aggressive price cuts of the other products (especially the second-generation TCAs) in response to the surge of Prozac, which resulted in the drop in the market share of Prozac in 1991 (See Figure 1). In that sense, the jump in 1990-1991 reflects another face of the drastic innovation caused by the introduction of the breakthrough drug. In addition, the positive values of D_W in 1991-1995 reflect the innovations resulting from subsequent entry of new molecules in the SSRI subclass as shown in Table 1.

The high values of D_W for 1980-1982 in Figure 4 reflect the entry of the second-generational branded drugs in the TCA subclass (and the "Other" subclass). The decline of weighted average real prices in 1981 may imply strategic price cuts of the existing antidepressant drugs in response to the introduction of the second-generational TCAs in 1980 and 1981. The big surge of D_W in 1985-1987 illustrates the value of innovations that resulted from a flurry of generic entries in the TCA subclass facilitated by the 1984 Waxman-Hatch Act. For 1982-1985 and for 1988-1990, D_W had negative values. In other words, the effects of price increases dominated the positive values of innovations and consumers' learning. During these years, there was no significant entry of new products.
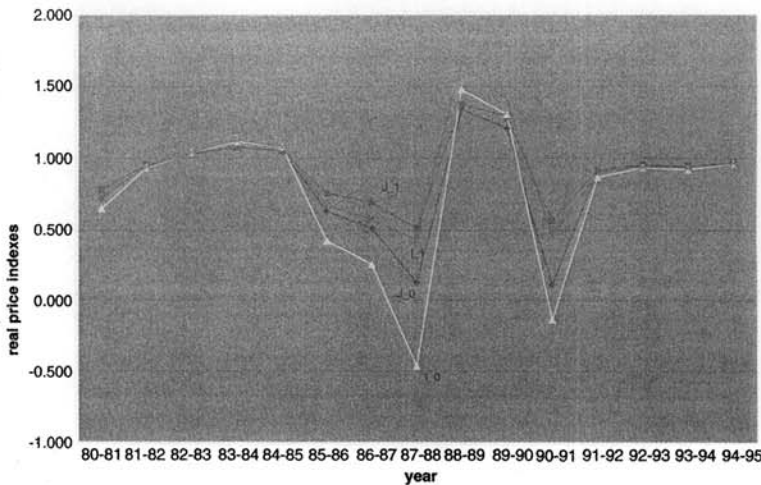

Figure 5. Real Price Indexes.

Figure 5 illustrates the four real price indexes defined in Section 3. Figure 6 calculates the (nominal) price indexes I^l and J^l in the sense that increases in the price levels of the

---

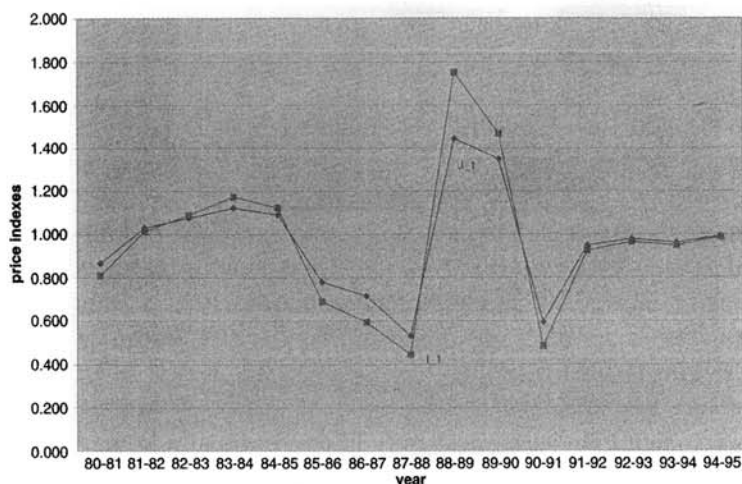24) In 1988, Prozac was about five times as expensive as the average TCA branded drugs.

Figure 6. Price Indexes.

outside alternative are taken into account. Figures 5 and 6 clearly indicate that these calculated price indexes truthfully reflect the changes in consumer surplus reported in Figure 4. Recall that in the case of prescription drugs, we do not have to worry about the overstatement of consumer welfare gains from new products in the typical logit model. Figure 4 shows that our focal index I^{l} has a more fluctuation over time than J^{l}. The four real price indexes show decreases in the innovation-and-learning-adjusted price for the entire years in question, except in the years 1982-1985 and 1988-1990 when the increases of social surplus are negative. In addition to these years, the nominal indexes have increased in 1981-1982. Overall, however, the values of innovations resulting from a series of entries of new products and learning about experience characteristics have been dominant. The price index I^{0} has negative signs, especially in 1987-1988 and 1990-1991, which implies the occurrence of dramatic innovations because of the entry of Prozac in 1988 and the second-generational TCAs' aggressive price cut in response to the surge of Prozac in 1991.

We now proceed to compare our price indexes reported in Figure 4 with those calculated in previous studies. In June 1980, the PPI program at the BLS began publishing a monthly price index for psychotherapeutics consisting of tranquilizers and antidepressants ("Cycle A" sample). In December of 1987, the BLS created a separate category for antidepressants ("Cycle B" sample) and updated its sample in December of 1993 ("Cycle C" sample). The PPI published by the BLS is calculated by a modified Laspeyres formula. However, the BLS PPI for antidepressants may have had an upward bias since the BLS had implicitly treated generic versions of a drug as entirely distinct and non-substitutable products.[25] The FDA, on the other hand, has argued the other extreme: "a pill, is a pill, is a pill." In this case, the relevant price for a molecule will be the weighted average price of all generic drugs and the branded drug within the

---

25) In addition, the BLS PPI had a problem caused by the employed weights. For example, Cycle C sample excludes Prozac, the largest selling antidepressant because Prozac is manufactured in Puerto Rico. Refer to Berndt, Cockburn and Griliches (1996) for a detailed discussion about the BLS PPI.

molecule. Beginning in May 1996, the BLS adopted a new procedure, which treats generics and their branded antecedents as perfect substitutes and calculates the Laspeyres index with fixed branded weight split into a 64.2% generic component and a 35.8% branded component. Emphasizing the importance of new products, Berndt, Cockburn and Griliches (1996) calculated a Paasche price index for antidepressants based on the Griliches-Cockburn adjusted Paasche Diffusion (GCPD) method developed in Griliches and Cockburn (1994). Based on estimates of hedonic price indexes, they predicted the reservation price of a new branded drug prior to market introduction, while they calculated the reservation price of a generic drug prior to the market introduction as the average six-month-later price of the generic drug and its branded antecedent.

Table 3. Average Annual Growth Rates of Price Indexes

Procedure	Entire period 1980-1995	Period I 1980-1987	Period II 1987-1993	Period III 1993-1995
weighted average price	10.38	9.45	12.16	8.41
I_1*	-9.52	-9.9	-11.08	-3.29
J_1*	-6.65	-5.85	-8.95	-2.39
FDA average price**	2.95	5.71	1.33	1.1
New BLS**	3.71	7.41	2.49	0.42
GCPD**	4.33	7.08	3.99	0.52
BLS index**	NA	NA	10.4	4.27

* For the definitions, refer to the text.
** source: Berndt, Cockburn and Griliches (1996). For these four indexes, the entire period covers from 1980:1 to 1996:2, period I from 1981:12-1987:12, period II from 1987:12 to 1993:12, period III from 1993:12 to 1996:2.

Table 3 shows the AAGRs for several different alternative price index calculations. Because of drastic innovations in 1988 and 1991 and the reasons discussed in section 3, we use I^I as our focal index. For the entire years in question, the AAGR of I^I is -9.52% (the AAGR of J^I is -6.65%).[26] In other words, there has been roughly a 9.5% annual decline of innovation-and-learning-adjusted average prices of antidepressant drugs. The AAGR of I^I, based on the estimates under the alternative experience signal in (6), is -13.27%. During the same time, the AAGR of weighted average nominal prices is 10.38%. These numbers indicate that despite substantial increases in prices, there has been an 'exceptional and remarkable' series of innovations in the market for antidepressant drugs during the years in question. Recall that a great number of people with chronic depression, who might have had no treatment at all or chosen expensive non-drug treatment such as psychotherapy, turned to antidepressant drugs since the

---

26) We calculate the AAGR from year 0 to year n, say g, by solving: $(1+g)^n = I_{01} I_{12} I_{n-1n}$, where I_{ii+1} is the price index from year i to year $i+1$.

introduction of the SSRI subclass in 1988 (See Figure 2). The AAGR of I^l for this market, -9.52 percent, is not a surprising figure, compared to those of price indexes for PC's and CT scanners. The AAGR of hedonic price indexes was -30% in the U.S. PC market over the 1988-1992 period (see Berndt, Griliches and Rappaport (1995)), and the AAGR of the price indexes J^l for CT scanners reported in Trajtenberg (1990) was -55.87% over the 1974-1982 period. Based on medical claims data and expert clinical opinion, it is also noteworthy that Berndt, Bir, Busch, Frank and Normand (2000) found that the expenditure index per incremental full or partial remission of treating chronic depression (including both psychotherapy and pharmacotherapy) has declined annually at about 4.8 to 2.1% over the years 1991-1996. When they considered the effects of changing patient mix in a hedonic-like equation, the annual reductions ranged from about -1.66% to -2.13% per year. In our calculation, the AAGR of I^l for antidepressant drugs over 1991-1995 is -4.43%.

As shown in Table 3, however, all the other indexes discussed in Berndt, Cockburn and Griliches (1996) reported positive values of AAGRs for antidepressants during the years 1980-1995: 2.95% of the FDA average price, 3.71% of the New BLS procedure, and 4.33% of the GCPD.[27] Moreover, these three indexes, along with the BLS index, have higher values of AAGRs in period II (1987 to 1993) than in period III (1993 to 1995). As discussed above, the most substantial and remarkable innovations occurred in period II. As shown in Figure 4, there are two big increases of consumer surplus in 1987-1988 and 1991-1992, which were the result of the introduction of Prozac and the dramatic price cuts of the other products in response to the surge of Prozac. Our innovation-and-learning-adjusted indexes I^l and J^l show faster declines of the growth rates in period II than in period III (although there were sharp increases of I^l and J^l in 1988-1990). The other three price indexes as well as the BLS index fail to capture these substantial and remarkable innovations caused by the introduction of the breakthrough subclass, SSRI. The prices of SSRIs were substantially higher than the other antidepressants, and steadily increased over this time period. These previous indexes overstated the increases in prices, and did not truthfully reflect the exceptional and remarkable innovations in period II. Note also that our innovation-and-learning-adjusted price indexes have substantially lower values in period I, compared to the previous three indexes. In period I, the price increases were slower than in periods II and III, but consumers substantially benefited from the introduction of the second-generational TCAs and a flurry of generic entries facilitated by the 1984 Waxman-Hatch Act.

CONCLUDING REMARKS

This paper has proposed innovation-and-learning-adjusted price indexes for prescription drugs, based on the characteristics approach to demand functions and the Bayesian learning process of experience characteristics. This structural approach not only resolves quality adjustment of new molecules, but also avoids arbitrary assumptions on the link-in of generic drugs to the originator branded drug. In this paper, the link-in problem

---

27) Note that these three indexes are calculated using monthly data.

was treated in a context of informational product differentiation between generic drugs and the originator branded drug. The suggested price indexes have been applied to the data for antidepressant drugs during the years 1980-1995. Our calculated innovation-and-learning-adjusted indexes truthfully reflect exceptional and remarkable innovations in the market for antidepressants during these years: the AAGR of the focal price index is -9.5%. We have also found that the effects of consumers' learning about experience characteristics were also significant in consumer welfare gains. These results suggest that all the other existing indexes discussed in Berndt, Cockburn and Griliches (1996) seriously understate the value of innovations and consumers' learning and thus substantially overstate the rate of inflation in a rapidly growing market with the entry of innovative new prescription drugs.

A desirable extension of this paper is closely related to the availability of consumer-level information. A downside of our structural approach is that our price indexes require estimating a complicated structural model. As discussed in Section 4, estimations of the structural model for prescription drugs, based on product-level aggregate data, usually end up with imprecise estimates. A reason for this may be the sample size of product-level aggregate data. Because of the patent protection and the regulation of the FDA, there are not many different prescription drugs in the market for a certain therapeutic class. With a sufficient number of consumer-level observations, however, we may expect to obtain reasonably precise estimates of the structural model. The consumer-level data will also be valuable for incorporating each individual patient's income and insurance status into the consumer's choice problem. In the model presented in Section 2, these individual patients' variations are assumed a part of the consumer's idiosyncratic tastes for products. The information of each patient's insurance status will be of particular use for obtaining a more accurate estimate of the coefficient of 'price'. As discussed by Keeler (1996), fee-for-service health insurance may exaggerate a consumer's apparent marginal willingness to pay for newer or more expensive drugs than managed care such as Health Maintenance Organizations. However, as discussed in Keeler (1996), there is little evidence whether this moral hazard problem is significant in the purchase of prescription drugs. The consumer-level information, coupled with our structural model, may be able to provide statistical evidence on this moral hazard problem.

While the topic of this paper is on prescription drugs, the problem discussed in the paper is of wider importance. The suggested innovation-and-learning-adjusted price indexes will enable us to quantify the effects of consumers' learning about experience goods in general. Furthermore, our innovation-and-learning-adjusted indexes may be adapted to the case of informational diffusion and introductory pricing.

Bibliography

Ackerberg, D. 1997. "*Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination.*" Mimeo, Boston University.

_____. 2001. "Empirically Distinguishing Informative and Prestige Effects of Advertising." *Rand Journal of Economics*, Vol. 32, pp. 443-481.

American Psychiatric Association. 1994. *Diagnostic and Statistical Manuel for Mental Disorders*. Four Edition, Washington DC.

Berndt, E. R. Ed. July 21, 1999. *Research on Price Index Measurement: Agendas for the Next Twenty Years*. Panel presentations at the NBER Productivity Program of the Summer Institute.

_____. 2000. Bir A., Busch S. H., Frank R. G., and Normand S.-L. T. "The Medical Treatment of Depression. 1991-1996: Productive Inefficiency, Expected Outcome Variations, and Price Indexes." NBER Working Paper No. 7816.

_____, Pindyck R. S., and Azoulay P. 2000. "*Consumption Externalities and Diffusion in Pharmaceutical Markets: Antiulcer Drugs*." NBER Working Paper No. 7772.

_____, Cockburn I., and Griliches Z. 1996. "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs." *Brookings Papers on Economic Activity: Microeconomics*, pp. 133-188.

_____, Griliches Z., and Rappaport N. J. 1995. "Econometric Estimates of Price Indexes for Personal Computers in the 1990's." *Journal of Econometrics*, Vol. 68, pp. 243-68.

Berry, S. 1994. "Estimating Discrete Choice Models of Product Differentiation." *RAND Journal of Economics*, Vol. 25, pp. 242-262.

_____ and Pakes, A. 1999. "Estimating the Pure Hedonic Discrete Choice Model." Mimeo, Yale University.

_____, Levinsohn, J., and Pakes, 1995. A. "Automobile Prices in Market Equilibrium." *Econometrica*, Vol. 63, pp. 841-890.

Cardell, N. S. 1997. "Variance Components Structures for the Extreme Value and Logistic Distributions with Applications to Models of Heterogeneity." *Econometric Theory*, Vol. 13, pp. 185-213.

Ching, A. 2000. "*Dynamic Equilibrium in the U.S. Prescription Drug Market after Patent Expiration*." Mimeo, Ohio State University.

Crawford, G. S. and Shum, M. 2000. "*Uncertainty and Learning in Pharmaceutical Demand*." Mimeo, Duke University.

Currie, G. R. and Park, S. 2000. "*The Effects of Marketing and Consumption Experience on Demand for Antidepressant Drugs*." Mimeo, SUNY at Stony Brook.

Depression Guideline Panel. 1993. *Depression in Primary Care*. Vol. 2, Rockville, MD: U.S. Department of Health and Human Services, AHCPR Publication No. 93-0551.

Erdem, T. and Keane, M. 1996. "Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." *Marketing Science*, Vol. 15, pp. 1-20.

Feenstra, R. C. 1997. "Generics and New Goods in Pharmaceutical Price Indexes: Comment." *American Economic Review*, Vol. 87, pp. 760-767.

Fisher, F. M. and Griliches, Z. 1995. "Aggregate Price Indexes, New Goods, and Generics." *Quarterly Journal of Economics*, Vol. 110, pp. 229-244.

Grabowski, H. G. and Vernon, J. H. 1992. "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act." *Journal of Law and Economics*, Vol. 35, pp. 331-350.

Griliches, Z. Ed. 1971. *Price Indexes and Quality Change*. Cambridge, MA: Harvard University Press.

_____ and Cockburn, I. 1994. "Generics and New Goods in Pharmaceutical Price Indexes." *American Economic Review*, Vol. 84, pp. 1213-1232.

Hausman, J. A. 1997. "Valuation of New Goods under Perfect and Imperfect Competition." In T. F. Bresnahan and R. J. Gordon, eds., *The Economics of New Goods*, Chicago: University of Chicago Press, pp. 209-237.

Keeler, T. E. 1996. "Comments." *Brookings Papers on Economic Activity: Micro-economics*, pp. 189-94.

Milgrom, P. and Roberts, J. 1986. "Price and Advertising Signals of Product Quality." *Journal of Political Economy*, Vol. 94, pp. 796-821.

Nelson, P. 1970. "Information and Consumer Behavior." *Journal of Political Economy*, Vol. 78, pp. 311-329.

Newey, W. and McFadden, D. 1994. "Large Sample Estimation and Hypothesis Testing." In Robert Engle and Daniel McFadden, eds., *Handbook of Econometrics IV*, Amsterdam: Elsevier.

Pakes, A. 2001. *"Some Notes on Hedonic Price Indices, with an Application to PC's."* Mimeo, Harvard University.

Reinsdorf, M. 1993. "The Effect of Outlet Price Differentials on the U.S. Consumer Price Index." In M. F. Foss, M. E. Manser, and A. H. Young, eds., *Price Measurement and Their Use*, NBER Studies in Income and Wealth, Vol. 57, Chicago: University of Chicago Press, pp. 227-54.

Trajtenberg, M. 1989. "The Welfare Analysis of Product Innovations with an Application to Computed Tomography Scanners." *Journal of Political Economy*, Vol. 97, pp. 444-479.

_____. 1990. *"Product Innovation, Price Indices and the (Mis)Measurement of Economic Performance."* NBER Working Paper No. 3261.

Schmalensee, R. 1982. "Product Differentiation Advantages of Pioneering Brands." *American Economic Review*, Vol. 72, pp. 349-365.

Stern, S. 1996. *"Product Demand in Pharmaceutical Markets."* Mimeo, Sloan School of Management, MIT.