ELSEVIER

International Journal of Industrial Organization
18 (2000) 1085–1104

International Journal of Industrial Organization

www.elsevier.com/locate/econbase

Barriers to entry, brand advertising, and generic entry in the US pharmaceutical industry

Fiona M. Scott Morton*

Graduate School of Business, 1101 East 58th Street, Chicago, IL 60637, USA

Received 31 October 1997; received in revised form 30 June 1998; accepted 30 November 1998

Abstract

This paper examines the entry decisions of generic pharmaceutical manufacturers into markets opened by patent expiration. In particular, I examine the role of pre-expiration brand advertising to see if it deters generic entry. Other drug characteristics affect the number of entrants; the most important of these is pre-expiration brand revenue. Drugs that treat chronic conditions and drugs that are oral solids attract more entry. The previous literature has assumed advertising is exogenous to the entry decision when analyzing the role of advertising. The results under this hypothesis indicate that brands may affect generic entry very slightly by advertising before patent expiration, but two opposing effects render the result nearly insignificant. When instrumented, the coefficient on advertising is completely insignificant. I conclude that brand advertising is not a barrier to entry by generic firms into the US pharmaceutical market.

Keywords: Advertising; Barriers to entry; Generic pharmaceuticals

1. Introduction

The question of whether advertising acts as a barrier to entry has been a subject of ongoing controversy in the industrial organization literature.[1] Advertising might

* *E-mail address:* fiona.scott.morton@gsb.uchicago.edu (F.M. Scott Morton).

[1] For example, Dixit (1980); Schmalensee (1982), (1983).

0167-7187/00/$ – see front matter
PII: S0167-7187(98)00057-5

disseminate information, thereby increasing market size, and help consumers make rational choices. On the other hand, advertising might merely persuade consumers of product differentiation where none exists. The second type of advertising could act as a barrier to entry. Such a barrier to entry might be profitable to construct if the incumbent has a long period of legal monopoly with a specific date when entry is permitted, as a patented product does in the pharmaceutical industry. Previous work on entry deterrence in the pharmaceutical industry has included advertising to physicians as an explanatory variable in a generic firm's entry decision. However, the brand's advertising choice is endogenous in this context, because both advertising before patent expiration and generic entry depend on expected profits in the post-entry period; and hence, both advertising and entry will be affected by the same forces and the same random errors. This paper analyzes the entry decision more carefully by instrumenting for endogenous, pre-expiration, brand advertising.[2] I conclude that brand advertising is not a barrier to entry by generic firms.

When a pharmaceutical patent expires, generic firms may enter that market and begin selling an exact replica of the original drug. The entry decisions made by many generic firms determine the observed outcome in any given market. Each potential entrant examines the characteristics of the available markets, including those affected by brand behavior such as advertising. The entry decision is based on expected profits, which themselves depend on market characteristics, brand advertising, and the potential entrant's expectations of other generic firms' behavior. For a given revenue stream, a brand producer benefits from fewer generic entrants. The availability of strategies that may be effective in deterring entry is an important issue for producers of branded pharmaceutical products.

Brands have at least two instruments with which to influence the state of the market and generic entry: price and advertising. Limit pricing is a well-known approach that might first occur to an economist thinking about deterrence. A lower branded pharmaceutical price forces a lower generic price; generics cannot sell their product unless they offer a significant price discount. At the time a generic is considering entry, the brand can charge a low price. Generic entrants may conclude that the expected profits do not warrant incurring the fixed costs of entry.

However, a branded product is unlikely to be able to effectively commit to charge a low price *after* patent expiration. Unless such a commitment is made, generic manufacturers calculating expected profits over the life of the drug will not be deterred by a temporarily low brand price. Moreover, because of severe agency problems in the industry, charging a low price in the pre-expiration period will not build up a consumer base that can be exploited later due to habit-formation or other switching costs. Doctors do not share in the cost of the prescriptions they

[2] Throughout the paper I will use the terms advertising and promotion interchangeably. Much 'advertising' in the pharmaceutical industry might more correctly be referred to as promotions to physicians.

write and thus do not have an incentive to consider the price of different products in making a prescription decision. Additionally, many consumers have some type of insurance that weakens incentives based on price. We might suspect that price is not playing its customary role in this market. However, innovator pharmaceutical firms do a great deal of advertising which is primarily directed at the physicians who decide which drugs a patient will use. Therefore, the analysis below will illustrate how the manufacturer of a branded pharmaceutical might employ advertising rather than price to prepare for post-patent competition.

I model the number of generic firms entering a market as a function of market characteristics and brand advertising; market characteristics that are hypothesized to affect entry include revenue of the brand before patent expiration, elasticity of demand, customer mix, switching costs, FDA regulations, and advertising. I test the model on a sample of 98 drugs that lost patent protection between 1986 and 1992. I find that revenue is consistently important in predicting the amount of generic entry, as are dummy variables representing regulatory stringency at the FDA. Increasing the number of consumers with elastic demand, such as hospital buyers and consumers with chronic diseases, increases generic entry.

Previous work has maintained that advertising plays a role in determining the amount of entry into a market.[3] If advertising previous to patent expiration is exogenous to the entry decisions of generic firms, advertising can be included in a model of entry as a regular explanatory variable without instruments. The estimation results under this simple hypothesis indicate that advertising by brands that are about to lose patent protection may actually encourage generic entry, providing support for the market-expanding effects of advertising, but the coefficient is small and unstable. When I assume that advertising before entry is endogenous and instrument for advertising, its coefficient falls and becomes statistically insignificant. I conclude that brand advertising is not a barrier to entry by generic firms.

The organization of the paper is as follows. Section 2 looks at features of demand for prescription pharmaceutical products. Section 3 discusses the previous literature in this area. The industry and its rules of entry are covered in Section 4 as well as a description of my data and data sources. Section 5 develops and estimates a reduced form model of generic entry. Concluding remarks follow in Section 6.

2. Demand for prescription drugs and agency problems

A critical feature of the demand for prescription pharmaceuticals is that the end consumer, the patient, does not select the drug he or she will consume. Instead, the

[3] Schmalensee (1983); Spence (1980).

physician picks the drug therapy and also chooses either the brand or generic form. Pharmaceutical firms spend as much on promotion of brands (much of it directed at physicians) as on R&D.[4] This promotion is intended to influence what the physician prescribes. Hellerstein (1998) in an examination of physician prescribing behavior, finds that ''all of the evidence indicates that physicians are indeed an important agent in determining whether patients receive either trade-name or generic drugs.''[5] Advertising emphasizes the brand's therapeutic advances, its absence of side effects, and the quality and service of the brand's firm.[6] Physicians are also often uninformed as to prices and availability of generics, since generic firms do not advertise very much.

Even if a doctor writes a brand name on the prescription, substituting a generic for the brand may be permitted. Each state in the US regulates generic substitution; it can be forbidden, permitted, or even mandated. By intervening between the physician and the patient, the pharmacy exerts an important influence on whether a brand or generic is sold.[7] The final consumer might also lack incentives; about 40% of prescriptions are covered by some sort of insurance. If the consumer does not pay for the drug, or pays a flat fee, and is at all concerned with quality, he or she may not want to buy a generic even if it is available. Due to these agency problems, price does not have as much influence in the pharmaceutical market as other markets. Instead, advertising is the instrument with which innovator firms can create switching costs and move future demand from the generic to the branded product.

3. Literature review

The structured competition of the pharmaceutical industry which results from the combination of regulation and patented products has attracted interest from economists. The works below attempt to explain the amount of generic entry while including advertising, but do not treat advertising as an endogenous variable.

Telser et al. (1975) investigates the relationship among entry, price changes, and advertising across therapeutic categories. The data he uses are now relatively old, 1963–1972, and since then the regulatory environment has changed significantly.[8] He finds that entry is positively correlated with advertising and the growth rate of sales in 17 therapeutic categories over time, while it is negatively correlated with price changes and the number of firms already in the therapeutic category. He strongly rejects the hypothesis that advertising is a barrier to entry.

[4] Hurwitz and Caves (1988).

[5] Conclusion of the abstract of Hellerstein (1998).

[6] See Leffler (1981) for more on pharmaceutical promotion.

[7] In the early 1990's the majority of prescription drugs, over 60%, were sold through independent drug stores or chain pharmacies.

[8] In addition, he examined entry into a therapeutic class rather than into a chemical compound. This requires a model of differentiated, rather than homogeneous, products.

Hurwitz and Caves (1988) use advertising, price, and number of generic manufacturers to explain market share of a branded drug after it has lost patent protection. They find the market share of the brand increases with its advertising and the number of years it was on patent. A large price differential between generic and brand decreases brand share, as does a larger number of generic suppliers. Their sample differs from those of later periods in that generic drugs were much rarer and conditions for entry stricter. Grabowski and Vernon (1992) use a sample of 18 products achieving over 50 million dollars in sales each at the time of patent expiration. By selecting their sample in this manner, they may find results relevant only to large drug markets. They explain the number of generic entrants with price cost markup, number of years on patent, and advertising of the brand. The markup coefficient is significant in their specification, but advertising is not. Their method of calculating marginal cost, taking the asymptote of price as more generic firms enter, could easily lead to bias. Should perfect competition fail in their two year window, (perhaps because the FDA holds up some entrants) fewer firms will enter, price competition will be less fierce, 'marginal cost' will be higher, the price cost markup will appear to be smaller, and the model will predict less entry. Additionally, endogeneity is clearly a problem because the unobserved expectation of future profits drives the markup, the advertising, and the number of entrants. They also conclude that there is no evidence of entry deterring behavior on the part of incumbents.

Caves et al. (1991) (hereafter CWH) examine 30 products that lost patent protection between 1976 and 1987.[9] Their main concern is to explain price movements, market shares, and quantities sold of both generic and branded drugs. Again, they show that a greater number of generic entrants (instrumented) depresses the generic price and lowers the quantity share of the brand, but they do not model the entry decision directly. The second result of CWH is that brand advertising starts declining two years before patent expiration and then falls substantially with generic entry. The interpretation they lend to the result is that brand advertising does not limit generic competition post-patent; otherwise it would continue after expiration.

4. The pharmaceutical market

Manufacturers fall into two main categories. The first category contains 'pioneer' firms; they undertake research and development to discover new drugs and bring them to market. In order legally to produce and sell the new product in the United States, a pioneer firm must first have an approved New Drug

[9] It is worth noting CWH's treatment of two problems that recur in pharmaceutical data: while they weed out patents that did not hold in practice, they include cases where cross-licensing or simultaneous discovery result in two brands for the same patented drug. I follow the same rules when constructing my dataset.

Application (NDA) which demonstrates safety and efficacy to the satisfaction of the FDA. The second type of firm is a generic or imitator firm that submits Abbreviated New Drug Applications, ANDAs, to the FDA. The ANDA demonstrates that the generic product is 'bioequivalent' to the original branded product. In addition, FDA inspectors examine the equipment the firm plans to use to manufacture the drug, and inspect early batches of the drug.[10] The ANDA application process takes about 18 months from first submission of the application to final granting of permission on average, although the variance is large. Variation in time to approval occurs because firms' applications differ in quality and because the FDA is an unpredictable bureaucracy.

4.1. Data

In this paper, a 'drug' refers to a specific chemical entity that may be called by its generic or brand (proprietary) name. A drug may be manufactured by either the NDA holder or one of many ANDA holders. A firm decides what forms and concentrations to produce; a different (A)NDA is required for each. In my data I characterize forms as solid (e.g. tablets or capsules), injectable, or topical (all other forms including, cream, patch and inhaler). Very different manufacturing equipment is needed for the different forms. Within a form, a drug can come in different concentrations (e.g. 250 mg or 500 mg tablets). A unit of observation is a drug–form combination.[11] The FDA began to keep a uniform record of patent expiration dates in 1984, which forms the base of my dataset. I exclude observations which I could not complete due to lack of data from other sources described below. There remain 81 drugs, but 98 branded drug–form combinations, that can be used as observations in the dataset. The observations break down as follows: 54 oral solid, 18 injectable, 12 other liquid, and 14 miscellaneous.

4.2. Data from IMS America

Revenue and quantity data were supplied by IMS America from their Drugstore and Hospital Audits.[12] I observe revenue and quantity in April, August and December from 2 years before patent expiration to 1 year after expiration for all the drugs that lost patent protection between 1986 and 1991. I also collected monthly advertising data from 3 years before, to 1 year after, patent expiration.

[10] Interviews with industry executives produced estimates of the cost of submitting an ANDA in the early 1990s ranging from 0.25 million to 20 million dollars, much less than the cost of an NDA. The large variation in cost is due to technical differences across drugs and firm–drug combinations.

[11] A more complete description of each observation is included in my original dataset: chemical, labeler, form, concentration, and quantity are listed.

[12] IMS America is a firm which collects and processes different types of pharmaceutical data and sells the information to customers, most of whom are pharmaceutical firms.

IMS records the monthly expenditure on journal advertising and detailing for each drug–manufacturer pair. Each advertising expenditure observation, therefore, encompasses all concentrations of a drug.[13] Detailing is the process whereby drug company representatives make office visits to doctors and present information about certain drugs in person. Printed advertising is published in medical journals that physicians read. Spending on journal advertising is about a third of the amount spent on detailing.[14]

The dollar figures are deflated by the Urban CPI to make prices comparable across years. I rearrange the data into 9 4-month periods, where periods 1–6 contain the data for the 2 years before patent expiration and periods 7–9 have data for the year after patent expiration. The exact date of patent expiration falls in period 7 for each drug, regardless of calendar date of the expiration.

4.3. Data from the FDA

I compiled information from the FDA's *Approved Drug Products* giving the number of ANDAs granted during the period 0, 12 or 24 months after patent expiration. These variables are named *Early, Middle* and *Total*. It is surprising, given the profits that could be earned, how few generic entrants manage to win approval of an ANDA before patent expiration. The mean number of firms entering a market within 1 year after patent expiration (*Middle*) is 1.26. A substantial number of markets, 64 out of 98, experienced no entry within 1 year following patent expiration. Table 1 contains descriptive statistics for the entry and advertising variables.

One problem with measuring approved ANDAs is that a firm may have expected to receive its ANDA at date t and encountered unexpected FDA delays. The months following the generic scandal were just such an occasion; many staff members of the FDA's Generic Drug Division were fired and others became very cautious and slow in approving applications. To control for this problem, I obtained from the FDA the application dates for all approved ANDAs. Instead of looking only at which firms were granted ANDAs near patent expiration, I can measure how many firms applied for that ANDA at some point in time. Unfortunately, application dates for applications that were not ultimately approved are not included in the publicly available dataset. I construct variables measuring how many firms applied (and were eventually approved) for a specific ANDA by

[13] Occasionally IMS divides up a product by form or concentration. For example, the extended release version of a drug is usually a different concentration and might have advertising expenditures that appear separately.

[14] Journal expenditures are measured more accurately because IMS subscribes to every existing medical (including nursing and dental) publication and counts and estimates the cost of the advertisements therein.

Table 1
Descriptive statistics

Dependent variables	Obs	Min	Max	Mean
Total: Number of generic approvals within 2 years of patent expiration	98	0	13	2.09
Middle: Number of generic approvals within 1 year of patent expiration	98	0	13	1.27
Early: Number of generic approvals by patent expiration	98	0	11	0.50
Late: Total-Early	98	0	11	1.59
Applic-Total: Number of generic applications by 6 months after patent expiration	98	0	14	2.05
Applic-Middle: Number of generic applications by 6 months before patent expiration	98	0	14	1.38
Applic-Early: Number of generic applications by 18 months before patent expiration	98	0	7	0.48
Applic-Late: Applic-Total-Applic-Early	98	0	14	1.57
Independent variables and advertising				
Duopoly dummy: 1 if patented market was a duopoly	98	0	1	0.17
Chronic dummy: 1 if drug treats chronic condition	98	0	1	0.55
Injectable dummy: 1 if drug is injected	98	0	1	0.18
Topical dummy: 1 if drug applied topically	98	0	1	0.19
Share Hospital Revenue: Percentage drug revenue from hospital sales (each obs equal weight)	98	0	1	0.23
Revenue: Drug-form revenue in year before patent expiration ('000s)	98	9	349 803	31 613
LnRevenue: Natural log of drug-form revenue (in 10s of millions of $)	98	−9.31	1.25	−2.45
Number Off-Patent Substitutes: Number off-patent brands in therapeutic class	98	0	12	3.62
Journal advertising expenditure 3–2 years before patent expiration (in 10 000 $)	98	0	2.48	0.210
Detailing expenditure 3–2 years before patent expiration (in 10 000 $)	98	0	2.73	0.462
NumberDocs: Number of physicians who might use drug 3 years before expiration	98	2160	246 021	141 546
Company Detail Minutes: Total minutes of corporate detailing, 3 years before patent expiration	92	0	8066	2906
Other Form on Patent Dummy: 1 if another form of the drug still has patent protection	98	0	1	0.112
Months on Patent: Number of months the drug has been sold while patent-protected	98	8	346	123

18 months preceding patent expiration (the average time to approval), by 6 months before expiration, and not later than 6 months after expiration. These variables are named *Applic-Early*, *Applic-Middle* and *Applic-Total*, respectively.

The average length of the approval process over all years and products in my data is 18 months. The generic scandal (when bribery was discovered at the FDA) significantly raised approval times; all years since 1989 have higher average approval lengths than those before the scandal (this also may be affected by the composition of drugs going through the process in any given year). In this paper I analyze both approval and application dates since approval time is affected by the amount of effort and resources put into the application. In addition, if a drug is known to have a short approval time, interested firms will submit ANDAs relatively late; in that case approvals will be a better measure of the amount of generic entry than applications. Table 2 shows the distribution of entry over time. The measures that are 18 months apart show considerable similarity: see *Applic-Middle* and *Middle*, for example. Since all the application dates I have represent ANDAs that were eventually approved, the entry and application variables are correlated; the correlations range from 0.42 to 0.82. In the interests of brevity, the results presented here only examine *Total* and *Applic-Total*.[15] Matching all datasets eliminated drugs that were not defined consistently across data sources or were omitted from one or more sources, leaving 98 usable drug forms.

4.4. The decision to enter

A generic firm deciding whether to enter a pharmaceutical market has to form an estimate of expected market profit by considering issues of demand, supply (other entrants), and regulatory uncertainty. Because there is a fixed cost of entry (the cost of activities associated with filing the ANDA), a firm will want to enter only if it expects the markup over variable costs and quantity sold to be large enough to cover its fixed costs. The larger expected profit in the market ($E[\Pi]$), the more generic firms (N) we expect to see entering. In equilibrium the correct number of entrants enters for the size of expected profit, and a zero-profit condition holds.[16] This relationship is shown in the simple reduced form equation that follows:[17]

$$E[N_i] = f(E[\Pi_i])$$

[15] Results using the other dependent variables are similar and can be obtained from the author.

[16] Future supply and the resulting markup are also critical to the entry decision, but I make a simplifying assumption of equilibrium outcomes in this paper.

[17] Bresnahan and Reiss (1988), (1990), (1991) develop a framework to determine how many firms a market will support and examine the ratio of monopoly market size to duopoly market size. In their work, the population of the different small-town markets serves as the primary measure of underlying market size. It is difficult to construct an analogous 'population' measure for a particular molecule and additionally model per-person demand using market characteristics and prices, so I do not use their approach.

Table 2
Distribution of markets by amount of entry at three moments in time (T = patent expiration date)

Generic entrant count	No. ANDAs $T+0$ months = Early	No. ANDAs $T+12$ months = Middle	No. ANDAs $T+24$ months = Total	Applications $T-18$ months = Applic_Early	Applications $T-6$ months = Applic_Middle	Applications $T+6$ months = Applic_Total
0	79	64	51	78	65	51
1	10	9	12	9	7	15
2	2	9	8	5	6	7
3	2	3	2	3	4	4
4	3	5	8	0	3	0
5	0	2	3	1	5	8
6	1	1	6	0	4	4
7	0	0	0	2	0	1
8	0	1	1	0	1	2
9	0	1	1	0	0	1
10	0	1	1	0	1	0
11	1	0	1	0	1	1
12	0	1	2	0	0	1
13	0	1	2	0	0	1
14	0	0	0	0	1	2
Obs	98	98	98	98	98	98

Number of later entrants in markets by amount of early entry

	Early = 0				Early > 0			
	Mean	S.D.	Max	Obs	Mean	S.D.	Max	Obs
Total	1.22	2.34	11	79	5.74	4.12	13	19
Late	1.22	2.34	11	79	3.16	2.89	10	19

The estimate of the future profitability of a particular compound is complex and difficult to form. For example, the demand for a specific drug might grow over time as the population ages. Therapeutic substitutes for a specific drug might not exist, or they might be about to lose patent protection themselves. A pre-existing drug could be found to have new properties, increasing the demand for it.[18] The price elasticities of demand of the drug's consumers will determine how readily they will buy a generic. Switching costs (if a patient is already using the brand) is one of the factors determining how quickly the generic gains market share. Another important part of total profitability is the technological difficulty of manufacturing the drug and receiving FDA approval.[19]

5. An empirical examination of entry

The most important variable in predicting entry will clearly be the proxy for market size since this is likely to be correlated with profits. The reason profit is correlated with revenue is because marginal costs tend to be low in the pharmaceutical industry. I created the variable *LnRevenue* by taking the log of the sum of the dollar revenue, measured in 10s of millions of dollars, of all strengths and quantities within one chemical-form (in other words, summed over concentration) during the 12 months before patent expiration. An additional factor that will affect entry is the form of the product. Generic firms usually specialize in producing injectables or non-injectables. Injectables require special equipment compared to solids and the plant must meet a high FDA standard of cleanliness. Also, the capital equipment required to produce solids is different from that required for other forms such as sprays or creams, which alters the fixed cost of entry. Two dummy variables, *Injectable* and *Topical*, control for form relative to the omitted category, oral solids.

Share Hospital is the revenue share of that drug form sold to hospitals rather than drugstores. The mean hospital share for injectables is 68.5% while it is 12.6% for non-injectables. If hospitals are more willing to buy generics, a market with a large *Share Hospital* will be more attractive to entry, all else equal.

Duopoly is a dummy variable given the value one when two firms make a

[18] For example, some ulcer drugs are also good antacids.

[19] The generic firm must also decide what strengths to make and the marginal cost of applying for an additional strength is low. However, a separate application must be submitted and appropriate lab work done and batches tested for each concentration, so concentration entry has a positive cost. In my dataset generic firms apply for all the concentrations of a drug in all instances, although approval for one concentration might lag a month behind. To avoid weighting the data by the number of concentrations associated with each drug, I do not include a separate observation for each concentration. Instead, multiple concentrations of the same drug are merged into one observation (revenue is combined) and included in the dataset used for estimation.

branded product under the same patent. Duopoly markets are an unusual feature of the pharmaceutical industry; two firms both have rights to the same patent. The products they produce are chemically identical, but have different brand names. Observations in this category (17 out of 98 observations) sum revenue and advertising values for both of the individual firms. However, we might expect duopoly markets not to act like two monopoly markets. Price–cost margins will be lower if the two firms are competing; this means that the relationship between revenue and attractiveness for entry may differ between monopoly and duopoly markets. Secondly, advertising may differ because the two firms compete on advertising, or free-ride on each other's advertising, or try to create differentiation with advertising. Because this is a complicated topic beyond the scope of the paper, I do not model it but simply include the dummy and its interaction with drug revenue to capture market size.

An important characteristic affecting entry is the extent of therapeutic substitution available for a particular drug. If there exists an off-patent therapeutic substitute that is a better treatment than the drug in question, fewer generic firms will be interested in entering the market. The timing of patent expirations of a drug's therapeutic substitutes will also affect how many generic firms enter. Therapeutic substitution is a difficult attribute to measure; drugs are good substitutes on some occasions for some patients, and not for others. It is therefore very difficult to find a good explanatory variable that will account for therapeutic substitution. A reasonable choice might be the number of brands in IMS' definition of therapeutic class.[20] I count the number of brands in a drug's therapeutic class and record their patent status. The number of already expired brands should affect the amount of generic entry negatively. I define *Substitutes Off Patent* to be the number of already expired brands in the same therapeutic class as defined by IMS.

I also determined which drugs in my dataset treated *Chronic* or acute conditions.[21] To the extent the patient has input about the choice of drug, medications for chronic diseases should have more price elastic consumers; repeat purchasers have greater opportunity to collect information about prices and substitutes. On the other hand, patients being treated for a chronic condition with a branded drug may be reluctant to risk adjustment costs incurred by switching to a generic.

Advertising expenditure is measured very simply by adding up expenditures on journal advertising and detailing in the second and third years before patent expiration, when generics are making their entry decisions.[22] Finally, I include

[20] IMS' categories are somewhat arbitrary, but largely consistent.

[21] Medical reference texts provided most information. David Bailey of MIT's Health Services was very helpful.

[22] Convergence is easier with roughly normalized variables. *Advertising* is expressed in 50 million dollar units so it lies between 0 and 4.

year fixed effects; a drug is identified by year of patent expiration. I am particularly concerned with capturing the effects of the generic scandal of 1989 and the subsequent decline in ANDA approvals. Table 1 presents descriptive statistics for independent and dependent variables and instruments.

5.1. Effect of advertising

How should we expect brand advertising to affect the amount of generic entry? A traditional interpretation of advertising is that it expands the market for the product. The market-expansion role of pharmaceutical advertising applies principally at the beginning of the product lifecycle when the drug is new to doctors. Detailing and journal advertising expenditures are an order of magnitude or more higher at the introduction of a new product than they are later. By the end of patent life, advertising that is 'expanding the market' is defending a drug's share against its therapeutic substitutes.

Standard models of entry deterrence such as excess capacity accumulation (Spence, 1977) can be applied using advertising as the choice variable. Reducing expected profits for entering generic firms by shrinking their market with advertising will deter entry. It is clear that two correlations between advertising and entry are possible. Existing advertising might be establishing switching costs with doctors and reducing profits available to generics in the future, making the market less attractive to enter. Alternatively, advertising might be high because the market looks profitable for the future; defending against therapeutic substitutes with investment in advertising is worthwhile (Dorfman and Steiner, 1954). Such an attractive market would encourage generic entry. In other words, the correlation between advertising and entry will be negative if brand advertising is a barrier to entry. A positive relationship will arise if attractiveness of the future market causes a high level of brand advertising and increased generic entry. Because it is difficult to include appropriate variables measuring the likely success of entry deterrence and the future attractiveness of therapeutic markets, the inclusion of advertising as an exogenous explanatory variable will likely result in biased coefficients.

5.2. Specification

Since we are counting number of entrants into each expired patent, one reasonable choice for $f(.)$ is the exponential, making the model a Poisson count data model.[23] I use a standard Poisson specification where $E(N_i|X_i)=f(X_i).f(.)$ is parametrized as follows:

[23] I have some overdispersion in my data; the variance of an entry measure is larger than its mean. I have corrected the standard errors in the tables to allow for misspecification of the conditional variance as in Wooldrige (1991).

$$N_i = \exp\Bigg(\beta_0 + \beta_1 \text{Advertising}_i + \beta_2 \text{LnRevenue}_i + \beta_3 \text{Duopoly}_i \cdot \text{LnRevenue}_i + \beta_4 \text{Duopoly}_i + \beta_5 \text{Chronic}_i + \beta_6 \text{Share Hospital}_i + \beta_7 \text{Topical}_i + \beta_8 \text{Injectable}_i + \beta_9 \text{Substitutes Off Patent}_i + \sum_{t=1987}^{1991} \alpha_t \text{Year}_{it}\Bigg)$$

Table 3 reports the specification using Total and Applic-Total as dependent variables and assuming advertising is exogenous.[24] Journal advertising appears to decrease the number of expected entrants into a market while detailing expenditures positively affect the number of entrants. The advertising coefficients reflect *correlation* not causality, since we do not know which way the causality runs. The coefficients are close to zero, perhaps because of the countervailing effects discussed above; a coefficient of 0.3 on detailing means even a 10 million dollar increase in detailing has a negligible effect on the expected size of *Applic-Total*.

The coefficients of a Poisson regression do not provide estimates of the marginal effect of changing an explanatory variable. When using a Poisson functional form, the marginal effect of variable *i* for the mean observation in the data is given by $\hat{\beta}_i \exp(\bar{X}\hat{\beta})$.[25] I report the value of the exponential term for each regression in the last row of the table. The marginal effect for the average observation can be calculated by multiplying the coefficient of interest times this constant. For example, the effect on the number of generic entrants of a dummy variable moving from zero to one is simply its marginal effect; from Table 3, column one, a *Chronic* market increases the number of generic entrants by 0.64 ($0.823 \times 0.783 \times 1$). Moving a continuous variable such as Hospital Share from its mean (0.23) to its max (1) using the coefficients in the same column would produce a marginal effect of 0.9 ($1.5 \times 0.783 \times 0.77$).

Ln Revenue is significant in explaining entry across all four dependent variables. The predicted number of entrants rises by about 0.5–0.7 if a drug with 10 million dollars of revenue doubles its revenue and its other characteristics stay fixed. *Share Hospital*'s coefficient is generally positive and significant. The results imply that an otherwise average drug with 100% hospital share attracts almost one more entrant than an equivalent drug with 0% hospital share. The coefficient on *Duopoly*×*LnRevenue* is significant and negative. A duopoly market attracts a smaller amount of generic entry as its revenue increases when compared to an

[24] The full set of results (all dependent variables and base case without any advertising variables) is available from the author on request.

[25] Because the exponential function is nonlinear, note that this marginal effect will, in general, be different from the average marginal effect in the sample.

Table 3
Determinants of generic entry assuming advertising exogenous[a]

Dependent variable	Total	Total	Applic-Total	Applic-Total
Detailing	0.564[b]	–	0.312[b]	–
	(0.165)		(0.148)	
Journal Advertising	–	−0.079	–	−0.420
		(0.359)		(0.427)
Ln Revenue	0.538[b]	0.674[b]	0.633[b]	0.771[b]
	(0.108)	(0.089)	(0.079)	(0.087)
Share Hospital	1.50	1.28[b]	1.30[b]	1.18[b]
	(0.628)	(0.462)	(0.425)	(0.428)
Duopoly	−0.567[c]	−0.206	−0.280	0.236
	(0.338)	(0.418)	(0.320)	(0.362)
Duopoly×Ln Revenue	−0.631[b]	−0.601[b]	−0.455[b]	−0.317[b]
	(0.111)	(0.122)	(0.141)	(0.143)
Topical	−1.46[b]	−1.35[b]	−1.61[b]	−1.64[b]
	(0.264)	(0.202)	(0.175)	(0.225)
Inject	−0.990[c]	−0.683[c]	−0.553	−0.300
	(0.542)	(0.398)	(0.378)	(0.340)
Chronic	0.823[b]	−0.975[b]	0.581[b]	0.673[b]
	(0.238)	(0.184)	(0.212)	(0.247)
Substitutes Off Patent	0.031	0.029	0.056[c]	0.054
	(0.044)	(0.039)	(0.033)	(0.037)
Dum87	−0.498	−0.301	−0.215	−0.086
	(0.301)	(0.320)	(0.283)	(0.277)
Dum88	1.02[b]	0.659[c]	0.985[b]	0.687[c]
	(0.348)	(0.366)	(0.290)	(0.385)
Dum89	−0.761	−0.443	−0.281	−0.052
	(0.507)	(0.639)	(0.411)	(0.677)
Dum90	−0.952[b]	−0.868[b]	−0.696[b]	−0.695[b]
	(0.267)	(0.282)	(0.251)	(0.329)
Dum91	−1.40[b]	−1.23[b]	−1.31[b]	−1.39[b]
	(0.289)	(0.274)	(0.259)	(0.236)
Dum92	−2.31[b]	−1.58[b]	−2.01[b]	−1.30[b]
	(0.342)	(0.384)	(0.335)	(0.395)
Constant	0.942	1.31	1.18	1.54
	(2.33)	(1.49)	(1.72)	(1.10)
N	98	98	98	98
Pseudo R^2	0.476	0.456	0.479	0.479
$\exp(\bar{X}\hat{\beta})$	0.783	0.840	0.780	0.753

[a] Standard errors are in parentheses.
[b] Indicates significant at at least the 5% level (2-tail test).
[c] Indicates significant at between the 5% and the 10% level.

otherwise identical market; the coefficient on the *Duopoly* dummy is generally insignificant. One explanation for this result is that competition after expiration is expected to be stronger due to the additional firm in the market. The existence of off-patent therapeutic substitutes (*Substitutes Off Patent*) in the brand's therapeutic

class does not appear to affect the amount of generic entry.[26] I conclude from this that it is very difficult to construct a meaningful measure of therapeutic substitutes; the coefficient is not even negative, as theory would suggest. Dummy variables for zero or one substitute did not perform any better. *Topical* drugs consistently show fewer entrants; perhaps the long approval times create higher fixed entry costs. *Chronic* has a positive and significant coefficient in all but one specification; the greater elasticity of demand of 'chronic' consumers makes chronic markets more attractive, but the unwillingness of customers to switch drugs works in the opposite direction. It appears from these results that the elasticity effect of *Chronic* is stronger.

The year dummies pick up two effects, the entry error common to all the drugs expiring in a given year and FDA approval speed. FDA approval speed is not related to the number of applicants, hence we can interpret the year coefficients as simple year fixed effects. The coefficients in Table 3 have the expected sign for each cohort of applicants, given the timing of the generic scandal. The effects of the scandal are evident in negative coefficients on *Dum90*, *Dum91* and *Dum92*.

The discussion above suggests that including advertising as an exogenous variable may produce biased coefficients on advertising. In the next set of regression results advertising is instrumented in order to estimate a consistent effect of advertising on the amount of generic entry. The instruments I use include the amount of time the drug has spent on patent, which is a measure of where expiration falls in the lifecycle of the drug. Another instrument is a dummy variable indicating whether the firm has other forms of the same drug still under patent protection. IMS lumps the advertising expenditures for the two different forms together and thus advertising expenditure will be higher than normal for these observations. The manufacturer's total detailing expenditure across all of its drugs during the third year before patent expiration is the third instrument. Firms that have big detailing forces in operation for promoting other drugs will find it less costly to undertake some marginal advertising for the drug losing patent protection. The final instrument is the number of physicians that could be expected to use a particular drug. This measure is simply the total number of physicians that report themselves in one of a number of subspecialties that would use the drug. It indicates how costly advertising the drug will be; journal and detailing expenditures must be higher to reach more physicians.[27]

There is no reason to expect these instruments—years on patent, other forms being sold, total firm promotions, and number of target physicians—to be

[26] *Substitutes Off Patent* was the only variable to lose significance after robust standard errors were calculated.

[27] The correlation between these instruments and journal advertising and detailing have the expected sign (years is negative) and are all significant except those for between journal ads and firm detailing and number of specialists.

correlated with the error in the promotion and entry decisions due to unobserved profitability of a market. Expected profit will drive generic entry and will also affect promotional expenditures on the drug. The cost-based measures of promotion (economies of scale and number of physicians) will affect the firm's choice of

Table 4
Determinants of generic entry assuming advertising endogenous[a]

NLIV, dependent variable	Total	Total	Applic_Total	Applic_Total
Detailing	–	0.136	–	−0.115
		(0.684)		(0.767)
Journal Advertising	−0.298	–	−0.300	–
	(1.50)		(1.40)	
Ln Revenue	0.668[b]	0.613[b]	0.768[b]	0.765[b]
	(0.268)	(0.235)	(0.293)	(0.308)
Share Hospital	1.17	1.27	1.20	1.20
	(1.26)	(1.23)	(1.23)	(1.21)
Duopoly	0.041	−0.194	0.254	0.182
	(0.829)	(0.548)	(0.866)	(0.736)
Duopoly×Ln Revenue	−0.519[c]	−0.543[b]	−0.288	−0.287
	(0.288)	(0.266)	(0.389)	(0.432)
Topical	−1.41[b]	−1.48[b]	−1.81[b]	−1.77[b]
	(0.720)	(0.741)	(0.915)	(0.885)
Inject	−0.595	−0.717	−0.326	−0.307
	(0.986)	(0.982)	(0.877)	(0.925)
Chronic	0.942[c]	0.885	0.625	0.621
	(0.552)	(0.517)	(0.505)	(0.518)
Substitutes Off Patent	0.010	0.003	0.020	0.023
	(0.079)	(0.081)	(0.076)	(0.083)
Dum87	−0.251	−0.227	−0.039	0.039
	(0.447)	(0.361)	(0.435)	(0.416)
Dum88	0.648	0.800[c]	0.750[c]	0.746
	(0.415)	(0.496)	(0.399)	(0.549)
Dum89	−0.182	−0.175	0.224	0.289
	(0.473)	(0.438)	(0.395)	(0.427)
Dum90	−0.825[b]	−0.832[b]	−0.652[c]	−0.635[c]
	(0.397)	(0.375)	(0.372)	(0.378)
Dum91	−1.20[b]	−1.17[b]	−1.31[b]	−1.20[b]
	(0.577)	(0.497)	(0.664)	(0.515)
Dum92	−1.31	−1.68	−1.31	−1.48
	(1.28)	(0.956)	(1.34)	(0.986)
Constant	1.45	1.31	1.64	1.61
	(0.794)	(0.762)	(0.801)	(0.782)
N	92	92	92	92
R^2	0.669	0.710	0.741	0.737
$\exp(X\beta)$	1.05	1.04	0.844	0.853

[a] Standard errors are in parentheses.
[b] Indicates significant at at least the 5% level (2-tail test).
[c] Indicates significant at between the 5% and the 10% level. Advertising variables are insignificant in specifications using all dependent variables.

the latter but will not be correlated with any demand-side error I estimate at the market level. Similarly, years on patent and other forms being sold reflect the firm's internal promotion pattern, but they are determined by technological advances and are therefore not correlated with the part of expected profit that the econometrician does not observe.

The most appropriate technique for estimating this regression with an endogenous variable is Nonlinear Two Stage Least Squares. The nonlinear function is simply the exponential, so the equations remain in a Poisson framework. The results are reported in Table 4. All coefficients decrease in significance compared to previous specifications without instrumentation, while their marginal effects rise slightly. These results show unambiguously that advertising has no effect on the amount of generic entry into a market.[28] As a final experiment, I examined whether advertising affects the total market share gained by generic firms, even if it does not affect the number of generic firms. I find that advertising is consistently unable to explain the market share lost by the brand to generic entrants.[29] However, the number of entrants and (instrumented) relative price are significant in predicting brand share.

6. Conclusion

Characteristics of a drug market before the brand's patent expires are significant predictors of generic entry. The revenue of the brand in the year before patent expiration is the most important factor determining the amount of generic entry; higher revenues attract a greater number of entrants. The extent to which a branded drug is purchased by hospitals positively affects the number of entrants. Drugs which treat chronic conditions also attract more generic firms; the patients buying these drugs may have a higher elasticity of demand. Topical forms experience one or two fewer entrants on average. Duopoly markets are relatively less attractive as market size increases. Further research into advertising and markup behavior in duopoly markets is needed to fully understand this finding. Not surprisingly, the stringency of FDA regulatory procedure negatively affects the number of generic entrants observed in a market. In particular, the generic scandal of 1989 caused a significant decrease in the number of generic firms approved around patent expiration.

The previous literature has assumed that advertising before patent expiration is

[28] The entry-deterring effect of advertising might only be present in a small market where the incumbent has a chance to deter the only entrant. I cannot interact advertising with certain size markets because the number of observations drops steeply. A researcher would need a much larger sample size to investigate sub-groups thoroughly.

[29] The full set of results is available from the author.

exogenous with respect to the amount of generic entry. I estimate that model and find that the coefficient on advertising is very small, and its sign varies with the type of advertising. Advertising may affect entry if it creates a barrier to entry for generic firms, or the advertising may reflect the same market conditions that affect entry: good markets are worth advertising in and entering. The small estimated coefficient on advertising thus may be the result of two opposing effects. Because both advertising and entry depend on the unobserved expectation of profits in the post-entry market, one should not assume advertising is exogenous. I assume it is endogenous and instrument for incumbent advertising in an equation predicting generic entry. I find that the coefficient on advertising is insignificantly different from zero. Additional regressions examining brand market share after patent expiration show no effect of brand advertising on the market share retained by the brand. I conclude that brand advertising is not a barrier to entry in the US pharmaceutical industry.

Acknowledgements

I would like to thank Ernie Berndt, Iain Cockburn, Sara Ellison, two anonymous referees, and seminar participants at MIT and the NBER for helpful comments. I appreciate the advice of Jerry Hausman and Nancy Rose in particular. All remaining errors are mine. I also thank IMS America for permission to use their data, and Merck and Co. for providing access to that data. I am grateful for financial support from the Sloan Foundation through the Program on the Pharmaceutical Industry, Sloan School, MIT.

References

Bresnahan, T., Reiss, P., 1988. Do entry conditions vary across markets? Brookings Papers on Economic Activity.

Bresnahan, T., Reiss, P., 1990. Entry in monopoly markets. Review of Economic Studies 57, 531–553.

Bresnahan, T., Reiss, P., 1991. Entry and competition in concentrated markets. Journal of Political Economy 90, 51.

Caves, R., Whinston, M., Hurwitz, M., 1991. Patent expiration, entry, and competition in the US pharmaceutical industry, Brookings Papers.

Dixit, A., 1980. The role of investment in entry-deterrence. Economic Journal 90, 95–106.

Dorfman, R., Steiner, P., 1954. Optimal advertising and optimal quantity. American Economic Review 44, 826–836.

Grabowski, H., Vernon, J., 1992. Brand loyalty, entry, and price competition in pharmaceuticals after the 1984 Drug Act, Journal of Law and Economics 35.

Hellerstein, J., 1998. The importance of the physician in the generic versus trade-name prescription decision. RAND Journal of Economics 29 (1), 108–136.

Hurwitz, M., Caves, R., 1988. Persuasion or information? Promotion and the shares of brand name and generic pharmaceuticals, Journal of Law and Economics 31.

Leffler, K., 1981. Persuasion or information? The economics of prescription drug advertising. Journal of Law and Economics 24, 45–74.

Schmalensee, R., 1982. Product differentiation advantages of pioneering brands. American Economic Review 72, 349–365.

Schmalensee, R., 1983. Advertising and entry deterrence: An exploratory model. Journal of Political Economy 91, 636–653.

Spence, A.M., 1977. Entry, capacity, investment and oligopolistic pricing. Bell Journal of Economics 8, 534–544.

Spence, A.M., 1980. Notes on advertising, economies of scale, and entry barriers. Quarterly Journal of Economics 95 (3), 493–507.

Telser, L., et al., 1975. The theory of supply with application to the ethical pharmaceutical industry. Journal of Law and Economics 18, 449.

Wooldrige, J., 1991. On the application of robust, regression-based diagnostics to models of conditional means and conditional variances. Journal of Econometrics 45, 5–46.