**Title**
Vertical Integration and Market Entry in the Generic Pharmaceutical Industry



**Author**
Kubo, Kensuke

**Publication Date**
2011

Peer reviewed|Thesis/dissertation

**Vertical Integration and Market Entry in the Generic Pharmaceutical Industry**

by

Kensuke Kubo

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Brian D. Wright, Chair

Professor Jeffrey M. Perloff

Professor Sofia B. Villas-Boas

Professor J. Miguel Villas-Boas

Spring 2011

**Vertical Integration and Market Entry in the Generic Pharmaceutical Industry**



by

Kensuke Kubo

## Abstract

Vertical Integration and Market Entry in the Generic Pharmaceutical Industry

by

Kensuke Kubo

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Brian D. Wright, Chair

This dissertation explores the relationship between vertical integration and market structure formation. It does so by combining the empirical industrial organization literature on vertical integration with that on market entry. The first empirical essay, Chapter 2, explores the motives for vertical integration in the US generic pharmaceutical industry. The industry is made up of numerous drug markets that open up to competition among generic manufacturers at different points in time. Each market consists of an upstream segment that manufactures active pharmaceutical ingredients, and a downstream segment that processes the active pharmaceutical ingredients into finished formulations and supplies them to final consumers. The econometric analysis shows that vertical integration in the generic drug industry is characterized by bandwagon behavior. While bandwagon effects have been widely discussed in the vertical integration literature, this study is one of the first to present empirical evidence on its existence. The analysis also indicates that vertical integration is partly driven by the need for a particular form of relationship-specific non-contractible investment – the early development of active pharmaceutical ingredients by upstream units. The relationship-specificity of such investments is greater in markets where generic firms try to enter by challenging the patents held by originator pharmaceutical companies. I find that in such markets, individual firms have a higher propensity to vertically integrate.

The second empirical essay, presented as Chapter 3, introduces an econometric model of a vertical entry game. The model is used to estimate rival effects – the effect of rival entry on the post-entry profits of individual firms. These estimates allow us to make inferences about the competitive effects of vertical integration. Application of the model to the generic pharmaceutical industry yields the following result: vertical integration has significant efficiency effects that benefit unintegrated downstream firms. This implies that vertical integration is likely to be procompetitive from a static point of view. The parameter estimates are used to simulate the impact of a hypothetical policy that bans vertically integrated entry. The results indicate that such a ban tends to reduce the equilibrium number of downstream entrants. This suggests that the effect of vertical integration on market structure formation is also procompetitive.

To My Wife Azusa.

# Contents

# List of Figures

# List of Tables

# Acknowledgments

It is hard to describe how grateful I am to each of my advisers as I complete this dissertation. Brian Wright has been the most supportive thesis adviser imaginable. I can never thank him enough for his patience and generosity during an arduous, albeit rewarding process. Brian's uncompromising pursuit of quality in research, as well as his belief in economics as a powerful tool for improving society, will be an inspiration throughout the remainder of my career.

Jeff Perloff, to whom I owe my interest in industrial organization, gave me words of encouragement when they were most needed. Current, past, and future students of ARE are fortunate to have him as a teacher and mentor. Sofia Villas-Boas taught me the excitement of applying new empirical methods to one's economic problem. In particular, she provided the crucial suggestion of using a market entry framework to study the effects of vertical integration. Her passion for economic research has amazing spillover effects upon all of her students. Miguel Villas-Boas provided advice, emanating from his encyclopedic knowledge of theory and empirical methodology, that has always proved very valuable. I benefited immeasurably from our sessions at Haas.

I would also like to mention the guidance I received from the late Jenny Lanjouw. In addition to introducing me to the world of pharmaceutical economics, she has been, and will continue to be, a shining role model. I have also been fortunate to receive excellent advice from scholars outside of Berkeley, especially Nancy Gallini, Toshiaki Iizuka, and my colleagues at the Institute of Developing Economies. I also wish to express my heartfelt thanks to the amazing staff of ARE, especially Gail Vawter and Diana Lazo.

The empirical analysis would not have been possible without high quality data. I thank Thomson Reuters for access to their pharmaceutical database; discussions with Mike Chace-Ortiz, Yoshiko Tanahashi, and Ken Ueki were especially helpful for understanding the database as well as the generics business. I would also like to acknowledge the support of the Agency for Healthcare Research and Quality (AHRQ); part of the research was conducted at its Center for Financing, Access, and Cost Trends (CFACT) Data Center. The results and conclusions in this dissertation are, however, those of the author and do not indicate concurrence by AHRQ or the Department of Health and Human Services.

My family deserves much of the credit for the completion of this dissertation. My mother made sure that my two brothers and I received the best education, and that we kept our priorities in order. I am infinitely grateful for her enduring support and encouragement. Finally, my wife Azusa has provided the optimism that has really pushed things forward. I am grateful for her selfless support, and I look forward to realizing all of our postponed adventures together.

# Chapter 1

# Introduction

Since the work of Bain (1956), the formation of market structure and its effect on market outcomes has been a central topic in industrial organization. In many industries, market structure is not only about the number of entrants and their respective market shares; there is also an important vertical aspect. In particular, the role played by vertical integration – the combination of two or more vertically related functions within the same firm – has been a topic of active theoretical research. Authors have recognized the ability of vertical integration to influence market structure formation by deterring or facilitating entry. Vertical integration may also have a direct impact on market outcomes. For instance, two markets with similar horizontal market structures may have different price levels if the firms in one of them have a higher degree of integration into a vertically related activity.

While many empirical studies have examined the relationship between vertical integration and market outcomes, with a few exceptions (e.g., Hortaçsu and Syverson, 2007), vertical integration has not been treated explicitly as part of the market structure formation process. Meanwhile, the empirical entry literature has so far focused on horizontal interactions among firms; vertical interactions, including decisions to integrate, have not been explicitly incorporated into the econometric analysis of entry.

The present dissertation fills this gap in the literature by combining the analysis of vertical integration with that of market entry. There are two benefits from doing so. First, the incorporation of vertical integration into the analysis of entry behavior lets us obtain a more accurate understanding of the market structure formation process. Second, utilizing an empirical framework based on market entry behavior allows us to investigate the motives for, and effects of, vertical integration in new and useful ways.

The two empirical essays in this dissertation analyze market entry and vertical integration in the US generic pharmaceutical industry. Generic pharmaceuticals are drug products that become available to consumers after the expiration of patents and other market exclusivities that protect the original product. The industry provides a good setting for studying vertical market structure formation because it consists of many markets – one for each original drug – made up of two vertical segments. The upstream segment manufactures active pharmaceutical ingredients (APIs) and the

downstream segment processes APIs into finished formulations to sell to final consumers. In each market, multiple generic firms simultaneously choose their entry and vertical integration actions. Therefore, the industry provides a large number of market observations where vertical market structure formation takes place through the simultaneous and collective actions of individual firms.

The first empirical essay, Chapter 2, seeks to explain the increased prevalence of vertically integrated entry in the generics industry since the late 1990s. Using a firm-level dataset covering 85 markets that opened up to generic competition between 1999 and 2005, I investigate the determinants of a generic firm's decision to vertically integrate. I find that a firm has a higher probability of vertically integrating, conditional on its decision to enter the downstream segment, if it has greater past entry experience in the upstream API segment. This suggests that a firm's upstream experience lowers its cost of vertical integration. I also find that a firm is more likely to vertically integrate when the average upstream experience among its rivals is higher. This effect can be divided into two parts. First, higher upstream experience among rivals implies a greater incidence of vertical integration in the equilibrium market structure. Second, the expectation of a more vertically integrated market structure raises the incentive for an individual firm to become vertically integrated. The latter effect suggests that vertical integration is characterized by bandwagon behavior. While bandwagon effects have been widely discussed in the theoretical literature, and anecdotal accounts of bandwagon behavior is not difficult to find, this result represents one of the first pieces of empirical evidence on its existence.

The analysis also finds that generic firms are more likely to be vertically integrated in markets where they try to enter by filing a "paragraph IV certification" that challenges the patents held by originator pharmaceutical companies. Generic entrants have an incentive to engage in such patent challenges, because the first-to-file paragraph IV entrant may be awarded a 180-day exclusivity in the generic market. I argue that in markets characterized by paragraph IV patent challenges, upstream investment into API development tends to be relationship-specific. This is because in such markets, the API has a much higher value if it is used by the first-to-file entrant than when it is used by some other firm. Such relationship specificity does not exist in other generic drug markets. Therefore, the higher relationship specificity of upstream investments in paragraph IV markets is likely to explain the higher incidence of vertical integration in such markets.

Chapter 3 is another empirical essay. It specifies the formation of vertical market structure in generic drug markets as the outcome of a simultaneous-move vertical entry game. Firms choose their actions from a set containing up to four elements: unintegrated downstream entry, unintegrated upstream entry, vertically integrated entry, and no entry. The actions of rival firms enter the payoff functions of potential entrants so that vertical rival effects are measured. The estimated rival effects are then used to make inferences about the competitive effects of vertical integration.

An econometric model of the vertical entry game is estimated using a dataset consisting of 85 markets that opened up during 1993-2005. Markets that are subject to paragraph IV patent challenges are not included in the analysis, because the entry process in such markets is characterized by a race to be first rather than a simultaneous-move game. The estimates suggest that vertical integration by rival entrants has a significantly positive impact on the payoffs of uninte-

grated downstream entrants. This implies that vertical integration has strong efficiency effects that spill over to benefit unintegrated downstream firms. I also find that the profit of an unintegrated upstream entrant falls when, in a market structure consisting of two upstream firms and one downstream firm, the other firms become vertically integrated. This finding is also consistent with the existence of efficiency effects.

The usefulness of the vertical entry model lies in its ability to accommodate policy simulations based on estimated parameters. In one such simulation, it is found that a policy that bans vertically integrated entry tends to decrease the number of downstream entrants in equilibrium. Combined with the finding that vertical integration has significant efficiency effects, this result supports the notion that vertically integration plays a procompetitive role in the generic drug industry.

# Chapter 2

# Explaining Vertical Integration in the Generic Pharmaceutical Industry

## 2.1 Introduction

While vertical integration is a feature of many businesses, its incidence or prevalence varies across industries, across different markets in the same industry, and among firms operating in the same market. Explaining such variation in vertical integration has long been an active area of industrial organization research.

The motives for vertical integration identified in the theoretical literature can be grouped into two major categories: (i) improvement of efficiency for the integrating firm and (ii) foreclosure of rival firms from the supply of an input or from access to consumers. Each category is further divided into sub-categories. For instance, efficiency motives include the elimination of double margins, the facilitation of relationship-specific non-contractible investments, and the assurance of an input supply.

In addition to these primary motives, a firm's decision of whether or not to vertically integrate may be influenced by the actions of its rivals. For instance, a downstream firm's incentive to integrate backward may be greater if a larger proportion of its rivals are vertically integrated. This would be the case if vertical integration has a foreclosure effect that raises the input price faced by the downstream firm. Thus, "bandwagon" behavior, where a firm vertically integrates in response to similar action by rivals, may be profitable under some circumstances.

Most of the empirical analysis on the determinants of vertical integration has focused on efficiency motives (Lafontaine and Slade, 2007). A common approach is to investigate the relationship between certain market characteristics – such as those associated with the importance of non-contractible relationship-specific investments – on the one hand, and the incidence or prevalence of vertical integration on the other. Numerous studies have found a significant relationship between non-contractible investment requirements and vertical integration.[1] This has provided support to the transaction cost and property rights theories of vertical integration represented by

[1] Recent examples include Woodruff (2002) and Ciliberto (2006). Whinston (2003) provides a useful review.

Williamson (1971), Klein et al. (1979), and Grossman and Hart (1986).

In principle, the vertical foreclosure motive can also be explored through similar methodology – for instance, by examining whether markets that are more susceptible to foreclosure are characterized by higher rates of vertical integration. However, market characteristics associated with vulnerability to foreclosure, such as the level of market concentration, also tend to be related to the degree of relationship specificity in investments. Thus, studies that find a positive relationship between market concentration and vertical integration attribute those findings to efficiency rather than foreclosure motives (e.g., Caves and Bradburd, 1988; Lieberman, 1991). [2]

To date, the significance of bandwagon effects as a cause of vertical integration has received little attention from empirical researchers. This is despite suspected cases of bandwagon behavior often being discussed and documented in business and legal circles. For example, industry executives in the cement and ready-made concrete industries, which experienced a vertical merger wave during the 1960s, justified their vertical integration decisions as an inevitable response to increasingly integrated rivals (Federal Trade Commission, 1966). A more recent example is the acquisition of Kinko's by Fedex in 2004. The shipping company's acquisition of the office services provider, which enabled the former to access small-business owners and other customers more directly, was seen by commentators as a response to rival shipper UPS's acquisition of Mail Boxes Etc., another office services provider (Deutsch, 2003).[3]

This chapter looks at the causes of vertical integration in the US generic pharmaceutical industry. This industry consists of a number of markets, each identified by a particular drug product. Each market starts off as a patent-protected monopoly served by an originator pharmaceutical company – also called an innovator or brand-name firm. New markets open up to competition by generic manufacturers at different points in time, following the expiration of patents and other exclusivities held by the originators. This competition has a significant impact on the market price of drugs. Berndt and Aitken (2010) find, in a sample of nine drug markets that went generic during 2006-2008, that the daily cost of drug treatment fell by 50.1 percent on average in the first two years after generic entry. The same study finds that since 2007, the average volume-based share of generic products has been higher than 90 percent in markets where they exist. By generating large cost savings for consumers and insurers, generic competition has successfully reversed an earlier trend – observed up to the early 2000s – where pharmaceutical expenditure growth outstripped growth in the quantity of drugs being prescribed (Berndt and Aitken, 2010).

The generic drug industry is a suitable setting for investigating the motives for vertical integration because each market exhibits a clear demarcation between the upstream and downstream segments, and each entrant decides whether or not to vertically integrate. Upstream plants produce active pharmaceutical ingredients (APIs), which are chemical compounds with therapeutic prop-

[2]Much of the recent empirical literature on foreclosure effects take the form of impact analysis. For example, Hastings and Gilbert (2005) and Suzuki (2008) measure the effect of vertical integration on intermediate good prices and product quality, respectively. The methodological focus of these studies is to find situations where the incidence or prevalence of vertical integration can be assumed to be exogenous. Aydemir and Buehler (2003) is a notable exception.

[3]The dearth of empirical research on bandwagon effects may be traced to the difficulty of collecting data; suspected cases of bandwagon behavior such as Fedex/Kinko's are few and far between in most industries.

erties, using raw materials such as basic and intermediate chemicals, solvents, and catalysts. The downstream segment manufactures finished formulations by combining APIs with inactive ingredients and processing them into dose forms such as tablets and injectables. There is a significant degree of vertical integration in generic drug markets and it has been rising over time. Since the late 1990s, markets opening up in later years have tended to exhibit a greater prevalence of vertical integration. Using a sample of 128 markets, I calculate the average proportion of vertically integrated entrants among all downstream entrants as 8.1 percent in markets that went generic during 1993-2000. The corresponding figure for markets that opened up during 2001-2005 is 24.2 percent.

Using firm-level data from generic drug markets, I estimate the determinants of a firm's decision to vertically integrate. The first finding is that a firm has a higher probability of vertically integrating, conditional on its decision to enter the downstream segment, if it has greater past entry experience in the upstream API segment. This suggests that a firm's upstream experience lowers its cost of vertical integration. In addition, a firm is more likely to vertically integrate when the average upstream experience level among its rivals is higher. This is equivalent to saying that a firm's vertical integration probability is decreasing in its rivals' cost of vertical integration. Employing a simple duopoly model, I show how such a finding would arise if the payoff function of an individual firm has the following characteristic: the firm gains more from vertical integration when more of its rivals are vertically integrated – which is equivalent to saying that firms' vertical integration decisions are strategic complements. Intuitively, when a firm is faced with rivals who have low vertical integration costs, it expects a higher degree of vertical integration in the equilibrium market structure. Given that the firm's gain from vertical integration is greater when more of its rivals are integrated, we should observe a higher probability of vertical integration by the firm itself. In sum, firms in the generic drug industry are responding to the expected prevalence of vertical integration among rivals by becoming vertically integrated themselves. This can be classified as a type of bandwagon effect. Put another way, firms in the generics industry have payoff functions that are conducive to bandwagon behavior.

A second set of findings pertains to the relationship specificity of investments as a determinant of vertical integration. I find that generic drug companies are more likely to be vertically integrated in markets where they try to enter with a "paragraph IV certification" – a certification that one or more patents held by the originator pharmaceutical firm are either invalid or not infringed. Generic entrants have an incentive to engage in such patent challenges, because the first one to enter with a paragraph IV certification may be awarded a 180-day exclusivity in the generic market. I employ a simple model to argue that in markets characterized by paragraph IV patent challenges, upstream investment into API development tends to be relationship-specific. This is because in such markets, the upstream product has a much higher value if it is used by the first-to-file paragraph IV entrant (who owns the 180-day generic market exclusivity) than when it is used by some other firm. Such relationship specificity does not exist in other generic drug markets. Therefore, it is likely that the higher relationship specificity of upstream investments in paragraph IV markets explains the higher incidence of vertical integration in such markets.

The remainder of the chapter is structured as follows. In Section 2.2, I describe the process of entry and vertical market structure formation in the generic drug industry. The section also examines how vertical integration patterns have evolved over time. Section 2.3 employs simple theoretical models to derive testable predictions. The first model shows that when a firm's payoff gain due to vertical integration is increasing in the vertical integration status of its rival, the firm's probability of vertical integration rises as its rival's cost of integration falls. The second model demonstrates that in a market where generic companies engage in a race to be the first-to-file, investment into API development is characterized by relationship specificity. It also demonstrates the advantage of being vertically integrated in such a market. In Section 2.4, I present the econometric specification used to analyze the determinants of vertical integration by individual firms. Section 2.5 describes the data for the US generics industry and Section 2.6 presents the empirical results. Section 2.7 concludes.

## 2.2 Entry and Vertical Market Structure in the Generic Pharmaceutical Industry

### 2.2.1 Marketing Exclusivity of New Drugs

A pharmaceutical product market is born when an originator company receives approval from the Food and Drug Administration (FDA) to market a new drug. The approval process involves the submission of a New Drug Application (NDA) by the originator, and the FDA's review of the NDA based on the criteria of safety and efficacy. Included under the definition of new drugs are formulations containing entirely novel active pharmaceutical ingredients (called new chemical entities), formulations containing new combinations of existing APIs, new dosage forms of existing APIs, and existing drugs for use in previously unapproved indications.

Most newly approved drugs are awarded a period of marketing exclusivity by the federal government. For example, a drug containing a new chemical entity is usually protected by a patent on the API as well as by a five-year period of data exclusivity. The term "data" in data exclusivity refers to the clinical trials information generated by the originator and submitted to the FDA as part of its NDA. The data are protected in the sense that the FDA is not authorized to use it for the purpose of reviewing marketing approval applications submitted by generic manufacturers. In fact, the FDA is not even allowed to accept applications from generic companies until one year before the expiration of the originator's data exclusivity period if, as is normally the case, those applications rely on the originator's clinical trials data. New drugs that do not contain new chemical entities are also subject to data exclusivity: new combinations, new formulations, and new uses are all eligible for three years of data protection (International Federation of Pharmaceutical Manufacturers and Associations, 2005).

In many cases, a new drug is protected by multiple patents. Each patent basically has a minimum term of twenty years from the date of filing so that patent protection usually outlasts the data

exclusivity period.[4] The one covering the API is often called a basic product patent. In addition, there are patents that protect new formulations (including new combinations of existing APIs) and new uses for existing drugs. Originators also employ additional patents relating to the API, such as those covering new processes of manufacture and those protecting new chemical forms of the same compound (e.g., novel salts). Such additional patents, sometimes called secondary patents, are especially valuable when a new drug is not protected by a basic product patent. This was the case for the antiviral drug zidovudine, whose basic product patent had already expired when it was developed as a pioneering treatment for HIV infection (Grabowski, 2004). Even in cases where a basic product patent exists, secondary patents are often used to extend the exclusivity of a new drug beyond the life of the basic patent (Mándi, 2003). This is done by filing the secondary patents during or after the drug development stage, when the life of the basic patent has already been eroded by several years (Hutchins, 2003).

From the viewpoint of originators, a limitation of secondary patents as an entry barrier is that, unlike data exclusivities and basic product patents, they tend to provide incomplete protection against generic entry. It is sometimes possible for generic companies to produce and sell a drug without infringing any of its secondary patents. For example, if a drug is protected only by a process patent, a generic firm can avoid infringement by employing an alternative process. Moreover, the patentability of innovations that underlie secondary patents is often open to question even after the patent is granted. For instance, combining an anti-hypertension compound and a cholesterol-lowering agent into the same pill creates significant benefits for some consumers, given that physicians often prescribe such combinations. However, it is a challenge to argue that the combination satisfies the non-obviousness requirement of patentability. Thus, the validity of Pfizer's patent on Caduet, a combination of amlodipine besylate and atorvastatin calcium, has been challenged by several generic firms (Harrison, 2008).

In this way, many secondary pharmaceutical patents belong to the category of what Lemley and Shapiro (2005) call "probabilistic patents". Lei and Wright (2009) shed light on the question of why such patents are allowed to exist in the first place. Their empirical analysis indicates that while patent examiners at the US Patent and Trademark Office generally have the ability to correctly judge the patentability of an application, the pro-applicant rules and procedures within the organization drive them to issue more patents than they should.

The proliferation of secondary patents creates a potential "patent minefield" where generic firms face the risk of being sued by the originator for infringing a patent that they did not even know existed. Such litigation risks are harmful not only for the generic firms but also for consumers, because they may lead to the abrupt removal of approved generic products from the market. Partly to prevent such situations, the FDA requires originator firms to provide information on the patents covering new drugs as part of their NDA filings. Typically, originators provide information on all relevant patents except for those that only claim manufacturing processes. Once an NDA is approved, a list of patents that are associated with the new drug is published in a FDA publica-

[4]For patents whose applications were filed before June 8, 1995, the patent term is seventeen years from the date of issue or twenty years from the date of first application, whichever expires later.

tion called “Approved Drug Products with Therapeutic Equivalence and Evaluations”, commonly known as the Orange Book.[5] The Orange Book is used by generic companies to learn about the existence and duration of originator patents in every drug market that they contemplate for entry.

### 2.2.2 Process of Generic Entry

#### Downstream Entry Through Abbreviated New Drug Applications

The entry process for generic pharmaceutical has greatly evolved over the last three decades. Prior to 1984, generic firms seeking marketing approval had to provide the FDA with the same type of information as originator firms, including data on clinical trials conducted on a large number of patients. As a result of the substantial entry costs that this entailed, entry by generic companies was limited: in 1984, roughly 150 drug markets were estimated to have been lacking generic entrants despite the expiration of patents (Federal Trade Commission, 2002).

The Drug Price Competition and Patent Restoration Act of 1984, also known as the Hatch-Waxman Amendments, drastically changed the process of generic entry. Most significantly, generic companies were exempted from submitting complete NDAs.[6] Instead, a generic entrant could file an Abbreviated New Drug Application (ANDA), which replaces full-scale clinical trial results with data on bioequivalence. Bioequivalence tests, which compare generic and originator drugs in the way that the active ingredient is absorbed into the bloodstream of healthy subjects, are much smaller in scale and far cheaper to conduct than conventional clinical trials. When the FDA reviews an ANDA for a generic product, its decision is based on the bioequivalence test results as well as the clinical trial results contained in the originator product’s NDA. The introduction of the ANDA system implied a huge reduction in product development costs, and generic entry surged after the mid-1980s; the volume-based share of generic drugs rose from 19 percent in 1984 to 51 percent in 2002, increasing further to 74 percent in 2009 (Grabowski, 2004; Berndt and Aitken, 2010).

ANDAs are prepared by downstream finished formulation manufacturers and submitted to the FDA some time before they plan to enter the generic market. In the case of a drug containing a new chemical entity, the earliest possible date for filing an ANDA is four years after the approval of the originator’s NDA (one year before the data exclusivity expires), but typical filing dates are later. If a generic firm plans to enter after all patents listed in the Orange Book have expired, it begins the ANDA filing process two to three years before the patent expiration date (Scott Morton, 1999). This reflects the expected time it takes the FDA to review an ANDA; the median approval time was 16.3 months in 2005, increasing in recent years to reach 26.7 months in 2009 (Buehler, 2006; Karst, 2010).[7]

[5] An electronic version of the Orange Book is accessible at `http://www.accessdata.fda.gov/scripts/cder/ob/default.cfm`.

[6] Another important aspect of the Hatch-Waxman Amendments is that it introduced patent term restorations of up to five years, in order to compensate for the delay in drug marketing that arises from the FDA’s regulatory process.

[7] The lengthening of generic approval times is due to a growing backlog of ANDAs. This backlog has been caused by a larger number of drugs going off patent and more firms entering each market (Buehler, 2006).

When unexpired patents are listed in the Orange Book at the time of ANDA filing, the generic firm must make a certification regarding each patent. The firm either indicates that it will wait until the patent expires to enter, or certifies that the patent is invalid or not infringed by its product. The first option is called a paragraph III certification and the latter is called a paragraph IV certification, named after corresponding passages in section 505(j)(2)(A) of the Federal Food, Drug, and Cosmetic Act. By filing an ANDA containing a paragraph IV certification, a generic firm preemptively counters any patent infringement claims that it expects from the originator. The FDA cannot give full approval to an ANDA until all patents listed in the Orange Book have expired or have been determined to be invalid or not infringed; a tentative approval, which does not permit the ANDA applicant to enter, can be issued in the mean time. The filing of an ANDA by a generic firm is not publicized by the FDA until the latter announces a tentative or full approval. Therefore, generic firms generally do not observe their rivals preparing and filing ANDAs in real time.

**Sourcing of Active Pharmaceutical Ingredients**

The preparation of an ANDA involves the development of the generic drug product by the applicant, who uses it to conduct bioequivalence tests.[8] A physical sample of the product is submitted to the FDA along with documents pertaining to bioequivalence and quality. An important part of generic product development is the sourcing of APIs. Here, the ANDA applicant faces a make-or-buy decision. If the firm has a plant equipped with specialized machinery such as chemical reactors, it can choose to produce its own API. If the ANDA applicant decides to buy its API from outside, it must find a supplier from among the many manufacturers located around the world. There is no centralized market for generic APIs, but international trade shows such as the Convention on Pharmaceutical Ingredients and Intermediates (CPhI) provide regular opportunities for buyers and suppliers to gather and transact. Once the API is obtained, the downstream firm develops the finished formulation and prepares documentation for the ANDA.

The ANDA documents, which are used by the FDA to evaluate the safety and efficacy of the generic product, must convey detailed information regarding the manufacture of the API to the agency. When the API is purchased from outside, the required information must be supplied by the upstream manufacturer. Basic information on the processes used for synthesizing the API is usually shared between the seller and buyer, but there remain trade secrets – such as the optimal conditions for chemical reaction – that the upstream firm may be unwilling to fully disclose to the downstream buyer. This is because the buyer might misuse the trade secrets by divulging them to other upstream firms who are willing to supply the API at a lower price.

To address such concerns among API manufacturers, and to maximize the quantity and quality of API-related information that reaches the FDA, the agency uses a system of Drug Master Files (DMFs). DMFs are dossiers, prepared by individual manufacturers, that contain information on manufacturing processes and product quality for APIs. By submitting the DMF directly to the FDA rather than to its downstream customer, the API manufacturer is able to convey all relevant

[8] Section 271(e)(1) of the Patent Act, also known as the Roche-Bolar provision, enables generic firms to develop their products during the originator's patent term without being sued for infringement.

information to the regulatory agency without risking the misuse of its trade secrets (Shaw, 2008).[9] Unlike ANDAs, the identities of submitted DMFs are published upon receipt by the FDA.[10]

If an ANDA applicant buys APIs from outside, it notifies the FDA about the source of the ingredient by referring to the serial number of a specific DMF. At the same time, the applicant contacts the DMF holder, who in turn informs the FDA that the ANDA applicant is authorized to refer to its DMF. In this way, the FDA reviewer knows where to find the API-related information for each ANDA. It is possible for the ANDA applicant to reference multiple DMFs at the time of filing, and for a single DMF to be referenced by multiple ANDAs. On the other hand, adding new DMF reference numbers after filing the ANDA is time-consuming. According to the Federal Trade Commission (FTC), it takes around eighteen months for an ANDA applicant to switch its API supplier by adding a new DMF reference.[11]

It would appear that a vertically integrated entrant has less of an incentive to use the DMF system than an unintegrated upstream firm. To the extent that the vertically integrated firm produces API exclusively for in-house use, concerns about the expropriation of trade secrets do not arise. In reality, however, many DMFs are filed by vertically integrated firms. One reason for this is that such firms often sell APIs to unintegrated downstream firms even if they are competing in the same market. For instance, Teva, a large Israeli generic drug company who is present in many US generic markets as a vertically integrated producer, sold 32 percent (in value terms) of its API output in 2008 to outside buyers (Teva Pharmaceutical Industries, 2009). Another reason is that generic companies often file separate ANDAs for multiple formulations containing the same API. By submitting a DMF to the FDA, an integrated firm can avoid the burden of including the same API information in multiple ANDAs. While one cannot rule out the possibility that vertically integrated firms sometimes refrain from submitting DMFs, the above discussion suggests that a DMF submission is a good indicator of upstream entry by both vertically integrated and unintegrated entrants.[12]

[9]The DMF system may have facilitated the vertical separation between the API and finished formulation manufacturing activities. The risk of expropriation of upstream trade secrets, had it not been addressed by the DMF system, may have motivated more firms to vertically integrate.

[10]The list of DMFs submitted to the FDA is available on the website of the FDA's Office of Generic Drugs at `http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/DrugMasterFilesDMFs/default.htm`.

[11]See Amended Complaint for Injunctive and Other Equitable Relief, *FTC v. Mylan Laboratories, Inc., et al.* (D.D.C., 1999), available at `http://www.ftc.gov/os/1999/02/mylanamencmp.htm`.

[12]There are two possible reasons why a vertically integrated firm may want to avoid filing a DMF, but neither of them appear to be substantial. First, filing a DMF might alert the originator firm to the entry plans of the integrated generic firm, causing the former to take defensive action. However, the generic firm can avoid giving such early warning by submitting the DMF immediately before filing its ANDA (the latter act is immediately observed by the originator if a patent challenge is involved, as described later). Second, by filing the DMF and exposing its intent to enter, the vertically integrated firm may reveal private information about the profitability of a market to other generic companies. Such information asymmetries are, however, unlikely in the generics industry where markets tend to be mature by definition. In fact, a vertically integrated firm may gain strategically by using a DMF submission to credibly indicate its intent to enter, possibly deterring the entry of some of its rivals. By contrast, the FDA's policy of keeping ANDA receipts confidential until approval implies that an unintegrated downstream firm can at best engage in cheap

A final note regarding DMFs addresses the possibility that a DMF submission does not necessarily imply entry into the API market. As Stafford (2006) suggests, some API manufacturers may file a DMF to attract the attention of potential buyers, but may not begin actual product development for the US market until buyer interest is confirmed. Such cases do appear to exist, but the practice is counterproductive for two reasons. First, a spurious DMF that is not backed by an actual product, while creating little real business for the firm, can be potentially damaging for an API manufacturer's reputation. Second, changing the content of an already-submitted DMF is time-consuming and requires notification to downstream customers (Food and Drug Administration, 1989). Thus, it seems safe to assume that a DMF submission by a relatively established API manufacturer indicates upstream market entry.[13]

**Stylized Description of Vertical Market Structure Formation**

In order to motivate the subsequent empirical analysis, I present a stylized description of the vertical market structure formation process in the generic industry. The process varies depending on whether or not a patent challenge is involved. I first consider the situation without patent challenges, and discuss the case involving patent challenges next.

When all generic entrants decide to wait until the expiration of originator patents (i.e., they make paragraph III certifications with respect to all unexpired patents), the vertical market structure of a given generic drug market is formed through a simultaneous entry game. Potential entrants simultaneously choose their actions from the following four alternatives: unintegrated downstream entry, unintegrated upstream entry, vertically integrated entry, and no entry. A firm's ANDA filing is not observed by the other players until the FDA announces its approval. This unobservability allows us to assume that firms make their downstream entry decisions simultaneously (Scott Morton, 1999). On the other hand, an entrant's submission of a DMF becomes observable when the FDA posts that information on its website. This creates the possibility that some firms choose their actions after observing the upstream entry decisions of other firms. However, since upstream manufacturers tend to submit DMFs later in the product development process, when they are already capable of producing the API on a commercial scale, it is reasonable to assume that upstream entry decisions are made simultaneously with downstream decisions.

Once the identities of the market entrants are fixed, we can envision a matching process where downstream manufacturing units are matched with upstream units. The matching process is not observed, because data from the FDA do not tell us which ANDAs refer to which DMFs.[14] After

talk – in the manner of Farrell (1987) – about its intention to enter a market.

[13]In a 2007 suit where a patent holder sought to prevent a generic API manufacturer from selling an infringing product, the plaintiff's attorney stated that "the act of filing a DMF indicates that the present intent of the DMF filer is to supply API in the United States". See Complaint for Declaratory Judgment, *Teva Pharmaceutical Industries, Ltd. v. Lupin Ltd.* (D.N.J., 2007), available at `http://patentdocs.typepad.com/patent_docs/files/teva_v_lupin_621.pdf`.

[14]In June 2005, I filed a Freedom of Information Act request for information on the linkages between specific ANDAs and DMFs. In July 2005, I received a reply from the FDA's Center for Drug Evaluation and Research stating that the requested information is proprietary and cannot be disclosed.

the matches are realized, firms invest in product development and document preparation. Upstream units develop their APIs and submit DMFs to the FDA, while downstream units develop finished formulations and file their ANDAs.[15] Downstream generic manufacturers market their products to consumers after the FDA approves their ANDAs and all patents and data exclusivities belonging to the originator expire. The payoffs of individual firms are realized when each downstream firm's revenue is split between itself and its upstream supplier, in the form of payment for APIs.

**Entry Process in the Presence of a Patent Challenge**

When entry into a generic drug market involves a paragraph IV patent challenge, the process of market structure formation can no longer be described as a simultaneous entry game. There are two reasons for this. First, there is no fixed date when generic firms begin to enter, due to the uncertain nature of patent litigation outcomes. Second, there exist regulatory rules that reward the first generic firm to initiate a successful patent challenge against the originator. This causes potential entrants to compete to become the first patent challenger.

The system of rewarding patent challenges was introduced in 1984 as part of the Hatch-Waxman Amendments. The rationale for providing such an incentive to generic firms is that the outcome of a successful patent challenge – the invalidation of a patent or a finding of non-infringement – is a public good (Miller, 2004). Suppose that one generic firm invests in research and spends time and money on litigation to invalidate an originator patent listed in the Orange Book. Suppose also that the patent is the only one protecting a particular drug market. Then, the act of invalidation benefits not only the generic firm who made the investment, but also others who seek to enter the market. Because such public goods tend to be undersupplied in a competitive market, Congress created a system to reward the first generic firm to invest in a patent challenge.

The reward is given out through a complex process that I summarize here. When a generic firm files an ANDA containing a paragraph IV certification to the FDA, it must directly notify the originator (the holder of the NDA for the original product), as well as the other holders of the patents being challenged, about its filing. The originator must then decide within 45 days whether or not to initiate a patent infringement suit. If the originator decides not to sue, then the FDA is allowed to approve the ANDA and the generic may enter the market. If the generic firm is the first to have filed a substantially complete ANDA containing a paragraph IV certification, it is awarded a 180-day exclusivity in the generic market. This means that the FDA is not allowed to approve any other ANDA until 180 days have passed since the first generic product's commercial launch.

If the originator decides to sue the generic entrant, then the FDA is stayed from giving final approval to the ANDA until 30 months have passed or until a court decides that the patent in question is invalid or not infringed, whichever comes sooner. The FDA may review the ANDA in the mean time, but it can only issue a tentative approval. Thus, the 30-month stay functions as an

[15]The existence of a time gap between entry decisions and actual investments (due to the inclusion of the matching stage) suggests that some firms may cancel their entry plans after finding out that the outcome of the entry and matching processes is not in their favor. Such reversals would create transactional risks for other firms, which in turn may affect the entry behavior of all potential entrants. To avoid this problem, I assume that entry decisions are irreversible.

automatic preliminary injunction against the paragraph IV ANDA applicant.

The main possible outcomes of the patent infringement suit between the originator and the paragraph IV applicant are the following: a victory for the generic entrant, a loss for the generic entrant, or a settlement between the two parties. If the generic applicant wins the patent infringement suit, its ANDA receives final approval from the FDA once the other patents listed in the Orange Book expire. If the generic firm is the first to have filed a substantially complete paragraph IV ANDA, it obtains the right to 180-day exclusivity. The exclusivity period starts when the first-to-file generic begins commercial marketing or when a court decides that the patent in question is invalid or not infringed, whichever is earlier.

If the generic firm loses the infringement suit for every challenged patent, then its ANDA is not approved until expiration of those patents or until the end of the 30-month stay. Even if the firm is the first-to-file paragraph IV applicant, it is not awarded the 180-day exclusivity, because the right to exclusivity disappears with the expiration of the challenged patents (Lietzan, 2004a). If the generic and originator firms decide to settle the patent infringement suit, the generic firm's ANDA is approved only after the 30-month stay. If the generic firm is the first-to-file paragraph IV applicant, it becomes eligible for 180-day exclusivity, which is triggered by the generic product's commercial launch.

The right to 180-day exclusivity is given only to the first-to-file paragraph IV applicant. If the first-to-file applicant loses in patent infringement litigation or otherwise forfeits its right to 180-day exclusivity, the right disappears; it is not rolled over to the next-in-line applicant (Korn et al., 2009). If multiple firms file ANDAs with paragraph IV certifications on the same day, and no prior ANDA has been filed, the right to generic exclusivity is shared between those firms.[16]

Although the Hatch-Waxman framework for rewarding patent challenges was introduced in 1984, it was not until the late 1990s that 180-day exclusivities began to be issued on a regular basis. Prior to 1998, the FDA's regulatory rules required a paragraph IV applicant to be sued by the originator, and to prevail in the ensuing infringement suit, in order to be eligible for generic exclusivity. This rule, called the "successful defense requirement", prevented most paragraph IV applicants from earning 180-day exclusivity because in many cases the originator did not sue and many patent disputes that were litigated ended in settlement. The Federal Trade Commission (2002) notes that between 1992 and 1998, not a single 180-day exclusivity was granted by the FDA. The system changed drastically following a pair of appellate court decisions: *Mova Pharmaceutical Corp. v. Shalala* (D.C. Cir., 1998) and *Granutec, Inc. v. Shalala* (4th Cir., 1998). These decisions struck down the FDA's successful defense requirement, and allowed paragraph IV applicants to be eligible for 180-day exclusivity even if they are not sued by the originator or if their suit ends in settlement (Lietzan, 2004b).

The regulatory change of 1998 had a dramatic impact. According to the Federal Trade Commission (2002), 180-day exclusivities were granted 31 times between 1998 and 2002. The generic

[16]Such "shared exclusivities" arise when multiple generic firms file on the first day that the FDA begins accepting ANDAs. For a drug containing a new chemical entity, that date is exactly four years after the approval of the originator's NDA.

exclusivity awarded to Barr Laboratories in 2000 for the antidepressant drug fluoxetine (Eli Lilly's Prozac) demonstrated the magnitude of profits at stake in the markets for so-called "blockbuster" drugs. Barr's stock price rose by two-thirds on the day of the appellate court decision invalidating the patent held by Eli Lilly. Barr proceeded to capture a 65 percent share of the market for fluoxetine within two months (Filson and Oweis, 2010).[17]

The large profits available from 180-day exclusivities have made generic firms more aggressive in their patent challenges. As Grabowski (2004) and Higgins and Graham (2006) note, the number of ANDAs containing paragraph IV certifications increased rapidly after the regulatory change: the average number of paragraph IV ANDA filings per year rose from thirteen during 1992-2000 to 94 in the 2001-2008 period. While this increase partly reflects the greater number of blockbuster drugs going generic in the latter period, observers agree that the regulatory change played a significant role (Grabowski, 2004; Filson and Oweis, 2010; Hemphill and Sampat, 2010). Table 2.1 presents the share of generic markets that were the subject of one or more paragraph IV ANDA filings in a sample of 128 markets that opened up during 1993-2005. As described more fully in Section 2.5, drug markets were selected for inclusion using the following criteria: (i) the drug product contains only one API; (ii) of the set of finished formulations containing the same API, the product is the first to experience generic entry; and (iii) there is at least one generic entrant in the market. The propensity of paragraph IV challenges suddenly jumps for markets that experienced first generic entry in 1999. This reflects expectations among generic firms that the FDA would give out more 180-day exclusivities following the 1998 court decisions. The share of generic markets with paragraph IV certifications remains high – at around one-half – in the subsequent years.[18]

Grabowski (2004) comments that the granting of more 180-day exclusivities has, in some cases, turned the generic entry process into a race to be first. Higgins and Graham (2006) note that as a result of more aggressive efforts by generic entrants, ANDA filings have come to take place earlier in a drug's lifecycle. Indeed, there have been many markets where multiple generic firms filed their paragraph IV ANDAs exactly four years after the approval of the originator's NDA – that is, on the earliest date allowed by the FDA (Grabowski, 2004). Also, Grabowski and Kyle (2007) show that drug markets with higher revenue tend to experience generic entry sooner, partly because they tend to be more heavily targeted for paragraph IV challenges. Interestingly, while ANDAs filings are being made increasingly early, Grabowski and Kyle (2007) find no evidence that generic product launches are occurring earlier in the drug's lifecycle in markets that opened up more recently. This may be because the Hatch-Waxman system has had an unintended side effect. As reported by the Federal Trade Commission (2002) and Bulow (2004), the system has been used by some originators, somewhat paradoxically, to delay generic entry through the use of so-called "pay-to-delay" settlements.[19]

[17] According to Garnett (2000), fluoxetine had global revenues of more than 2.5 billion dollars in 1999.

[18] Using a larger dataset of generic drug approvals, Hemphill and Sampat (2010) shows that new drugs approved during the 1990s were more likely to be the subject of paragraph IV patent challenges than those approved earlier. These drugs are likely to have experienced generic entry after the 1998 court decisions.

[19] To see how such a settlement might be employed, suppose that an originator and a first-to-file paragraph IV ANDA applicant begin a patent suit and approval of the ANDA is stayed by 30 months. By settling or prolonging the

Table 2.1: Incidence of Paragraph IV Certification

| Year | Number of markets | Share of markets with Paragraph IV Certification (%) |
|---|---|---|
| 1993 | 8 | 12.5 |
| 1994 | 5 | 0.0 |
| 1995 | 10 | 20.0 |
| 1996 | 4 | 0.0 |
| 1997 | 9 | 11.1 |
| 1998 | 7 | 14.3 |
| 1999 | 6 | 66.7 |
| 2000 | 9 | 22.2 |
| 2001 | 12 | 50.0 |
| 2002 | 17 | 52.9 |
| 2003 | 14 | 42.9 |
| 2004 | 16 | 56.3 |
| 2005 | 11 | 18.2 |

Notes:
The second column shows the number of markets in the dataset to experience first generic entry in each year. The selection of markets is explained in Section 2.5.

The third column shows the percentage of markets where one or more ANDAs containing a paragraph IV certification was filed.

trial, the two parties can prevent the FDA from approving the first-to-file applicant's ANDA for the duration of the stay. Under the regulations that were in place until 2003, the originator and the generic challenger could delay the approval of subsequent ANDAs even after the expiration of the stay and the approval of the latter's ANDA. This was because the first-to-file applicant's right to 180-day exclusivity was not triggered until the applicant began commercial marketing as long as a court decision could be avoided. Thus, originators were able to delay generic competition indefinitely by convincing first-to-file applicants to hold off entry – often with the help of settlements involving payments to the generic side. While court decisions have been permissive of such pay-to-delay settlements (see, e.g.,*Schering-Plough v. FTC*, 11th Cir., 2005), their legality has been challenged by the FTC (Federal Trade Commission, 2002). Based on the FTC's recommendations, the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003

Given that the existence of a patent challenge turns the generic entry process into a race to be first, econometric analysis of generic firm behavior would ideally be based on a model that takes the timing of entry into account. Unfortunately, the data that I use do not contain accurate information on the timing of entry by each generic firm.[20] Also, I do not observe whether or not each ANDA filing contains a paragraph IV certification because this information is not disclosed by the FDA. On the other hand, the FDA publishes a list of drug markets that were the subject of one or more ANDAs containing a paragraph IV certification. Therefore, it is possible to distinguish between paragraph IV markets and non-paragraph IV markets, and to see if firm behavior differs across the two groups.

Our interest in this study is in seeing if paragraph IV patent challenges are associated with generic firms' vertical integration decisions. How might such an association arise? As I argue in Section 2.3, when generic entry involves a race to be first, investments made by upstream API manufacturers tend to become specific to a particular downstream buyer. If contracts between unintegrated upstream suppliers and downstream buyers are incomplete and payoffs are determined through *ex post* bargaining, this increase in relationship specificity could enhance the role of vertical integration as a way to facilitate investments. In the empirical analysis, I examine whether the occurrence of paragraph IV certification at the market level is associated with higher incidence of vertical integration at the firm level.

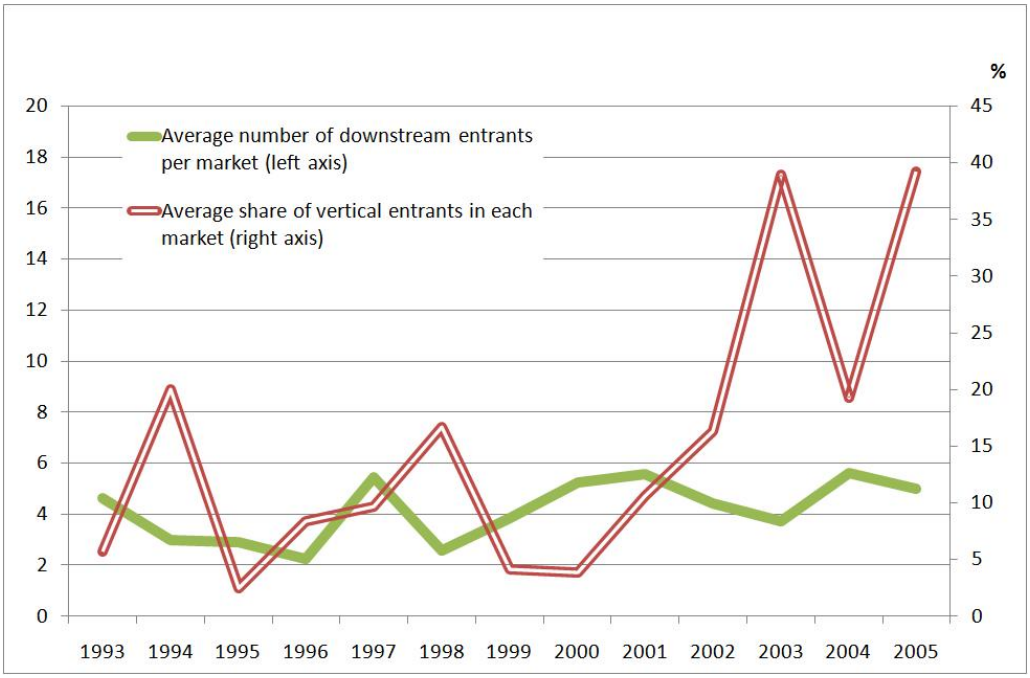
### 2.2.3 Trend in Vertical Integration

Before turning to the formal analysis, let us examine the pattern of vertical integration in the generics industry. Figure 2.1 shows how the prevalence of vertical integration at the market level has changed over time. It is based on the sample of 128 markets that opened up between 1993 and 2005. It can be seen that the average number of downstream entrants (including vertically integrated ones) per market has remained stable at around five. On the other hand, the share of those downstream entrants that are vertically integrated has increased over time. For markets that opened up in the 1993-2000 period, the average share of vertically integrated entrants, as a percentage of the number of downstream entrants, was 8.1 percent. In 2001-2005, the figure rose to 24.1 percent and the difference between the sub-periods is highly significant (the p-value is 0.001).

The incidence of vertical integration has similarly risen over time. In each of the years from 1993 to 2000, 24.0 percent of the sample markets opening up each year, on average, had one or more vertically integrated entrants. For the years 2001-2005, the average share of markets having

introduced several amendments to the Federal Food, Drug, and Cosmetic Act to limit the scope for collusive delays. Under the MMA provisions, the first-to-file ANDA applicant forfeits its right to 180-day exclusivity if the right is not exercised within 75 days of a settlement in the patent infringement suit or a court decision of invalidity/non-infringement (Korn et al., 2009). In addition, whereas originators were previously able to use multiple 30-month stays to delay the approval of the first-to-file paragraph IV ANDA, the MMA allows only one stay per drug product.

[20]This is because the FDA, whose data I use to measure entry, publicizes the approval dates of ANDAs but not their filing dates.




Notes:

(a) The selection of markets is explained in Section 2.5.

(b) The number of markets opening up each year is presented in Table 2.1.

(c) For each year, the average number of downstream entrants (including vertically integrated entrants) and the average share of vertically integrated entrants in terms of entrant count are calculated for the sample markets that opened up in that year.

Figure 2.1: Market-Level Share of Vertically Integrated Entrants

any vertically integrated entry was 64.6 percent (the p-value for the inter-period difference is less than 0.001).

An interesting fact about the US generic pharmaceutical industry is that it started off as being vertically separated. When the industry began its growth in the 1980s, finished formulation manufacturers procured most of their API requirements from outside suppliers located in Italy, Israel, and other foreign countries. This was mainly due to differences in patent protection across countries: while strong patent protection in the US (and the lack of Roche-Bolar-type exemptions until 1984) made it difficult for domestic companies to develop APIs before the expiration of originator patents, the weak patent regimes in Italy and other countries at the time allowed firms located there to develop generic APIs early (Bryant, 2004).

In addition to these historical origins, the nature of the generics business also made vertical

separation a natural outcome. Different downstream manufacturers of generic drugs produce near-identical products, because, by definition, they are all bioequivalent to the original product. Therefore, the APIs manufactured by different upstream firms are also expected to be homogeneous. This implies that in general, investments into API development by an upstream manufacturer are not specific to a particular downstream user. In other words, the investment facilitation effects of vertical integration are unlikely to be important in this industry under normal circumstances. This is analogous to Hart and Tirole's (1990) observation that the efficiency benefits of vertical integration were unlikely to have been strong in the cement and ready-mixed concrete industries during the 1960s when the vertical merger wave took place. Nevertheless, as Figure 2.1 demonstrates, vertical integration has become more prevalent over time in the generics industry. Several possible reasons for this can be found from industry reports.

One is that early development and procurement of APIs has become more important to the profitability of downstream manufacturers in recent years, particularly in markets characterized by paragraph IV patent challenges. For example, the annual report of Teva, the industry's largest firm, describes the motive for vertical integration as follows: "to provide us with early access to high quality active pharmaceutical ingredients and improve our profitability, in addition to further enhancing our R&D capabilities." (Teva Pharmaceutical Industries, 2008, p.15). Karwal (2006) mentions that "having access to a secure source of API can make a significant difference, particularly relating to difficult-to-develop API, when pursuing a potential Paragraph IV opportunity, and to secure sufficient quantities for development" (p.274). Similarly, Burck (2010) notes that "Access to API and control of the development and manufacturing process to support patent challenges has often been cited as a reason for backward integration" (p.34). These comments suggest that vertical integration allows downstream manufacturers to obtain APIs sooner than they otherwise would, and that this aids them in attaining first-to-file status in paragraph IV markets. This would partly explain why the increased prevalence in vertical integration appears to have followed closely behind the increase in paragraph IV patent challenges.

A second possible cause of increased vertical integration pertains to bandwagon effects. A former purchasing executive at Sandoz, one of the largest firms, mentions that firms vertically integrate to "avoid sourcing API from a competitor" (Stafford, 2006, p.302). Karwal (2006) points out that "Many key API suppliers, especially from India, China and Eastern Europe, are moving up the value chain and decreasing their supply activities, becoming direct competitors in finished form generics" (p.274).[21] He suggests that this is one of the factors behind increased backward

[21]During the 1990s, traditional API suppliers from Italy and other south European countries lost market share to new entrants from India and Eastern Europe. A major reason for this shift was that stricter patent protection in Western Europe – most notably the term extensions given to pharmaceutical patents through the introduction of Supplementary Protection Certificates in 1991 – made it more difficult for firms located there to develop their generic APIs early (Bryant, 2004; Stafford, 2006). Meanwhile, Indian pharmaceutical firms – who honed their product development skills under a weak patent regime that lasted from 1972 to 2005 and who became more open to the outside world under the economic liberalization policies of the early 1990s – focused on the US generics market as a target for their exports. As Lanjouw (1998) documents, Indian drug companies initially entered the US and other Western markets as API suppliers. By the mid-2000s, several of them, including Ranbaxy and Dr. Reddy's Laboratories, had also become major players in the downstream segment.

integration by established downstream manufacturers.

In the mid-2000s, traditionally unintegrated US firms in the downstream segment began acquiring API manufacturing assets. Examples include the acquisition of Indian API manufacturers by Mylan and Watson, both large US finished formulation companies.[22] It is important that these actions, by two of the main players of the industry, took place *after* vertically integrated entry became common. It is unlikely that Mylan and Watson were slower than their rivals at noticing the efficiency effects of vertical integration, given their long histories and large scale of activities.[23] More plausibly, their decisions were made in response to the expectation that generic drug markets were going to become increasingly vertically integrated.

The next section discusses how we can test the two leading explanations for the increase in vertical integration within the generic pharmaceutical industry: (i) the existence of bandwagon effects, and (ii) the importance of relationship-specific investments to support patent challenges.

## 2.3 Testing the Motives for Vertical Integration

### 2.3.1 Bandwagon Effects

#### Bandwagon Behavior and Strategic Complementarity

In the existing theoretical literature on vertical integration, bandwagon behavior is deemed to occur when a firm integrates in response to vertical integration by rivals (e.g., Hart and Tirole, 1990). In generic drug markets, firms make their entry and vertical integration decisions more or less simultaneously so that we do not observe firms choosing their vertical structures in response to the actions of their rivals.[24] Nevertheless, bandwagon effects can still exist in the sense that firms may become vertically integrated in response to the expected prevalence of vertical integration among rivals.

Such a possibility can be examined by seeing if the change in a firm's payoff from becoming vertically integrated is increasing (becoming either more positive or less negative) in the incidence or prevalence of vertical integration among rivals. In other words, we can check whether firms' payoff functions exhibit strategic complementarity in vertical integration decisions. As Buehler and Schmutzler (2005) point out, vertical integration decisions are shown to be strategic substitutes rather than complements in most theoretical models. However, there are a few important studies such as Ordover et al. (1990), Hart and Tirole (1990, p.227), and McLaren (2000) that demonstrate the possibility of strategic complementarity.[25] Anecdotal evidence also suggests the existence of

[22] Mylan acquired a majority stake in a large Indian API manufacturer called Matrix in September 2006 (Roumeliotis, 2006). In the same month, Watson acquired a smaller firm called Sekhsaria (Barnes, 2006).

[23] Mylan and Watson were founded in 1961 and 1984, respectively. As of 2006, both firms were among the top six firms in the global generic pharmaceutical industry in revenue terms (Stafford, 2006).

[24] In markets characterized by paragraph IV patent challenges, firms' decisions are not necessarily made simultaneously. Even in such markets, however, firms' actions tend to be unobserved until each firm makes its own decision so that the simultaneity assumption is justified.

[25] Algebraic analysis of the Ordover et al. (1990) model shows that integration decision are strategic complements.

strategic complementarity in certain industries. For instance, one US cement company's annual report for 1963 mentioned that while it was not inclined to acquire assets in the ready-made concrete industry, the wave of vertical integration among its rivals was forcing the firm to follow suit.[26]

I now show, using a simple duopoly model, that when firms' payoff functions are characterized by strategic complementarity in vertical integration decisions, the following testable prediction arises: a firm's probability of vertical integration decreases with its rival's cost of vertical integration. When vertical integration decisions are strategic substitutes, the opposite result holds: the firm's vertical integration probability increases with the rival's cost of vertical integration. These results allow us to design a simple econometric test of strategic complementarity.

**Duopoly Model of Equilibrium Vertical Integration**

Consider a market consisting of an upstream and a downstream segment. Assume that there are two potential downstream entrants indexed by 1 and 2. The firms simultaneously choose between unintegrated downstream entry ($D$), vertically integrated entry ($V$), and no entry.[27] When both firms 1 and 2 decide to enter as unintegrated downstream producers, I assume that two unintegrated suppliers enter the upstream segment. When one potential downstream entrant chooses unintegrated downstream entry while the other chooses vertically integrated entry, it is assumed that a single unintegrated upstream supplier also enters.

Each firm's payoff can be expressed as a function of its own action and the action of its rival. I assume for simplicity that the cost of unintegrated downstream entry is zero. On the other hand, the cost of vertically integrated entry is $K_i > 0, \;\; i = 1,2$. This includes the cost of developing the upstream product as well as any overhead costs that arise from holding upstream assets. Firm $i$'s payoff, net of entry cost, is $\pi(a_i, a_{i'}) - \mathbf{1}(a_i = V)K_i$, where $i' = 3 - i$ and $\mathbf{1}(\cdot)$ is the indicator function.

In the following, I employ the shorthand $\pi_{a_i,a_{i'}}$ to represent the post-entry payoff function $\pi(a_i, a_{i'})$. I assume that post-entry payoffs are greater than zero under any market structure, so that both of the potential downstream entrants always enter in one way or another. Thus, all realized market structures are characterized by two upstream units and two downstream units. The following assumptions are made about the post-entry payoff function:

---

Nevertheless, bandwagon behavior may not arise in their model because of an *ad hoc* ordering of integration decisions and strategic pricing by the first mover. Specifically, the first firm to vertically integrate sets the intermediate good price low enough so that its rival will not find it profitable to integrate.

[26] Annual Report of Alpha Portland Cement Company for 1963 as quoted in Federal Trade Commission (1966).

[27] The assumption that "unintegrated upstream entry" is not in the firms' choice set can be justified by the existence of independent upstream suppliers with lower costs.

$$\pi_{VD} > \pi_{DD}, \tag{2.1}$$

$$\pi_{VV} > \pi_{DV}, \tag{2.2}$$

$$\pi_{VD} > \pi_{VV}, \tag{2.3}$$

$$\pi_{DD} > \pi_{DV}. \tag{2.4}$$

Inequalities (2.1) and (2.2) say that a firm's post-entry profit, conditional on its rival's action, is higher if it is vertically integrated, whether the other firm chooses unintegrated entry or vertical integration.[28] Inequalities (2.3) and (2.4) say that a firm's post-entry payoff is decreasing in the other firm's vertical integration choice. Both sets of assumptions can be justified by the existence of efficiency effects due to vertical integration, such as the elimination of double marginalization. I also assume the following:

$$\pi_{DD} > \pi_{VV} - K_i, \quad i = 1,2, \tag{2.5}$$

which says that firms prefer to be in a market where both entrants are unintegrated than in one where both are vertically integrated.[29]

Two separate cases are considered with regard to the magnitude of payoff differentials. In the first case, $\pi_{VV} - \pi_{DV} > \pi_{VD} - \pi_{DD}$, so that vertical integration decisions are strategic complements. In the second case, $\pi_{VV} - \pi_{DV} < \pi_{VD} - \pi_{DD}$, implying that vertical integration actions are strategic substitutes.

The payoff matrix in Table 2.2 can be used to find the Nash equilibrium market structures. To see how one firm's equilibrium behavior is affected by the other firm's cost of vertical integration, let us assume that $K_1$ is fixed at some value $\bar{K_1}$ that falls between $\pi_{VD} - \pi_{DD}$ and $\pi_{VV} - \pi_{DV}$ and see how the equilibrium changes as $K_2$ varies.

Table 2.3 presents the results when the firms' vertical integration decisions are strategic complements. When the value of $K_2$ is at or below $\pi_{VD} - \pi_{DD}$ so that vertically integrated entry is a dominant strategy for firm 2, firm 1 also chooses vertically integrated entry in equilibrium.[30] On the other hand, when $K_2$ is greater than or equal to $\pi_{VV} - \pi_{DV}$ so that unintegrated downstream entry is firm 2's dominant strategy, firm 1 likewise chooses unintegrated downstream entry. For intermediate values of $K_2$, there are three possible Nash equilibria: the two pure strategy equilibria

[28] Note that the profit from vertically integrated entry is not necessarily higher than that from unintegrated entry once the cost of vertical integration, $K_i$, is subtracted out.

[29] This suggests that the equilibrium of the vertical integration game might be characterized as a Prisoner's Dilemma, a common result found in Ordover et al. (1990), Hart and Tirole (1990) and other representative models.

[30] Given that $\bar{K_1} \in [\pi_{VD} - \pi_{DD}, \pi_{VV} - \pi_{DV}]$, firm 1's optimal action is $D$ when firm 2 chooses $D$, and $V$ when firm 2 chooses $V$.

Table 2.2: Payoff Matrix of Vertical Entry Game

| | | Firm 2's action | |
|---|---|---|---|
| | | $D$ | $V$ |
| Firm 1's action | $D$ | $\pi_{DD}, \quad \pi_{DD}$ | $\pi_{DV}, \quad \pi_{VD} - K_2$ |
| | $V$ | $\pi_{VD} - K_1, \quad \pi_{DV}$ | $\pi_{VV} - K_1, \quad \pi_{VV} - K_2$ |

Notes:

(a) $D$ denotes unintegrated downstream entry and $V$ denotes vertically integrated entry.

(b) In each cell, the first element is firm 1's payoff and the second element is firm 2's payoff.

(c) The first subscript of $\pi$ represents the firm's own action; the second subscript is its rival's action.

$(a_1^*, a_2^*) = (D,D)$ and $(V,V)$, and one mixed strategy equilibrium.[31] By solving for the vertical integration probabilities that make both firms indifferent between vertically integrated entry and unintegrated downstream entry, the following mixed strategy equilibrium is derived:

$$\begin{aligned} &( \mathit{Prob}(a_1 = V),\ \mathit{Prob}(a_2 = V) ) \\ &= \left( \frac{K_2 - (\pi_{VD} - \pi_{DD})}{(\pi_{VV} - \pi_{DV}) - (\pi_{VD} - \pi_{DD})},\ \frac{K_1 - (\pi_{VD} - \pi_{DD})}{(\pi_{VV} - \pi_{DV}) - (\pi_{VD} - \pi_{DD})} \right). \end{aligned}$$

If we only look at the range of $K_2$ where the equilibrium is unique, firm 1's vertical integration probability is decreasing in firm 2's cost of vertical integration. As for the intermediate range characterized by multiple equilibria, we cannot say how firm 1's vertical integration probability changes with $K_2$.[32] We would like to say more about the relationship between the two variables – preferably, one-to-one mappings which would help us to derive testable predictions. This can be

[31] When $K_2 \in [\pi_{VD} - \pi_{DD}, \pi_{VV} - \pi_{DV}]$, each firm prefers to match the other firm's action. This gives rise to the two pure strategy equilibria.

[32] Firm 1's vertical integration probability in the mixed strategy equilibrium is increasing in $K_2$, but one cannot conclude from this that firm 1 is more likely to be vertically integrated when $K_2$ is high. The positive relationship between the two variables is an artifact of the indifference condition that characterizes the mixed strategy equilibrium. When $K_2$ is low and vertically integrated entry is relatively more attractive for firm 2, firm 1's vertical integration probability must be low enough in the mixed strategy equilibrium so that firm 2 stays indifferent between vertical

Table 2.3: Equilibrium Vertical Integration Probabilities Under Strategic Complementarity

| Range of $K_2$ | Firm 1's equilibrium vertical integration probabilities |
|---|---|
| $[\,0,\ \pi_{VD}-\pi_{DD}\,]$ | 1 |
| $(\,\pi_{VD}-\pi_{DD},\ \pi_{VV}-\pi_{DV}\,)$ | $\left\{0,\ \frac{K_2-[\pi_{VD}-\pi_{DD}]}{[\pi_{VV}-\pi_{DV}]-[\pi_{VD}-\pi_{DD}]},\ 1\right\}$ |
| $[\,\pi_{VV}-\pi_{DV},\ \infty\,)$ | 0 |

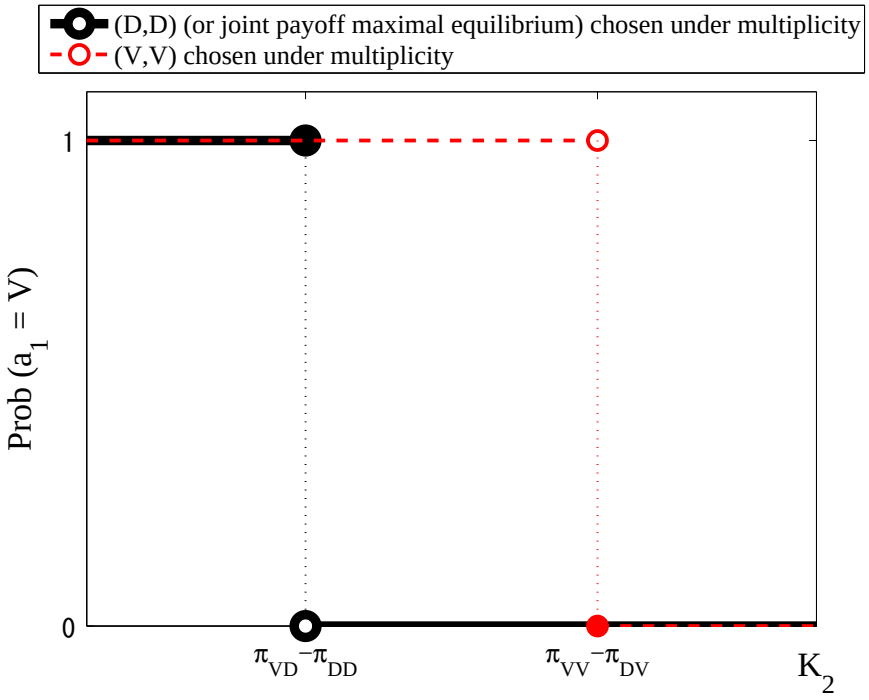
Notes:
Firm 1's vertical integration cost is fixed at $\bar{K}_1 \in (\pi_{VD}-\pi_{DD}, \pi_{VV}-\pi_{DV})$.
When $K_2 \in (\pi_{VD}-\pi_{DD}, \pi_{VV}-\pi_{DV})$, there is one mixed strategy equilibrium and two pure strategy equilibria: $(a_1^*, a_2^*) = (D,D)$ and $(a_1^*, a_2^*) = (V,V)$.

done by specifying different equilibrium selection rules. One simple rule is to let a particular pure strategy equilibrium be chosen for all values of $K_2$ in the intermediate range. This rule yields two possibilities. The first is that $(D,D)$ is always chosen for $K_2 \in (\pi_{VV}-\pi_{DV}, \pi_{VD}-\pi_{DD})$. The other possibility is that $(V,V)$ is always chosen in the intermediate range. Figure 2.2 shows how firm 1's vertical integration probability can be presented as decreasing step functions of $K_2$ under the two cases.

An alternative rule – one that is often employed in the empirical literature on entry games (e.g., Berry, 1992) – is to select the equilibrium (possibly one in mixed strategies) that yields the highest joint payoffs. Using inequalities (2.3), (2.4), and (2.5), it can be shown that under this rule, the pure strategy equilibrium $(a_1^*, a_2^*) = (D,D)$ is chosen when $K_2 \in (\pi_{VD}-\pi_{DD}, \pi_{VV}-\pi_{DV})$.[33] In other words, firm 1's vertical integration probability stays at zero when $K_2$ is in the intermediate

integration and unintegrated downstream entry. As $K_2$ rises, higher vertical integration probabilities for firm 1 are needed to maintain the mixed strategy equilibrium.

[33]Joint payoffs are greater under $(D,D)$ than under $(V,V)$ by (2.5). In the mixed strategy equilibrium, firm $i$'s payoff, given that firm $j$'s vertical integration probability is $q_j$, is $\pi_{VD} - q_j(\pi_{VD}-\pi_{VV}) - K_i = \pi_{DD} - q_j(\pi_{DD}-\pi_{DV})$. The equality indicates firm $i$'s indifference between vertically integrated and unintegrated downstream entry. The left-hand side is greater than $\pi_{VV}-K_i$ by (2.3), and the right-hand side is less than $\pi_{DD}$ by (2.4). Therefore, the joint payoffs under the mixed strategy equilibrium are between that under $(V,V)$ and that under $(D,D)$.




Notes:

(a) The horizontal axis represents firm 2's cost of vertical integration.

(b) The graphs represent firm 1's vertical integration probabilities under different equilibrium selection rules.

Figure 2.2: Firm 1's Vertical Integration Probability Under Strategic Complementarity

range.

Figure 2.2 demonstrates that, under each of the equilibrium selection rules considered, firm 1's vertical integration probability is a decreasing function of $K_2$ when the firms' vertical integration decisions are strategic complements. The fact that this result holds under different equilibrium selection rules suggests its generality. An intuitive interpretation is that when firm 2's vertical integration cost rises, firm 1 expects less vertically integrated entry by its rival. Under strategic complementarity, this expectation is translated into a lower probability of vertical integration by firm 1 itself.

Table 2.4 shows how firm 1's vertical integration probability changes with $K_2$ when the firms' integration decisions are strategic substitutes. In this case, the unique pure strategy equilibrium is $(a_1^*, a_2^*) = (D,V)$ for very low values of $K_2$ and $(V,D)$ for very high values of $K_2$.[34] The intermedi-

[34]Given that $\bar{K}_1 \in [\pi_{VV} - \pi_{DV}, \pi_{VD} - \pi_{DD}]$, firm 1's optimal action is $V$ when firm 2 chooses $D$, and $D$ when firm 2

Table 2.4: Equilibrium Vertical Integration Probabilities Under Strategic Substitutability

| Range of $K_2$ | Firm 1's equilibrium vertical integration probabilities |
|---|---|
| $[\,0,\ \pi_{VV}-\pi_{DV}\,]$ | 0 |
| $(\,\pi_{VV}-\pi_{DV},\ \pi_{VD}-\pi_{DD}\,)$ | $\left\{0,\ \frac{[\pi_{VD}-\pi_{DD}]-K_2}{[\pi_{VD}-\pi_{DD}]-[\pi_{VV}-\pi_{DV}]},\ 1\right\}$ |
| $[\,\pi_{VD}-\pi_{DD},\ \infty\,)$ | 1 |

Notes:
Firm 1's vertical integration cost is fixed at $\bar{K}_1 \in (\pi_{VV}-\pi_{DV}, \pi_{VD}-\pi_{DD})$.
When $K_2 \in (\pi_{VV}-\pi_{DV}, \pi_{VD}-\pi_{DD})$, there is one mixed strategy equilibrium and two pure strategy equilibria: $(a_1^*, a_2^*) = (D,V)$ and $(a_1^*, a_2^*) = (V,D)$.

ate values of $K_2 \in (\pi_{VV}-\pi_{DV},\ \pi_{VD}-\pi_{DD})$ are characterized by the two asymmetric pure strategy equilibria, $(D,V)$ and $(V,D)$, and the mixed strategy equilibrium

$$\begin{aligned}&(\,Prob(a_1=V),\ Prob(a_2=V)\,)\\ &=\left(\frac{(\pi_{VD}-\pi_{DD})-K_2}{(\pi_{VD}-\pi_{DD})-(\pi_{VV}-\pi_{DV})},\ \frac{(\pi_{VD}-\pi_{DD})-K_1}{(\pi_{VD}-\pi_{DD})-(\pi_{VV}-\pi_{DV})}\right).\end{aligned} \tag{2.6}$$

As before, let us consider different equilibrium selection rules for the intermediate range. If $(D,V)$ is always chosen for $K_2 \in (\pi_{VV}-\pi_{DV},\ \pi_{VD}-\pi_{DD})$, firm 1's vertical integration probability jumps from zero to one at $K_2 = \pi_{VD}-\pi_{DD}$, as seen in Figure 2.3. The jump occurs at $K_2 = \pi_{VV}-\pi_{DV}$ if $(V,D)$ is always chosen instead. If we assume that the pure strategy equilibrium with the highest joint payoffs is chosen, then the function exhibits a jump from zero to one at $\bar{K}_1$. Thus, as long as we restrict attention to pure strategy equilibria, as do most of the existing empirical studies on entry games (e.g. Berry, 1992; Mazzeo, 2002b; Ciliberto and Tamer, 2009), the function that maps from $K_2$ to firm 1's vertical integration probability is an increasing one when strategic substitutability holds.

The same function becomes more complicated if we allow for mixed strategy equilibria and apply the joint payoff-maximality rule. Let us simplify the analysis by setting firm 1's vertical integration cost at $\bar{\bar{K}}_1 = \frac{1}{2}(\pi_{VV}-\pi_{DV}+\pi_{VD}-\pi_{DD})$. First, consider the case where firm 2 has a

chooses $V$.

vertical integration cost that is less than or equal to firm 1's so that $K_2 \in (\pi_{VV} - \pi_{DV},\ \bar{\bar{K}}_1]$. The joint payoff maximal outcome in this case is the pure strategy equilibrium $(a_1^*, a_2^*) = (D,V)$. The proof involves taking the difference between the joint payoffs under the mixed strategy equilibrium and that under $(D,V)$, which is the pure strategy equilibrium with the highest joint payoffs. [35]

Next, consider the case of $K_2 \in (\bar{\bar{K}}_1,\ \pi_{VD} - \pi_{DD})$. It can be shown that the pure strategy equilibrium $(V,D)$ maximizes joint payoffs for $K_2 \in (\bar{\bar{K}}_1,\ \bar{\kappa}]$, where

$$\bar{\kappa} = \pi_{VV} - \pi_{DV} + \frac{(\pi_{VD} - \pi_{VV})[(\pi_{VD} - \pi_{DD}) - (\pi_{VV} - \pi_{DV})]}{2(\pi_{DD} - \pi_{DV})} > \bar{\bar{K}}_1.$$

The inequality follows from the strategic substitutability condition, $\pi_{VV} - \pi_{DV} < \pi_{VD} - \pi_{DD}$. For $K_2 \in (\bar{\kappa},\ \pi_{VD} - \pi_{DD})$, the mixed strategy equilibrium (2.6) maximizes joint profits. [36] As Figure 2.4 shows, firm 1's vertical integration probability is a non-monotonic function of $K_2$ in this case.

The monotonicity of firm 1's vertical integration probability with respect to firm 2's vertical integration cost under strategic substitutability depends on the choice of equilibrium selection rule. Nevertheless, there are grounds to expect the relationship to be increasing in practice. First, it is not very likely that in real world industries, firms actively switch between pure strategy entry equilibria and mixed strategy equilibria based on the criterion of joint profit maximality, as in Figure 2.4. Second, the range of $K_2$ in Figure 2.4 where firm 1's vertical integration probability is a decreasing function is narrow. Thus, we can state with some confidence that firm 1's equilibrium vertical integration probability is likely to be an increasing function of $K_2$ when vertical integration decisions

[35] The joint payoffs under the mixed strategy equilibrium can be written as

$$\Pi_{MS} = (1-q_2)\pi_{DD} + q_2\pi_{DV} + (1-q_1)\pi_{VD} + q_1\pi_{VV} - K_2,$$

where $q_i$ stands for firm $i$'s vertical integration probability. Taking the difference with $\Pi_{PS} = \pi_{DV} + \pi_{VD} - K_2$, the joint payoffs under $(D,V)$, and collecting terms gives

$$\begin{aligned}\Pi_{MS} - \Pi_{PS} &= (1-q_2)(\pi_{DD} - \pi_{DV}) - q_1(\pi_{VD} - \pi_{VV}) \\ &= \frac{1}{2}(\pi_{DD} - \pi_{DV}) - \frac{\pi_{VD} - \pi_{DD} - K_2}{(\pi_{VD} - \pi_{DD}) - (\pi_{VV} - \pi_{DV})}(\pi_{VD} - \pi_{VV}) < 0.\end{aligned}$$

The second equality is obtained by plugging in the expressions for $q_1$ and $q_2$ and rearranging. The last inequality follows from $\pi_{DD} - \pi_{DV} < \pi_{VD} - \pi_{VV}$, which is derived from the condition for strategic substitutability, and $\pi_{VD} - \pi_{DD} - K_2 \geq \pi_{VD} - \pi_{DD} - \bar{\bar{K}}_1 = \frac{1}{2}[(\pi_{VD} - \pi_{DD}) - (\pi_{VV} - \pi_{DV})]$.

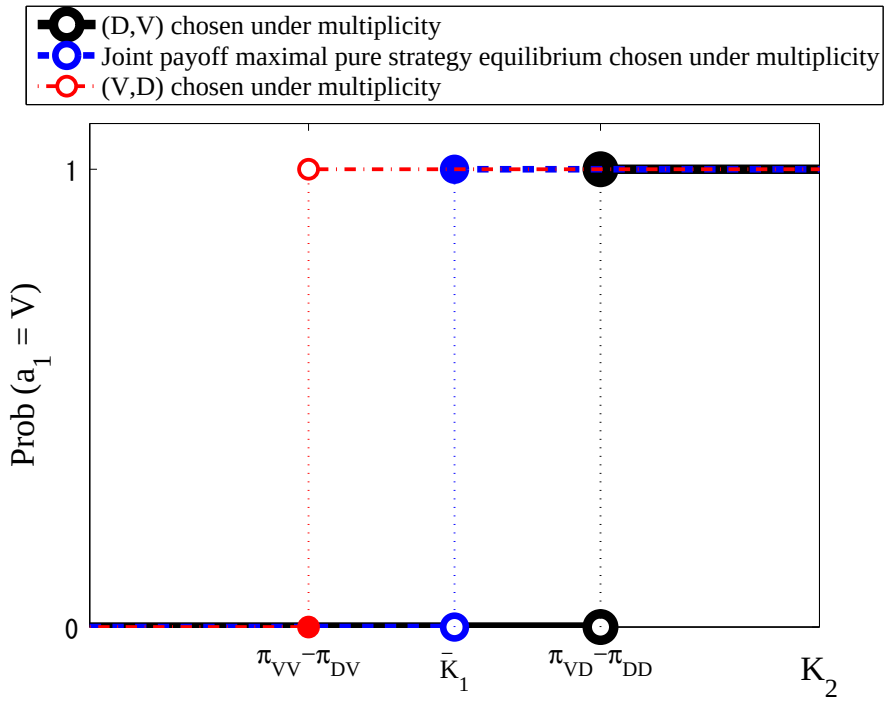
[36] Let us rewrite the joint payoffs under the mixed strategy equilibrium as

$$\Pi_{MS} = (1-q_1)\pi_{DD} + q_1\pi_{DV} + (1-q_2)\pi_{VD} + q_2\pi_{VV} - \bar{\bar{K}}_1.$$

Subtract from it $\Pi_{PS} = \pi_{DV} + \pi_{VD} - \bar{\bar{K}}_1$, the joint payoffs under $(V,D)$, which is the joint payoff maximal pure strategy equilibrium when $K_2 > \bar{\bar{K}}_1$:

$$\begin{aligned}\Pi_{MS} - \Pi_{PS} &= (1-q_1)(\pi_{DD} - \pi_{DV}) - q_2(\pi_{VD} - \pi_{VV}) \\ &= \frac{K_2 - (\pi_{VV} - \pi_{DV})}{(\pi_{VD} - \pi_{DD}) - (\pi_{VV} - \pi_{DV})}(\pi_{DD} - \pi_{DV}) - \frac{1}{2}(\pi_{VD} - \pi_{VV}).\end{aligned}$$

Rearranging terms shows that this expression is negative if and only if $K_2 < \bar{\kappa}$.



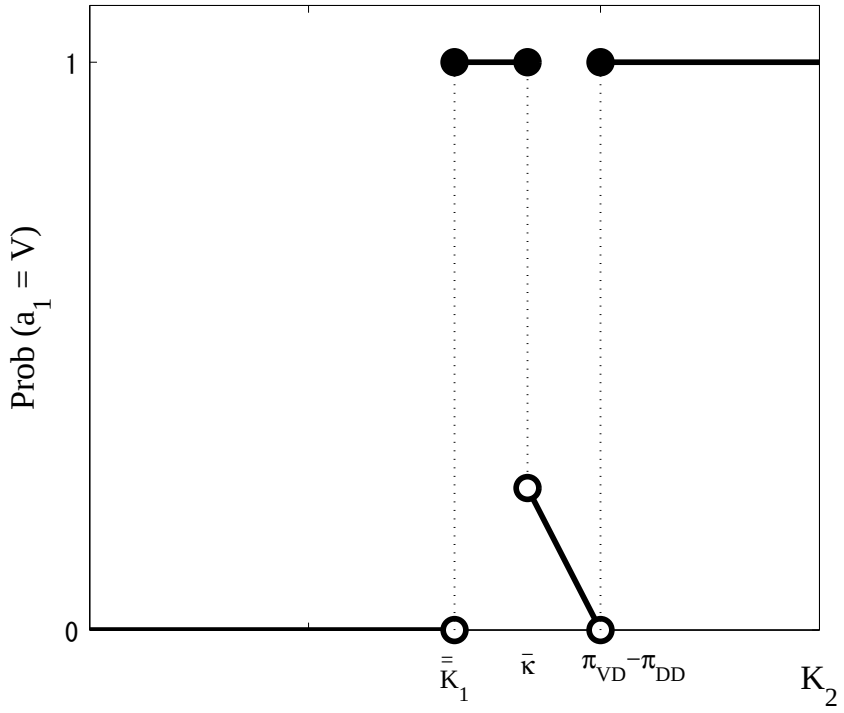

Notes:
The graphs represent firm 1's vertical integration probabilities under different equilibrium selection rules.

Figure 2.3: Firm 1's Vertical Integration Probability in Pure Strategy Equilibria Under Strategic Substitutability

are strategic substitutes. Intuitively, a higher vertical integration cost for firm 2 is interpreted by firm 1 as a lower probability of integration by its rival. Under strategic substitutability, this results in a higher probability of vertical integration by firm 1.

**Testing for Bandwagon Effects**

The main results represented by Figures 2.2 and 2.3 can form the basis for an empirical test of strategic complementarity or substitutability in vertical integration decisions. Suppose that one has data on multiple markets where a number of firms make entry and vertical integration decisions simultaneously. Suppose also that one has prior information that a particular firm characteristic – call it $z$ – affects the cost of vertical integration. Then, the test consists of measuring the effect of $\mathbf{z}_{-i}$, the vector containing the characteristics of firms other than $i$, on the probability that firm $i$ chooses to enter vertically. If $z_i$ has a cost-lowering effect and vertical integration decisions are strategic complements, we would expect the elements of the vector $\partial Prob(a_i = V)/\partial \mathbf{z}_{-i}$ to




Notes:
The graph represents firm 1's vertical integration probability when the equilibrium with the highest joint payoffs is always chosen.

Figure 2.4: Firm 1's Vertical Integration Probability Under Strategic Substitutability When Mixed Strategy Equilibrium Is Possible

be positive. If vertical integration is characterized by strategic substitutability, the derivatives are expected to have a negative sign. This suggests that the existence of strategic complementarity – and by association, bandwagon effects – can be tested in a reduced-form regression framework similar to the one used to analyze peer effects in youth behavior (e.g., Case and Katz, 1991; Evans et al., 1992).

A good candidate for *z* is the firm's previous entry experience. Earlier studies on the generic drug industry by Scott Morton (1999), Gallant et al. (2008), and others have shown that previous experience in entering similar markets has a significantly positive effect on entry probabilities. They conclude from this that previous entry experience lowers current entry costs. While these authors only examine downstream finished formulation markets, it is likely that previous entry experience lowers current entry costs in the upstream API segment as well.

If a firm's previous upstream entry experience is indeed associated with a lower cost of upstream entry, then it should also be associated with a lower cost of vertical integration by down-

stream entrants. In the subsequent empirical analysis, we use the potential downstream entrant's own upstream experience, as well as the upstream experience of the other potential downstream entrants, as covariates in order to test the strategic complementarity of vertical integration decisions.

### 2.3.2 Relationship Specificity of Investments to Support Patent Challenges

As discussed in Section 2.2.2, generic entrants engage in a race to be the first-to-file ANDA applicant when a market is characterized by a paragraph IV patent challenge. In such markets, early access to APIs, which enables early ANDA filings, is particularly important for the profitability of downstream entrants. Here, we examine how vertical integration might provide downstream entrants with earlier access to APIs than would be possible under vertical separation. According to the transaction cost and property rights theories of the firm (e.g., Williamson, 1971; Klein et al., 1979; Grossman and Hart, 1986), vertical integration facilitates investments that are characterized by relationship specificity and non-contractibility. A relationship-specific investment is one that has a greater value within a particular vertical relationship than in others (Grossman and Hart, 1986). Using a simple model, I show that the development of APIs to support a paragraph IV patent challenge fits this definition. I then demonstrate how vertical integration may facilitate the early development of APIs when supply contracts are incomplete *ex ante*.

#### Product Development in a Paragraph IV Market

In the previous subsection, where we focused on the vertical integration decisions of downstream entrants, we implicitly assumed the timing and cost of product development to be fixed. Here, we take the potential entrants' entry and vertical integration decisions as given and focus on the process of product development. We allow both the timing and cost of development to vary.

The timeline of events for a particular generic drug market is depicted in Figure 2.5 where one period is equal to one year. Activity begins at time 0 when the FDA approves the original product. The basic product patent for the drug expires at time $T_b$. The drug is also covered by a secondary patent that expires at $T_s > T_b$. Generic drug companies become aware of the product at time 0 and make plans for product development and regulatory filings. If a generic firm makes a paragraph IV patent certification with respect to the secondary patent, it aims to file its ANDA several periods before $T_b$. There are two reasons for the early filing. First, the target entry timing of a paragraph IV applicant is $T_b$, and the ANDA must be filed a few periods before that to give the FDA sufficient review time. To simplify, I assume that the ANDA review process takes two periods. Second, and more importantly, the paragraph IV applicant files the ANDA early in order to be ahead of its rivals.

A paragraph IV certification is met with a patent infringement suit by the originator, but it is assumed that the generic defendant invalidates the secondary patent and wins the suit with certainty. The FDA gives final approval to the first-to-file paragraph IV applicant's ANDA at $T_b$, whereupon the firm begins commercial marketing. The first-to-file firm enjoys exclusivity in the

generic market from $T_b$ to $T_b + \frac{1}{2}$, which corresponds to the 180-day generic exclusivity period. If a generic firm plans to enter without a patent challenge, it files an ANDA at or near $T_b - \frac{3}{2}$. The FDA spends two periods to review the ANDA, and gives final approval at $T_b + \frac{1}{2}$. Non-challengers begin commercial marketing at $T_b + \frac{1}{2}$, as do paragraph IV applicants who fail to be the first to file.

I assume that the timing of successful API development by upstream units is a random variable, and that downstream units are able to develop finished formulations immediately after the API becomes available. Following Loury's (1979) model of a patent race, the timing of successful API development by an upstream unit depends on the level of investment chosen by it at time 0. Let $y_i$ be the level of investment by an unintegrated upstream entrant $i$. The probability that the timing of success, denoted by $\vartheta(y_i)$, is earlier than $t$ is $Pr[\vartheta(y_i) < t] = 1 - e^{-h(y_i)t}$. Following Reinganum (1983), the hazard function $h(\cdot)$ is assumed to have the following characteristics: $h(0) = 0$, $h'(y) > 0$, and $h''(y) < 0$ for $y \in \mathbb{R}_+$. For the upstream unit of a vertically integrated entrant, the corresponding hazard function is $g(\cdot)$, with $g(y) < h(y), \ \forall y \in \mathbb{R}_+$. That the vertically integrated entrant has a lower hazard rate than the unintegrated entrant at the same level of investment implies the lower efficiency of the former in API development. This reflects the organizational inefficiencies due to vertical integration. The random variable $\vartheta(y_i)$ is assumed to be independent across firms.

I also make the following assumption: the hazard functions are such that upstream units face little uncertainty with respect to investment outcomes when they are not involved in challenging the secondary patent. In other words, if an upstream unit invests to maximize its expected payoff from supporting a non-challenger's ANDA, the probability that its API is successfully developed by $T_b - \frac{3}{2}$, the time for filing the ANDA, is very close to one. This implies that upstream units supporting a paragraph IV applicant, who are likely to invest more into API development and who expect their downstream users' ANDAs to be filed by $T_b - 2$ at the latest, face a probability near one of successfully developing its API by $T_b - 2$. In other words, there is almost always some paragraph IV applicant who is successful at obtaining generic exclusivity. This assumption simplifies the subsequent analysis.

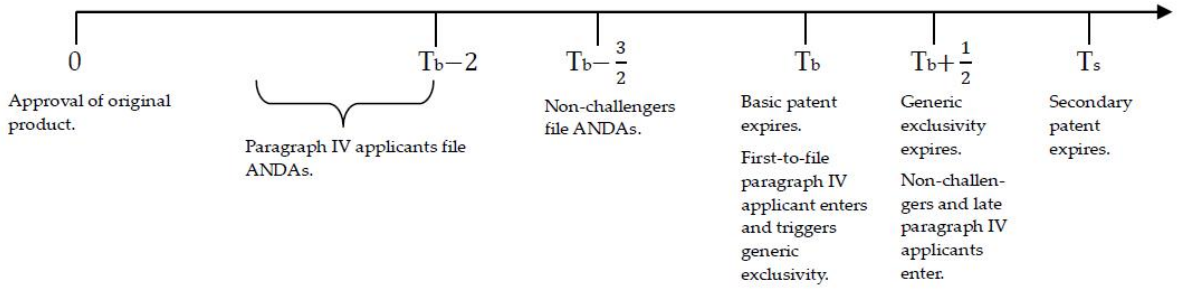



Figure 2.5: Timeline of Events in a Market Characterized by Patent Challenge

**Firm Revenues**

The expected revenue of a non-challenger in the finished formulation market, net of downstream processing costs, is

$$\pi_{pe}\left\{1-e^{-[\sum_{i\in\mathcal{U}_c}h(y_i)+\sum_{i\in\mathcal{V}_c}g(y_i)](T_b-2)}\right\}+\pi_{ps}e^{-[\sum_{i\in\mathcal{U}_c}h(y_i)+\sum_{i\in\mathcal{V}_c}g(y_i)](T_b-2)}\approx\pi_{pe},$$

where $\pi_{pe}$ and $\pi_{ps}$ are the revenues receivable from entering at $T_b+\frac{1}{2}$ and $T_s$, respectively. Both revenue figures are in present values as of time zero. The subscript *pe* stands for "post-exclusivity" and *ps* stands for "post-secondary patent". $\mathcal{U}_c$ is the set of unintegrated upstream firms that are involved in a paragraph IV patent challenge (the subscript *c* stands for "challenger") and $\mathcal{V}_c$ is the set of upstream units belonging to vertically integrated paragraph IV applicants. $e^{-[\sum_{i\in\mathcal{U}_c}h(y_i)+\sum_{i\in\mathcal{V}_c}g(y_i)](T_b-2)}$ is the probability that none of the paragraph IV applicants succeed at developing their APIs by $T_b-2$. The approximation follows from the assumption that this probability is close to zero.

The expected revenue of a paragraph IV applicant, after it finds out that it is not the first to file, is $\pi_{pe}$. On the other hand, the revenue of the first-to-file paragraph IV applicant is $\pi_e+\pi_{pe}$, where $\pi_e$ is the revenue earned during the generic exclusivity period.

A set of assumptions is employed regarding the way revenues and profits are shared between upstream and downstream units. For simplicity, let trades of API occur only within pairs consisting of one upstream unit and one downstream unit. In the case of a pair consisting of two separate firms, downstream revenue is divided through *ex post* bargaining. I assume that the bargaining takes place after investments have been made and the identity of the first-to-file paragraph IV applicant is known. *Ex ante* contracts for the supply of API are ruled out. This assumption may require some justification. In particular, readers may wonder why the firms don't enter into a contingent contract – e.g., one that specifies a high payment to the upstream firm only in the event of the pair winning first-to-file status.

The problem with a contingent contract in practice is that the upstream firm's investment performance (timing of successful API development relative to rivals) and its contribution to the final outcome (the timing of filing the paragraph IV ANDA) may be unverifiable. This is because in reality, unlike in the present model, the speed of generic product development depends to some extent on investment by the downstream unit. In addition, there is some anecdotal evidence, contained in court records, that buyers often breach API supply contracts without having to pay penalties.[37]

For a vertically integrated entrant, I assume that profit (revenue minus API development cost) is divided between the vertical units in fixed proportions. This profit-sharing assumption, borrowed from Hart and Tirole (1990), essentially assumes that "under integration, profits of the parent and subsidiary are commingled in such a way that profit sharing is inevitable" (Hart and Tirole, 1990, p.217).

[37]Many instances of API contracts being breached are described in the court's opinion for *Geneva and Apothecon v. Barr et al.* (S.D.N.Y., 2002). In one particular case in 1995, a downstream firm made a 1.8 million dollar purchase order for 2,500 kilograms of API from an upstream supplier. The purchase order was canceled eighteen months later, with more than 1,500 kilograms yet to be delivered.

Let us consider the *ex post* division of revenue by an unintegrated upstream-downstream pair who do not make a patent challenge. Assuming that there is some other firm who succeeds in a patent challenge, $\pi_{pe}$ becomes available to the pair at $T_b + \frac{1}{2}$ if the upstream firm succeeds at API development by $T_b - \frac{3}{2}$. In principle, bargaining over the division of this revenue can take a very complex form, involving all possible trading partners of the two firms (de Fontenay and Gans, 2005). For the sake of simplicity, however, I assume that the revenue is split evenly between the two firms. Thus, the upstream entrant – call it firm $i$ – solves the following maximization problem when it chooses its investment level:

$$\max_{y_i} \frac{1}{2}\pi_{pe}\left[1 - e^{-h(y_i)\left(T_b - \frac{3}{2}\right)}\right] - y_i. \tag{2.7}$$

The earlier assumption that upstream entrants succeed at API development by $T_b - \frac{3}{2}$ with probability near one is equivalent to assuming that $1 - e^{-h(y_i^*)\left(T_b - \frac{3}{2}\right)} \approx 1$, where $y_i^*$ is implicitly defined by the first order condition of (2.7).

Now, consider the division of revenue by an unintegrated pair who pursues a paragraph IV patent challenge. When the pair loses in the race to be first-to-file, each firm's revenue is the same as when there is no patent challenge: both receive $\frac{1}{2}\pi_{pe}$. When the pair wins generic exclusivity, I assume that the extra revenue during the exclusivity period, $\pi_e$, is split according to the Nash bargaining solution.

Let us consider the firms' outside options. The downstream firm, who owns the first-to-file paragraph IV ANDA, can enjoy the generic exclusivity revenue even if it does not trade with its original partner. Between winning first-to-file status (some time before $T_b - 2$) and receiving final approval (at $T_b$), the downstream firm has sufficient time to find another API manufacturer who is willing to supply during the exclusivity period. It must cede some portion (say, $\gamma < 1$) of revenue during the exclusivity period to the alternative supplier, but the downstream firm keeps the major share.[38]

Meanwhile, the original upstream partner has no claim on generic exclusivity except through its relationship with the downstream firm. Therefore, its outside option during the exclusivity period is zero. Note that the upstream firm's investment into API development is characterized by relationship specificity: its product generates a revenue of $\pi_e$ if supplied to the first-to-file ANDA applicant during the exclusivity period, but if supplied to another user, it generates zero revenue. Investment by the upstream firm is crucial for winning the right to generic exclusivity. Yet, the firm has no ownership claim over this valuable asset.

Given the outside options, the Nash bargaining solution for the upstream partner's revenue when the pair wins generic exclusivity is $\frac{1}{2}[\pi_e - (1-\gamma)\pi_e] + \frac{1}{2}\pi_{pe} = \frac{1}{2}(\gamma\pi_e + \pi_{pe})$. The first term

[38]The existence of an alternative supplier is assured with near certainty if there are multiple upstream units pursuing a paragraph IV patent challenge, because given the earlier assumption, each successfully develops its API by $T_b - 2$ with probability close to one. Even if there is only one upstream firm that pursues a patent challenge (i.e., the downstream firm's original partner), the downstream firm can contract with other potential suppliers, at relatively low cost, to develop the API by $T_b - 2$.

on the left-hand side is the revenue from supplying the downstream partner during the generic exclusivity period and the second term is the revenue during the post-exclusivity period. $\gamma$, the share of exclusivity period revenue that the downstream firm must pay to an alternative API supplier, is expected to be small (i.e., not much larger than zero). Therefore, the upstream firm's revenue from supplying the first-to-file paragraph IV applicant is only slightly larger than its revenue from supplying a non-challenger.

**Equilibrium Investments**

The equilibrium level of investments by upstream units can be derived as the solution of an investment race. For simplicity, I assume that two upstream units are present in the market. The first unit, labeled $u$, is an unintegrated firm who supplies API to an unintegrated downstream firm. The second unit is a subsidiary of a vertically integrated entrant labeled $v$. It produces API exclusively for in-house use. Both upstream units participate in a paragraph IV patent challenge. Following Reinganum (1983), the race outcome is derived as a Nash equilibrium with investment as the strategic variable.

Given the formula for the probability of successful API development, the probability density of firm $u$ succeeding at time $t$ is $h(y_u)e^{-h(y_u)t}$. The density for the probability of firm $u$ winning the race at time $t$ is therefore

$$h(y_u)e^{-h(y_u)t}\left[1-\left(1-e^{-g(y_v)t}\right)\right] = h(y_u)e^{-[h(y_u)+g(y_v)]t}.$$

Integrating from 0 to $T_b-2$ gives the probability that $u$ wins the race:

$$\begin{aligned} Prob(u \text{ wins}) &= \int_0^{T_b-2} h(y_u)e^{-[h(y_u)+g(y_v)]t}dt \\ &= \frac{h(y_u)}{h(y_u)+g(y_v)}\left[1-e^{-[h(y_u)+g(y_v)](T_b-2)}\right]. \end{aligned}$$

The probability of firm $v$ winning the race is analogously derived.

The profit maximization problem for firm $u$ is the following:

$$\max_{y_u} \frac{1}{2}\left\{\gamma\pi_e \frac{h(y_u)}{h(y_u)+g(y_v)}\left[1-e^{-[h(y_u)+g(y_v)](T_b-2)}\right]+\pi_{pe}\right\}-y_u.$$

Firm $u$'s best response to firm $v$'s investment is implicitly defined by the following first-order condition:

$$\begin{aligned} &\frac{1}{2}\gamma\pi_e\left\{\frac{h'(y_u)g(y_v)}{[h(y_u)+g(y_v)]^2}\left[1-e^{-[h(y_u)+g(y_v)](T_b-2)}\right]\right. \\ &\left.+\frac{h'(y_u)h(y_u)(T_b-2)}{h(y_u)+g(y_v)}e^{-[h(y_u)+g(y_v)](T_b-2)}\right\}=1. \end{aligned}$$

Applying the assumption that both upstream firms are successful at developing API by $T_b - 2$ with probability near one, this expression simplifies to

$$\frac{1}{2}\gamma\pi_e \frac{h'(y_u)g(y_v)}{[h(y_u)+g(y_v)]^2} \approx 1. \tag{2.8}$$

The upstream unit of the vertically integrated firm $v$ faces the following profit maximization problem:

$$\max_{y_v} \xi \left\{ \pi_e \frac{g(y_v)}{h(y_u)+g(y_v)} \left[1 - e^{-[h(y_u)+g(y_v)](T_b-2)}\right] + \pi_{pe} - y_v \right\},$$

where $\xi$ is the upstream unit's share of profits. The first-order condition that implicitly defines firm $v$'s best response to firm $u$'s investment level is

$$\pi_e \left\{ \frac{g'(y_v)h(y_u)}{[h(y_u)+g(y_v)]^2} \left[1 - e^{-[h(y_u)+g(y_v)](T_b-2)}\right] \right.$$

$$\left. + \frac{g'(y_v)g(y_v)(T_b-2)}{h(y_u)+g(y_v)} e^{-[h(y_u)+g(y_v)](T_b-2)} \right\} = 1.$$

which simplifies to

$$\pi_e \frac{g'(y_v)h(y_u)}{[h(y_u)+g(y_v)]^2} \approx 1. \tag{2.9}$$

We assumed earlier that $g(y) < h(y)$, to represent the organizational inefficiency of a vertically integrated entrant in terms of API development. Let us simplify by assuming that $g(y) = \varphi h(y)$ with $\varphi < 1$. Equating the left-hand sides of (2.8) and (2.9) yields the following result regarding the equilibrium level of investment by the upstream units:

$$\frac{h'(y_v^*)/h(y_v^*)}{h'(y_u^*)/h(y_u^*)} = \frac{\gamma}{2} < 1. \tag{2.10}$$

where the inequality follows from the definition of $\gamma$. Inequality (2.10) and the assumption of $h(\cdot)$ being an increasing and concave function imply that $y_v^* > y_u^*$. Therefore, firm $v$ invests more into API development than firm $u$ in equilibrium.

In order to compare the two firms' probability of winning the race, we require knowledge regarding the functional form of $h(\cdot)$. For the purpose of illustration, let us assume that $h(y) = \sqrt{y}$. Then, (2.10) simplifies to $y_v^* = \frac{2y_u^*}{\gamma}$. The hazard rate defining firm $v$'s success probability, $g(y_v^*)$, is

equal to $\varphi\sqrt{\frac{2y_u^*}{\gamma}}$. This is greater than $h(y_u^*) = \sqrt{y_u^*}$ if and only if $\varphi > \sqrt{\frac{\gamma}{2}}$. Therefore, the vertically integrated entrant has a higher probability of winning the investment race than the unintegrated upstream entrant as long as the organizational inefficiency due to vertical integration is not too severe.

The result that the vertically integrated firm is likely to have a higher probability of winning is driven by the assumption of *ex post* bargaining between the members of the unintegrated pair. The relationship specificity of API development investments in the context of a patent challenge, combined with the fact that the ANDA is owned by the downstream firm, contributes to the weak bargaining position of the unintegrated upstream firm. Expecting lower profits than its vertically integrated counterpart, the unintegrated upstream firm invests less in equilibrium.

**Implication for Empirical Analysis**

The prediction that vertical integration facilitates early API development during a patent challenge can be tested by seeing if ANDA applicants who make a paragraph IV certification are more likely than other applicants to be vertically integrated. However, my dataset only records whether or not each market is subject to one or more entrants making a paragraph IV certification. I therefore construct a market-level variable that indicates the occurrence of a paragraph IV patent challenge. This indicator variable essentially signifies a switch in the entry process: markets with no paragraph IV patent challenge are characterized by simultaneous entry, while paragraph IV markets are characterized by a race to be first. The empirical strategy is to see whether this switch in the entry process affects firms' incentives to become vertically integrated.

Using the market-level paragraph IV indicator variable as a determinant of firm-level behavior introduces a potential endogeneity problem: markets that are the subject of paragraph IV certification may be attractive to generic entrants in unobservable ways, and those unobserved factors may also influence entry and vertical integration decisions. This endogeneity can be taken care of by modeling the process of paragraph IV certification, and allowing the error term in the firm-level equations and that in the paragraph IV equation to be correlated.

Many authors note that paragraph IV patent challenges have become more common in recent years (Grabowski, 2004; Grabowski and Kyle, 2007; Higgins and Graham, 2006; Hemphill and Sampat, 2010). Patent challenges may also be more likely in larger markets that offer greater profits to the first-to-file entrant during the exclusivity period. In addition, Grabowski (2004) and Hemphill and Sampat (2010) note that certain types of secondary patents – particularly those that cover formulations and new uses – tend to be more vulnerable to patent challenge, presumably because it is easier to invalidate or avoid infringing such patents. This suggests the following as possible market-level determinants of paragraph IV certification: market size, the number of originator patents of different types, and year dummy variables.

## 2.4 Econometric Specification

The object of estimation is the set of payoff equations for potential entrants in the generic drug industry. One equation is defined for each alternative in the firms' choice set: unintegrated downstream entry ($D$), unintegrated upstream entry ($U$), vertically integrated entry ($V$), and no entry. Let $i = 1,2,...,I$ index potential entrants and $m = 1,2,...,M$ index drug markets. Also, let the vector $\mathbf{N} = (N_D, N_U, N_V)$ summarize a vertical market structure, with $N_j$ denoting the number of entrants in category $j$. Then, the payoffs of individual firms are represented by the following functions:

$$\begin{aligned}
\text{Payoff of firm } i \text{ from unintegrated downstream entry} &= \pi_D(\mathbf{w}_m, \mathbf{x}_i, \mathbf{N} - \iota_1) \\
\text{Payoff of firm } i \text{ from unintegrated upstream entry} &= \pi_U(\mathbf{w}_m, \mathbf{x}_j, \mathbf{N} - \iota_2) \\
\text{Payoff of firm } i \text{ from vertically integrated entry} &= \pi_V(\mathbf{w}_m, \mathbf{x}_i, \mathbf{N} - \iota_3) \\
\text{Payoff from no entry} &= 0
\end{aligned}$$

where $\mathbf{w}_m$ is a vector consisting of the characteristics of market $m$ and $\mathbf{x}_i$ is a vector containing the characteristics of firm $i$. $\iota_n$ is a three dimensional unit vector containing one as the $n$th element and zeros for the other elements. The reason for subtracting $\iota_n$ from the market structure vector is to avoid including a firm's own action as an argument of its action-specific payoff function. These payoffs are net of product development investments – i.e., sunk entry costs – that are functions of firm and market characteristics. Following common practice in the empirical entry literature (e.g., Berry, 1992), I assume that a firm's payoff is affected by a rival firm only through the latter's action, so that the payoff functions do not contain the characteristics of rivals as arguments. I also assume that the payoff impact of one rival's entry is identical to that of another's. This allows us to aggregate the payoff impact of rivals into a term involving the three dimensional vector $\mathbf{N}$.

I assume that $\mathbf{N}$ is generated as an equilibrium of an entry game into vertical oligopoly. Such games are generally characterized by multiple equilibria (Elberfeld, 2002). To simplify the analysis, I assume that potential entrants follow a common equilibrium selection rule such as one where the equilibrium with the highest joint profits is realized (e.g., Berry, 1992; Scott Morton, 1999). Therefore, the same unique equilibrium is always chosen for a given set of values for the exogenous variables. This implies that we can define a function that maps from the exogenous variables (market characteristics and firm characteristics of every potential entrant) to market structure outcomes. The existence of such a function allows us to rewrite payoffs in the following reduced form:

$$\begin{aligned}
\text{Payoff from unintegrated downstream entry} &= \pi_D(\mathbf{w}_m, \mathbf{x}_i, \mathbf{X}_{-i}) \\
\text{Payoff from unintegrated upstream entry} &= \pi_U(\mathbf{w}_m, \mathbf{x}_i, \mathbf{X}_{-i}) \\
\text{Payoff from vertically integrated entry} &= \pi_V(\mathbf{w}_m, \mathbf{x}_i, \mathbf{X}_{-i}) \\
\text{Payoff from no entry} &= 0
\end{aligned}$$

where $\mathbf{X}_{-i}$ is a matrix containing the firm characteristics of all potential entrants excluding firm $i$. These reduced-form payoff equations can be used for examining the determinants of the vertical integration decision.

To keep the estimation tractable, and to facilitate the interpretation of the results, I focus only on backward integration by downstream entrants. I assume that potential downstream entrants, defined to include all firms that are potential entrants in the downstream segment, follow a two-stage decision process. In the first stage, they decide whether or not to enter the downstream finished formulation segment of the market based on the following criterion:

$$\text{Enter downstream if and only if } \max(\pi_D, \pi_V) > \max(0, \pi_U),$$

where the arguments of the payoff equations have been abbreviated for brevity. Conditional on entering downstream, firms then decide whether to enter the upstream API segment as well – in other words, whether to vertically integrate. Thus, the second stage decision is as follows:

$$\text{Vertically integrate if and only if } \pi_V - \pi_D > 0.$$

This framework suggests the use of a bivariate discrete choice model with sample selection – for example, the censored probit model of Meng and Schmidt (1985).

The model is slightly complicated by the inclusion of an indicator for paragraph IV certification as a covariate. The potential endogeneity of this variable leads us to employ a trivariate discrete choice model with sample selection and endogeneity. By assuming a normal distribution for the error term vector, the following trivariate probit model is specified:

$$\begin{aligned}
y^*_{1mi} &= \beta_1' \mathbf{x}_{1mi} + \alpha\, PF_m + \varepsilon_{1mi} \\
y^*_{2mi} &= \beta_2' \mathbf{x}_{2mi} + \varepsilon_{2mi} \\
y^*_{3m} &= \beta_3' \mathbf{x}_{3m} + \varepsilon_{3m},
\end{aligned}$$

$$\begin{aligned}
VI_{mi} &= \mathbf{1}(y^*_{1mi} > 0) \times DE_{mi} \\
DE_{mi} &= \mathbf{1}(y^*_{2mi} > 0) \\
PF_m &= \mathbf{1}(y^*_{3m} > 0),
\end{aligned} \tag{2.11}$$

$$(\varepsilon_{1mi}, \varepsilon_{2mi}, \varepsilon_{3m}) \sim \mathcal{N}(0, \Sigma).$$

The model contains three dichotomous endogenous variables. $DE_{mi}$ is an indicator for firm $i$'s entry into the downstream segment of market $m$. $VI_{mi}$ indicates that firm $i$ enters as a vertically integrated firm. It is observed only if firm $i$ enters the downstream segment. $PF_m$ is a market-level indicator of paragraph IV certification by one or more downstream entrants. The $\beta$ vectors and $\alpha$ are unknown parameters to be estimated, and the covariance matrix of the normally distributed error term vector is assumed to have the following form:

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & 0 \\ \rho_{13} & 0 & 1 \end{bmatrix},$$

where $\rho_{mn}$ is the correlation coefficient between $\varepsilon_m$ and $\varepsilon_n$ and $\rho_{23} = 0$ is assumed.

The row vector of covariates in the outcome equation, $\mathbf{x}_{1mi}$, contains both market and firm characteristics. The market characteristics represent the revenue potential of the market as perceived by generic firms as well as the costs required for entering. They include measures of market size, the willingness of patients and other payers (e.g., insurers) to pay for the drug, and dummy variables for different therapeutic classes and dosage forms. The first two variables measure the market's revenue potential, while the dummy variables capture both revenue potential and magnitude of entry costs. The sole firm characteristic contained in $\mathbf{x}_{1mi}$ is the firm's experience in entering the upstream segment of markets that opened up previously. We can expect higher values of this variable to be associated with lower vertical integration costs.

Another set of variables in $\mathbf{x}_{1mi}$ is generated from the characteristics of other potential entrants in the same market. The first variable is the mean level of upstream entry experience among potential downstream entrants $i' \neq i$. Inclusion of this variable is motivated by the discussion in Section 2.3.1. There, it was shown that lower vertical integration cost among rivals raises the probability of integration by a potential downstream entrant if and only if vertical integration decisions are strategic complements. To the extent that past upstream experience lowers vertical integration costs, the mean upstream experience of rivals can be used to test for strategic complementarity.

The second variable to be constructed from the characteristics of other firms is the number of potential upstream-only entrants, defined as firms who are capable of entering the upstream segment but not the downstream segment. This variable represents the strength of the unintegrated upstream industry. A greater number of potential independent upstream suppliers is expected to lower firm $i$'s probability of vertically integrating.

The paragraph IV indicator enters the vertical integration equation as a potentially endogenous variable. A positive coefficient on this variable in the outcome equation indicates support for the hypothesis that vertical integration facilitates the early development of API when pursuing a patent challenge.

$\mathbf{x}_{3m}$ consists of variables that influence the incidence of paragraph IV certification at the market level. In addition to the market characteristics contained in $\mathbf{x}_{1mi}$, I include the following: the number of potential downstream entrants in market $m$, the mean level of upstream experience among potential downstream entrants, and the number of potential upstream-only entrants.[39] These three variables are expected to affect the post-entry market structure. To the extent that they also affect post-entry profits, the variables are also likely to affect the firms' paragraph IV decisions. For

[39]The "mean level of upstream experience among potential downstream entrants" variable is slightly different from the "mean level of upstream experience among potential downstream rivals" contained in $\mathbf{x}_{1mi}$ in that firm $i$'s upstream experience is excluded from the calculation of the latter.

instance, a firm may be more likely to engage in a patent challenge if it expects stiffer downstream competition.

Two variables related to the number of originator patents are also included in $\mathbf{x}_{3m}$. The first one measures the number of patents pertaining to the API – namely, product patents and process patents. The second variable measures the number of formulation patents and new use patents, which are more closely associated with the finished drug product. These variables can be used to check whether patent challenges are more likely when there are more patents to serve as targets for paragraph IV certification. Of particular interest is whether formulation and new use patents are more likely to attract challenges as suggested by Grabowski (2004) and Hemphill and Sampat (2010). The two patent-related variables are excluded from $\mathbf{x}_{1mi}$ based on the assumption that originator patents affect the vertical integration decisions of generic entrants only through their effect on the paragraph IV status of the market. The justification for this exclusion restriction is as follows. First, patents that are not the subject of paragraph IV certification are either those that are too strong to be challenged, or those that are clearly incapable of blocking generic entry.[40] In either case, they are unlikely to influence the vertical integration decisions of generic firms. Second, given that a market is subject to paragraph IV certification, the number of patents is unlikely to matter for the vertical integration decision.

The vector of covariates for the selection equation, $\mathbf{x}_{2mi}$, contains all of the variables in $\mathbf{x}_{1mi}$. Additional variables that are expected to influence the downstream entry decision, but not the vertical integration decision, are also included. First, the firm's downstream entry experience in past markets is included to represent its downstream entry cost. Second, the number of rival potential entrants in the downstream segment, representing the intensity of competition in the entry game, is included.[41] Although paragraph IV certification is expected to have an influence on the downstream entry decision, instead of including it in the selection equation, I put the two patent variables contained in $\mathbf{x}_{3m}$ into $\mathbf{x}_{2mi}$.[42] Thus, the selection equation can be thought of as being in a reduced form with respect to the effect of paragraph IV certification. Year dummy variables are included in $\mathbf{x}_{1mi}$, $\mathbf{x}_{2mi}$, and $\mathbf{x}_{3m}$ to control for unobserved time effects that may be correlated with some of the market and firm characteristics.

The inclusion of previous entry experience in the covariate vectors gives rise to two econometric concerns. The first is the possible correlation between past entry experience on the one hand, and $\varepsilon_{1mi}$ and $\varepsilon_{2mi}$ on the other. This would arise, for instance, if the error terms contain the effect of the firm's unobserved proficiency at developing certain types of products (e.g., injectable drugs),

[40] The patent data that I use to construct the two variables contain both patents that are listed by the originator in the Orange Book as well as those that are not. While listed patents become the subject of paragraph IV certification even if they are clearly non-blocking, patents that are not listed and that are non-blocking can be ignored by generic entrants.

[41] The difference between "number of potential downstream entrants" in $\mathbf{x}_{3m}$ and "number of potential downstream rivals" in $\mathbf{x}_{2mi}$ is that the latter does not count firm $i$.

[42] By replacing the paragraph IV indicator with the variables in $\mathbf{x}_{3m}$, I can assume that $\varepsilon_{2mi}$ and $\varepsilon_{3m}$ are uncorrelated. This facilitates estimation by preventing numerical problems, similar to the one pointed out by Butler (1996), that arise in the estimation of correlation coefficients.

which may be positively correlated with the firm's past entry experience. Ignoring the possible correlation may lead to upwardly biased estimates for the coefficients on the experience variables. The second concern is the possibility of forward-looking behavior by the firms. As Gallant et al. (2008, 2010) argue, generic drug manufacturers may consider, when making their entry decisions, how their actions in the current market affect their entry costs in future markets. For example, a firm may decide to enter a market this year, even though it earns no direct profit from doing so, just because the resulting accumulation of experience would lower its costs and raise the profitability of entering another market next year. Ignoring such forward-looking behavior may introduce bias into the coefficient estimates, but the direction of bias is not clear *a priori*.[43]

By employing the specification in (2.11), which ignores the potential endogeneity of the experience variables as well as the possible dynamics in firm behavior, I am implicitly assuming that the above concerns are not severe. The grounds for doing so are the following. First, if a firm is especially proficient at developing a certain type of product, it is most likely due to the accumulation of experience in developing such products. In other words, the past entry experience variable can be interpreted as a proxy for unobserved proficiencies. Second, unless the managers of generic drug companies are compensated based on their firms' long-term performance, the entry decisions made by them are unlikely to reflect dynamic solutions that are optimal for the firms' shareholders. Given the large number of mergers and acquisitions in this industry and the resulting high rate of employee turnover, it is likely that managers' decisions are more myopic than what their shareholders would like them to be.[44]

Before deriving the estimator, it is important to note that the paragraph IV equation is defined at the market level whereas the other equations are defined at the level of individual firms. In addition, it is possible that the firm-level error terms are correlated within markets due, for instance, to the existence of unobserved market effects. In this setting, the true likelihood function must be based on likelihood contributions defined at the market level. Each market's likelihood contribution is calculated by integrating over the joint distribution of $\varepsilon_{3m}$ and all the elements of $\{\varepsilon_{mi}\}_{i\in\mathfrak{P}_{Dm}}$, where $\varepsilon_{mi} = (\varepsilon_{1mi}, \varepsilon_{2mi})$ and $\mathfrak{P}_{Dm}$ is the set of potential downstream entrants in market $m$. Thus, estimation based on the true likelihood function requires the calculation of complicated integrals with high dimensionality.

Fortunately, consistent estimates of the parameters can be obtained by maximizing a "partial likelihood" rather than the true likelihood (Wooldridge, 2002, p.401). The partial log likelihood function is based on likelihood contributions defined at the firm level, as follows:

[43] Biases arising from forward-looking behavior can be avoided by estimating a model in which firms' decisions are based on "continuation payoffs" rather than on payoffs in the current market. Bajari et al. (2007) offer one way to implement such a strategy.

[44] Erdei (2004) notes that "the generics sector has been one of the most mergers and acquisitions (M&A)-driven subsectors within the pharmaceutical industry" (p.18). Karwal (2006) contains a list of the major M&A deals in the generics industry during 2004-2006.

$$
\begin{aligned}
\ell(\theta) = & \sum_{m=1}^{M} \sum_{i \in \mathfrak{P}_{Dm}} (1 - DE_{mi}) \ln \int_{-\infty}^{-\beta_2' \mathbf{x}_{2mi}} \phi(\varepsilon_2) d\varepsilon_2 \\
& + VI_{mi}\, DE_{mi}\, PF_m \ln \int_{-\beta_3' \mathbf{x}_{3m}}^{\infty} \int_{-\beta_2' \mathbf{x}_{2mi}}^{\infty} \int_{-\beta_1' \mathbf{x}_{1mi} - \alpha}^{\infty} f_3(\varepsilon_1, \varepsilon_2, \varepsilon_3; \Sigma)\, d\varepsilon_1 d\varepsilon_2 d\varepsilon_3 \\
& + VI_{mi}\, DE_{mi} (1 - PF_m) \ln \int_{-\infty}^{-\beta_3' \mathbf{x}_{3m}} \int_{-\beta_2' \mathbf{x}_{2mi}}^{\infty} \int_{-\beta_1' \mathbf{x}_{1mi}}^{\infty} f_3(\varepsilon_1, \varepsilon_2, \varepsilon_3; \Sigma)\, d\varepsilon_1 d\varepsilon_2 d\varepsilon_3 \\
& + (1 - VI_{mi}) DE_{mi}\, PF_m \ln \int_{-\beta_3' \mathbf{x}_{3m}}^{\infty} \int_{-\beta_2' \mathbf{x}_{2mi}}^{\infty} \int_{-\infty}^{-\beta_1' \mathbf{x}_{1mi} - \alpha} f_3(\varepsilon_1, \varepsilon_2, \varepsilon_3; \Sigma)\, d\varepsilon_1 d\varepsilon_2 d\varepsilon_3 \\
& + (1 - VI_{mi}) DE_{mi} (1 - PF_m) \ln \int_{-\infty}^{-\beta_3' \mathbf{x}_{3m}} \int_{-\beta_2' \mathbf{x}_{2mi}}^{\infty} \int_{-\infty}^{-\beta_1' \mathbf{x}_{1mi}} f_3(\varepsilon_1, \varepsilon_2, \varepsilon_3; \Sigma)\, d\varepsilon_1 d\varepsilon_2 d\varepsilon_3.
\end{aligned} \tag{2.12}
$$

where $\theta = (\beta, \alpha, \rho)$ is the vector of parameters, $M$ is the number of markets in the dataset, and $\phi(\cdot)$ is the standard normal probability density function. $f_3(\cdot; \Sigma)$ is the density for a trivariate normal distribution with zero mean vector and covariance matrix $\Sigma$.

The parameter point estimates can be obtained as if the firm-level observations were independent by maximizing (2.12). However, estimates of their standard errors must be adjusted to account for the clustering of firm-level observations into markets. Following Wooldridge (2002, pp.406-407), the cluster-adjusted asymptotic covariance matrix for the parameters can be written as $\text{Asy.Var}\sqrt{N}(\hat{\theta} - \theta_0) = \mathbf{A}_0^{-1} \mathbf{B}_0 \mathbf{A}_0^{-1}$, where $\theta_0$ is the true parameter value and $\hat{\theta}$ its estimate,

$$
\begin{aligned}
\mathbf{A}_0 &= - \sum_{i \in \mathfrak{P}_{Dm}} \mathrm{E}\left[ \nabla_\theta^2 \ell_{mi}(\theta_0) \right], \\
\mathbf{B}_0 &= \mathrm{E}\left\{ \left[ \sum_{i \in \mathfrak{P}_{Dm}} \mathbf{s}_{mi}(\theta_0) \right] \left[ \sum_{i \in \mathfrak{P}_{Dm}} \mathbf{s}_{mi}(\theta_0) \right]' \right\}, \\
\mathbf{s}_{mi}(\theta) &= \nabla_\theta \ell_{mi}(\theta)',
\end{aligned}
$$

and $\ell_{mi}(\theta)$ is the log likelihood contribution of firm $i$ in market $m$. The expectation is taken over markets. I use the following estimators for $\mathbf{A}_0$ and $\mathbf{B}_0$, as suggested by Wooldridge (2002):

$$
\begin{aligned}
\widehat{\mathbf{A}} &= M^{-1} \sum_{m=1}^{M} \sum_{i \in \mathfrak{P}_{Dm}} \mathbf{s}_{mi}(\hat{\theta}) \mathbf{s}_{mi}(\hat{\theta})', \\
\widehat{\mathbf{B}} &= M^{-1} \sum_{m=1}^{M} \sum_{i \in \mathfrak{P}_{Dm}} \sum_{i' \in \mathfrak{P}_{Dm}} \mathbf{s}_{mi'}(\hat{\theta}) \mathbf{s}_{mi}(\hat{\theta})'.
\end{aligned}
$$

The scores, $\mathbf{s}_{mi}(\hat{\theta})$, are calculated numerically by a finite difference method. Asymptotic standard errors are obtained by taking the square root of the main diagonal of $M^{-1}\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{B}}\widehat{\mathbf{A}}^{-1}$.

Additional calculations are required to obtain the marginal effect of changes in the covariates on outcome probabilities. As noted by Greene (2008, p.821), several types of marginal effects can be defined for multivariate discrete choice models. The simplest one in the current setting is the marginal effect on the marginal probability that a potential downstream entrant vertically integrates. For continuous covariates, it is defined as

$$\frac{\partial Prob\left(VI = 1 \mid \bar{\mathbf{x}}_1,\, \overline{PF}\right)}{\partial x_{1k}} = \phi\left(\beta_1'\bar{\mathbf{x}}_1 + \alpha\overline{PF}\right)\beta_{1k}, \tag{2.13}$$

where the bar shows that the variables are evaluated at their sample averages or some other representative values. $x_{1k}$ is the $k$th element of $\mathbf{x}_1$, $\beta_{1k}$ is the corresponding element of $\beta_1$, and the market and firm subscripts have been omitted for simplicity. For the dichotomous covariates in $\mathbf{x}_1$, the marginal effect on the marginal probability is calculated as

$$Prob(VI = 1 \mid \bar{\mathbf{x}}_{1,-k},\, x_{1k} = 1) - Prob(VI = 1 \mid \bar{\mathbf{x}}_{1,-k},\, x_{1k} = 0),$$

where $\bar{\mathbf{x}}_{1,-k}$ consists of representative values for the covariates excluding the $k$th one. The marginal effect of $PF$, the paragraph IV indicator, is

$$Prob(VI = 1 \mid \bar{\mathbf{x}}_1,\, PF = 1) - Prob(VI = 1 \mid \bar{\mathbf{x}}_1,\, PF = 0).$$

Another type of marginal effect that is advocated by Greene (1996) relates to the conditional outcome probability. In the current setting, it is defined as the marginal effect of the covariates on the probability of vertical integration by a potential downstream entrant, conditional on the firm having entered the downstream segment and on the paragraph IV status of the market. The expression for this set of marginal effects is quite involved and it is contained in Appendix A.1. The standard errors for both sets of marginal effects are calculated by the delta method, using finite-difference numerical derivatives of the marginal effects with respect to the parameters (Greene, 1996).

## 2.5 Data

The generic drug markets used for analysis are selected from a database of the US Food and Drug Administration (FDA), called the Orange Book, which contains the population of all drug approvals. I begin by selecting a subset of drug markets that opened up to generic competition between January 1, 1993 and December 31, 2005.[45] The set of markets is further narrowed down to

[45] Appendix A.2 explains how generic products are identified in the Orange Book.

those where the relationship between the upstream and downstream segments is relatively straightforward. This is done by first restricting the downstream products to finished formulations containing only one API. When there are multiple single-ingredient formulations containing a given API, I choose only the first of these to open up to generic competition. This is based on the belief that when generic companies make their entry decisions in the first downstream market for a given API, the upstream market structure is not yet formed. Therefore, it makes sense to view downstream and upstream entry decisions as being made simultaneously. By the time the other downstream markets using the same API open up, the upstream market structure may already be fixed. Because it is not realistic to assume that upstream and downstream actions are decided simultaneously in such markets, they are excluded from the analysis.

I also restrict the sample to the following dosage forms which constitute the majority of generic drugs: oral solids, injectables, and topicals. This leaves 177 downstream markets, each defined by a distinct combination of an API and a dosage form. 128 markets remain after removing observations for which market characteristics data could not be obtained. There are 125 corresponding upstream markets, each defined by a distinct API. For three APIs (acyclovir, fluconazole, and gabapentin), two different dosage forms went generic on the same day. In these cases, I consider different dosage forms of the same API to constitute independent markets, and combine each of them with data for their respective API markets. Thus, for the three APIs mentioned above, the same upstream market data are used twice. Table A.1 in the Appendix contains a list of the drugs in the sample. A processed version of the FDA data was obtained from a proprietary database called Newport Sourcing, developed and maintained by Thomson Reuters.

Table 2.1 and Figure 2.1 presented in Section 2.2 are constructed from the dataset of 128 markets. The econometric model is estimated using observations on 85 of those markets that opened up to generic competition between 1999 and 2005. The reason for restricting the time period in this way is as follows. Between 1992 and 1998, the FDA did not grant 180-day generic exclusivity to the first-to-file paragraph IV applicant. Therefore, during this period generic firms had little incentive to develop their products early in order to engage in patent challenges. Thus, the paragraph IV status of a market is likely to have been irrelevant for the decision to vertically integrate. By limiting the sample to the post-1998 period, we can analyze the role of paragraph IV certification more accurately.

### 2.5.1 Entry Indicators and Potential Entrant Status

To record the two firm-level outcomes – downstream entry and vertical integration – it is first necessary to pinpoint the date when each market opens up to generic competition. Previous authors such as Scott Morton (1999) define the market opening date as the approval date of the first ANDA. After comparing ANDA approval dates with the dates when the generic products actually began to be marketed, I find that this definition is not always appropriate.[46] In some cases, the first generic product is not marketed until several months after its ANDA is approved. During those months,

[46]The product marketing dates are obtained from the Newport Sourcing database.

subsequent generic products are not approved by the FDA. I also find a few cases where drugs that appear to be generics are marketed before their ANDAs are approved. The first phenomenon arises when pending patent litigation between the generic entrant and the originator firm, or a settlement between the two, prevent the generic from entering immediately upon ANDA approval. The latter phenomenon is related to a practice called "authorized generics": the originator gives the generic company a license to sell the product based on the former's New Drug Application rather than the latter's ANDA. To accommodate these special cases, I define the market opening date as the first generic approval date or the first generic marketing date, whichever is later.

Firm-level entry actions are defined on the basis of market opening dates. Specifically, a potential downstream entrant is considered to have entered the downstream segment if its ANDA is approved by the FDA either before the market opening date or not later than one year after the market opening date. The relatively narrow window is justified on the grounds that entry timing is an important determinant of profits in generic drug markets; because prices fall rapidly in response to additional entry, most firms enter in the first few months after market opening (Caves et al., 1991; Reiffen and Ward, 2005). As for actions in the upstream segment, a downstream entrant is deemed to have vertically integrated if it submits a Drug Master File (DMF) to the FDA before the market opening date or no later than one year after the market opening date.

I identify a potential downstream entrant in market $m$ as a firm who has entered the downstream segment of any other generic market, including one outside the sample, on a date that is earlier than market $m$'s opening date but that is no more than five years before that date. Thus, I allow a firm to remain a potential downstream entrant for five years after its last entry. Similarly, a firm is identified as a potential upstream entrant of market $m$ if it has entered the upstream segment of another generic market prior to, but not more than seven years before, market $m$'s opening date. Therefore, potential entrant status in the upstream segment is allowed to last for seven years after the last entry event. The reason for setting a wider window for potential upstream entrants is that DMF submissions sometimes occur a few years before the market opening date. Firm $i$ is a potential upstream-only entrant in market $m$ if it is a potential upstream entrant but not a potential downstream entrant.

To evaluate the potential entrant status of a given firm, it is necessary to accurately identify its previous entries. This requires correct names for the ANDA applicants and DMF holders contained in the FDA data. Similarly, identifying firms' vertical integration actions, which involves matching the firms found in the downstream ANDA database with those in the upstream DMF database, requires accurate data on firm names. These tasks are complicated by the several mergers and acquisitions that took place in the generics industry during the observation period. As described in Appendix A.2, I use the Newport Sourcing database to attach accurate firm names to the FDA data. Changes in firm ownership are taken into account by assuming that the past entry experience of an acquired firm is fully carried over to the acquiring firm.

Table 2.5 presents the distribution of actual entry actions taken by potential downstream entrants in the dataset. The data consist of 92 firms facing 2,539 choice situations spread across 85 markets. 406 of these choice situations (15.99 percent) result in downstream entry. 76 of the

Table 2.5: Distribution of Entry Actions in Dataset

| | | Vertical Integration | |
|---|---|---|---|
| | | Not Integrate | Integrate |
| Downstream Entry | Not Enter | 2,133 | 0 |
| | Enter | 330 | 76 |

Notes:
The table shows the distribution of outcomes observed at the firm level. The dataset contains 2,539 firm-level observations from 85 markets that opened up to generic competition between 1999 and 2005.

downstream entries (18.72 percent) lead further to vertical integration.

### 2.5.2 Covariates

#### Market Characteristics

Table 2.6 presents summary statistics for the covariates. The first fourteen variables are market characteristics. "User Population" is a measure of market size, which is expected to have a positive impact on a firm's probability of downstream entry (Bresnahan and Reiss, 1991b). However, its impact on a firm's propensity to vertically integrate is an open question: while Stigler (1951) hypothesizes that vertical integration would occur less frequently in larger markets, others note that under certain conditions, the incidence of vertical integration may actually rise with market size.[47] The user population variable is defined as the estimated number of users of each drug in the US during the period immediately before generic entry. It is constructed from results of the National Ambulatory Medical Care Survey (NAMCS) and the National Hospital Ambulatory Medical Care Survey (NHAMCS). These surveys are conducted by the National Center for Health Statistics (NCHS) to assess the use of ambulatory medical care in the US, through questionnaires sent to randomly selected hospitals and physicians' offices. One part of the survey asks for information on "drug visits" during a fixed reference period. A drug visit occurs when a patient visits a health care facility and a drug is prescribed. I estimate the total number of drug visits in the US for

[47] Perry and Groff (1988), Elberfeld (2002), and Dufeu (2004) indicate that larger markets may be characterized by more vertical integration if entry does not increase in proportion to market size.

Table 2.6: Summary Statistics for Covariates

| Variable Name | Unit | Mean | Min. | Max. |
|---|---|---|---|---|
| *Market Characteristics* | | | | |
| User Population | 1 million people | 2.566 | 0.022 | 18.127 |
| Per-User Expenditure | 1,000 US dollars | 0.979 | 0.018 | 10.726 |
| Anti-infective | Dummy | 0.235 | | |
| Cardiovascular | Dummy | 0.247 | | |
| Central Nervous System | Dummy | 0.200 | | |
| Gastrointestinal / Endocrine-Metabolic | Dummy | 0.141 | | |
| Oncology | Dummy | 0.082 | | |
| Other Therapeutic Class | Dummy | 0.094 | | |
| Oral Solid | Dummy | 0.824 | | |
| Injectable | Dummy | 0.129 | | |
| Topical | Dummy | 0.047 | | |
| Paragraph IV (*PF*) | Dummy | 0.447 | | |
| Upstream Originator Patents | Count | 3.353 | 0 | 24 |
| Downstream Originator Patents | Count | 3.506 | 0 | 24 |
| *Firm Characteristics* | | | | |
| Own Upstream Experience | Count (depreciated) | 8.870 | 0 | 71.423 |
| Own Downstream Experience | Count (depreciated) | 8.407 | 0.616 | 55.162 |
| *Potential Entrants' Characteristics* | | | | |
| Potential Downstream Entrants' Mean Upstream Experience [a] | Count (depreciated) | 7.964 | 5.198 | 19.461 |
| Number of Potential Upstream-Only Entrants | Count | 53.499 | 41 | 73 |
| Number of Potential Downstream Entrants [b] | Count | 35.973 | 6 | 42 |

Notes:
The data consist of 2,539 firm-level observations in 85 markets.

[a] In firm-level equations, the mean experience level of potential downstream entrants excluding firm *i* is used to construct the value for firm *i*.

[b] In firm-level equations, firm *i* is not counted when constructing the value for firm *i*.

each drug in the sample for every year between 1992 and 2004, based on the number of drug visits recorded by the surveys.[48] Then, the total number of drug visits during the one- to five-year period before generic market opening is used to represent the size of the user population for each drug market.[49] Because the focus of NAMCS/NHAMCS is on outpatient services, drugs that are primarily used in inpatient settings (e.g., anesthetics) are not captured by the surveys. Such drugs are therefore excluded from the sample. The average user population for the drugs in the sample is 2.57 million people.

"Per-User Expenditure" is a measure of patients' and insurers' willingness to pay for a drug product. Willingness-to-pay varies greatly across drug products because medical conditions (illnesses and injuries) vary in terms of morbidity and mortality for the patient, while pharmaceuticals vary in their effectiveness at preventing or treating those conditions as well as in the number of available substitutes (e.g., different drugs that treat the same condition). Such variation may influence generic companies' incentive to enter a market because it is likely to affect the number of firms that can profitably enter. As a proxy for the willingness to pay for a drug, I use the per-user average annual expenditure on the drug, including out-of-pocket expenses as well as payments made by insurers and other payers, during the year immediately prior to generic entry. This is estimated in two steps. In the first, the average consumed quantity per user is estimated for each drug using data from the Medical Expenditure Panel Survey (MEPS). Co-sponsored by the Agency for Healthcare Research Quality and the NCHS, MEPS is a nationwide survey that collects data on households' use of medical goods and services, supplemented with information from the respondents' health care providers and pharmacies. Using MEPS data for the period 1996-2005, I calculate the average quantity of each drug consumed by a user in one year. Instead of producing separate values for each year, ten years' worth of observations are pooled together to generate

[48]The NAMCS/NHAMCS data identify drugs only by their APIs and not their dose forms. Therefore, drug visits are counted for each API. Because the reference period for collecting drug visit information is relatively short (one week for NAMCS and four weeks for NHAMCS), I assume that each drug visit represents a unique patient. Sampling weights provided by the NCHS are used when adding up drug visits across different facilities. Detailed information on the surveys is available at `http://www.cdc.gov/nchs/ahcd.htm`.

[49]Due to sampling error, drug visit estimates based on a small number of records in the NAMCS/NHAMCS data tend to be inaccurate. According to Hsiao (2010), the reliability of the estimates can be raised by pooling together multiple years to yield a sufficiently large number of records. Thus, the following steps are taken to generate the user population for drug product $m$ whose generic market opens up in year $t$. First, I construct the following estimates of total drug visits at the national level, using different numbers of years up to $t-1$:

$$TotVisit_{mt,\tau} = Pop_{t-1}\frac{\sum_{s=t-\tau}^{t-1}\sum_h \omega_{hs} Visit_{mhs}}{\sum_{s=t-\tau}^{t-1} Pop_s}, \qquad \tau = 1,2,3,4,5.$$

The subscript $s$ indexes year and $h$ indexes health care facility. $\omega_{hs}$ is the sampling weight for facility $h$ in year $s$, $Visit_{mhs}$ is the number of unweighted drug visits recorded for drug product $m$ at facility $h$ in year $s$, and $Pop_s$ is the US civilian non-institutionalized population in year $s$. Then, the value of the user population variable is chosen as $TotVisit_{mt,\underline{\tau}}$ where $\underline{\tau} = \min \tau \;\; s.t. \;\; \sum_{s=t-\tau}^{t-1}\sum_h Visit_{mhs} \geq V_\tau$. In words, the value of $\tau$, the number of years used for generating the data, is raised until the cumulative number of unweighted drug visit occurrences reaches a prespecified threshold. The threshold value $V_\tau$ is set at 25 for $\tau \in \{1,2\}$, 20 for $\tau \in \{3,4\}$, and 17 for $\tau = 5$.

one figure for each drug to cover the entire observation period.[50] In the second step, the average wholesale price of each drug in the year immediately before generic market opening is obtained from different editions of the *Red Book*.[51] The per-user consumed quantity (rescaled to pricing units) is then multiplied by the average wholesale price to generate the average per-user annual expenditure. The mean of this variable for the drugs in the sample is 979 US dollars.

The drugs in the sample are grouped into six broad therapeutic classes: anti-infectives, cardiovascular agents, central nervous system agents, gastrointestinal and endocrine-metabolic agents (endocrine-metabolic agents include anti-diabetic drugs), oncology drugs, and others.[52] The first three categories each make up between one-fifth and one-quarter of the markets in the sample. The drugs are also classified into three distinct dose form groups: oral solids, injectables, and topicals. Oral solids, which make up 82.4 percent of the in-sample drugs, consist of tablets and capsules including extended-release and other enhanced versions. Injectables are liquids that are usually contained in vials and ampoules. Topicals include creams, lotions, and gels.

There are two reasons for including indicators for therapeutic classes and dose form groups as covariates. First, they are expected to capture unobserved factors that are related to the revenue potential and product development costs for each market, and that may affect generic entry behavior. For instance, patients may be more willing to switch from originator products to generics in certain therapeutic classes than in others. Second, technological economies due to vertical integration may be stronger for certain drug types than for others. For instance, the production of injectables is subject to quality and manufacturing standards that are generally more stringent than the ones for oral solids (Surendar, 2009). Thus, the returns to vertical integration, which enables tighter control over manufacturing processes, may be higher for injectables.

The remaining market characteristics pertain to paragraph IV patent challenges. The paragraph IV indicator variable is equal to one if the market experiences paragraph IV certification by one or more ANDA applicants, and zero otherwise. This information is available from the FDA's website.[53] To construct the two patent-related variables, I obtain a list of patents from the Newport Sourcing database for each drug in the sample. Using this information in conjunction with data on drug approvals and marketing, I identify the originator firms for each drug. Specifically, a firm is identified as an originator of a drug if it fulfills one or more of the following criteria: (i) the firm holds a constraining patent for the drug,[54] (ii) the firm holds the earliest product patent

[50] For many of the drugs in the sample, the number of users contained in a single year's MEPS data is too small to serve as a basis for estimation. By pooling observations from ten years, it is possible to obtain more accurate estimates. The procedure relies on the assumption that per-user consumed quantity does not vary greatly over time. Details of the MEPS data are available at `http://www.meps.ahrq.gov/mepsweb/`.

[51] The *Red Book* is a standard reference for drug prices. During the 1992-2004 period for which data were obtained, it was published by the Medical Economics Company, Thomson Medical Economics, and Thomson PDR.

[52] The therapeutic class of each drug was obtained from Thomson Reuters' Micromedex database.

[53] A list of drugs that have been subject to paragraph IV certification is posted at `http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/ucm047676.htm`.

[54] Constraining patents are defined in the Newport Sourcing database as those that are difficult to circumvent and are likely to prevent generic firms from entering.

(likely to be the basic product patent) for the drug, (iii) the firm is the applicant of the first New Drug Approval for the drug, (iv) the firm is the first marketer of the drug in the US, UK, France, Germany, or Japan. The "upstream originator patents" variable for market $m$ is constructed as the number of product patents and process patents that belong to one of originators of product $m$ and that cover the API used in the product. In addition, the application dates of the patents must be earlier than the generic market opening date, because otherwise they will not affect generic entry. The "downstream originator patents" variable is similarly constructed by counting the number of formulation patents and new use patents that cover product $m$, that belong to its originators, and whose application dates are earlier than the market opening date. The mean number of upstream originator patents in the sample markets is 3.353, and the mean number of downstream patents is 3.506.

**Firm Characteristics**

Following Scott Morton (1999) and Gallant et al. (2008), firm characteristics are generated from the same data source used to generate entry indicators and to determine the potential entrant status of firms. Specifically, a firm's past entry history is used to construct its experience variable for both the upstream and downstream segments. The value of firm $i$'s upstream experience variable for market $m$ is constructed from the firm's DMF submissions during the seven-year period leading up to the market opening date of market $m$. Let $\Upsilon_{i,-m} = \{d_1, d_2, ...\}$ be the sequence of firm $i$'s DMF submission dates, excluding its submission for market $m$. The number of firm $i$'s DMF submissions during the $s$th year prior to market $m$'s market opening date is

$$DMF_{is,-m} = \#\{d_l \in \Upsilon_{i,-m} : \ d_m^o - d_l \in [365(s-1),\ 365s]\},$$

where $s$ is a positive integer, $d_m^o$ is the market opening date for $m$, and $\#(\cdot)$ is a function that counts the number of elements in a set. The "own upstream experience" variable for firm $i$ in market $m$ is then constructed as $\sum_{s=1}^{7} \delta_U^{s-1} DMF_{is,-m}$, where $\delta_U \in [0,1]$ is the depreciation factor for upstream experience.

For constructing the downstream experience variable, the drug product's dose form type is taken into consideration. Suppose that drug $m$ is an oral solid formulation. Then, firm $i$'s downstream experience variable is constructed from its ANDA approvals for oral solid formulations, excluding the one for the current market, during the five-year period leading up to the market opening date. Let $F$ index dose form types, and let $\Delta_{i,-m}^F = \{d_1, d_2, ...\}$ be the sequence of firm $i$'s ANDA approvals for type $F$ dose forms, excluding its approval for market $m$. During the $s$th year prior to the opening of market $m$, firm $i$ obtains the following number of ANDA approvals for dose form type $F$:

$$ANDA_{is,-m}^F = \#\{d_l \in \Delta_{i,-m}^F : \ d_m^o - d_l \in [365(s-1),\ 365s]\}.$$

The "own downstream experience" variable for firm $i$ in market $m$, which is of dose form type $F$, is constructed as $\sum_{s=1}^{5} \delta_D^{s-1} ANDA_{is,-m}^F$, where $\delta_D \in [0,1]$ represents the depreciation factor for downstream experience.

I refer to Gallant et al. (2008) for the value of the depreciation factors. Using data from generic drug markets that opened up during 1990-1994, they estimate a dynamic model in which a firm's entry into one market reduces its cost of entering future markets. Under the simplifying assumption that generic markets open up sequentially in fixed intervals of 1.5 months, Gallant et al. (2008) estimate that fixed entry costs have a persistence parameter of 0.985.[55] In other words, 98.5 percent of the stock of cost reductions realized through past entry is carried over from one market opening to the next. Here, I assume that the depreciation factor of entry experience is equal to the rate of persistence of costs. Therefore, the depreciation factor over a one year interval is calculated to be $0.985^{12/1.5} = 0.886$. I set $\delta_U = \delta_D = 0.886$ and use it to construct the experience variables.

The mean of the own upstream experience variable is 8.870 and that for the own downstream experience variable is 8.407. While the means are similar and both are positively skewed, the upstream experience variable has a higher variance and is more highly skewed. This suggests that firms are more strongly differentiated in terms of their vertical integration capabilities than in terms of their downstream entry capabilities. The mean upstream experience level among all potential downstream entrants, calculated separately for each market, has a sample mean of 7.964. The firm-level counterpart of this variable is the mean upstream experience level among rivals. For firm $i$, it is calculated as the mean upstream experience level among the following set of potential downstream entrants: $\{i' \in \mathfrak{P}_{Dm} : i' \neq i\}$.

The last two covariates in Table 2.6 count the number of potential entrants in each market. The mean number of potential upstream-only entrants (53.499) is greater than that of potential downstream entrants (35.973).[56] This is partly a reflection of the higher degree of globalization in the upstream API industry, which in turn may be due to stricter demands for product quality – both from the FDA as well as consumers – in the downstream finished formulation segment. When drug manufacturers from developing countries such as India first enter the generics markets of the US and other developed countries, they find it easier to enter the upstream segment than the downstream segment (Lanjouw, 1998; Chaudhuri, 2005). As a result, the generic API industry is characterized by a larger number of firms that are more geographically dispersed than in the generic formulation industry.

## 2.6 Results

Table 2.7 presents an informal measure of goodness-of-fit for our estimates of the trivariate probit model. For each observation (firm-market pair) in the dataset, the alternative with the highest predicted probability is identified. These "highest probability alternatives" are then tabulated according to the alternative that is actually observed.[57] The percentage figures in the diagonal

[55] See the first column of Table 2 in Gallant et al. (2008).

[56] The number of potential downstream entrants, when used as a covariate in the firm-level equations, counts potential downstream entrants $i' \neq i$.

[57] For expositional purposes, the five possible alternatives in the trivariate probit model are reduced to three by aggregating across the two paragraph IV outcomes.

Table 2.7: Comparing Observed and Predicted Alternatives

| | | Highest Probability Alternative | | |
|---|---|---|---|---|
| | | *NDE* | *DE/NVI* | *DE/VI* |
| Observed Alternative | No Downstream Entry (*NDE*) | 2,093 (98.12%) | 32 (1.50%) | 8 (0.38%) |
| | Downstream Entry / No Vertical Integration (*DE/NVI*) | 266 (80.61%) | 61 (18.48%) | 3 (0.91%) |
| | Downstream Entry / Vertical Integration (*DE/VI*) | 62 (81.58%) | 5 (6.58%) | 9 (11.84%) |

Notes:

The rows represent observed alternatives and the columns represent alternatives with the highest predicted probability. Each cell shows the number of observations in the dataset that have the observed alternative specified by the row and the highest probability alternative specified by the column. The percentages add up to 100 across columns.

cells represent the proportion of observations for which the highest probability alternative and the observed alternative are the same – in other words, they represent the "percent correctly predicted" (Train, 2009).

Alternatives other than "No Downstream Entry" appear to be under-predicted by the model. If we just look at the two alternatives corresponding to "Downstream Entry" ("No Vertical Integration" and "Vertical Integration"), the model predictions seem fit the observed patterns fairly well. Each alternative has a higher frequency of being the highest probability alternative when it is also the observed alternative. For instance, "Downstream Entry / No Vertical Integration" is the highest probability alternative for 18.48 percent of the observations where it is also the observed alternative. This frequency falls to 1.50 percent and 6.58 percent when the observed alternative is "No Downstream Entry" and "Downstream Entry / Vertical Integration", respectively.

Table 2.8 presents the coefficient estimates for the trivariate probit model and Table 2.9 presents the corresponding marginal effects. The marginal effects are evaluated at representative values of the covariates. For a market characteristic variable that is continuous, the simple average across markets is used. The representative value of a continuous firm characteristic variable $x_k$

is obtained as the sample average of the mean among potential downstream entrants in a market: $\bar{x_k} = \frac{1}{M}\sum_m \left(\frac{1}{\#(\mathfrak{P}_{Dm})}\sum_{i\in\mathfrak{P}_{Dm}} x_{mik}\right)$. For the two variables that are defined differently at the firm level and at the market level – namely, the mean upstream experience of potential downstream entrants and the number of potential downstream entrants – the sample average of the market-level variable is used as the representative value. Therefore, the mean upstream experience of all potential downstream entrants, averaged across markets, is plugged into the firm-level equations as well as the market-level equation. Similarly, the average number of potential downstream entrants is used in all of the equations.[58]

The dichotomous variables are given values that are most commonly observed in the data. With regard to therapeutic class, the cardiovascular category is chosen as the baseline for measuring marginal effects and the dummy variables for the remaining classes are set to zero. Accordingly, the coefficient on each therapeutic class dummy is recalculated so that it measures the difference between that category and the cardiovascular category.[59] Similarly, the oral solid dose form group is chosen as the baseline and the dummy variables for injectables and topicals are set to zero. The most common market opening year in the data is 2002. Therefore, 2002 is chosen as the baseline year and dummy variables for the other years are set to zero (and their coefficients adjusted accordingly). Finally, the paragraph IV indicator variable is set to zero.

The predicted probabilities evaluated at representative values of the covariates are as follows: the marginal probability of vertical integration, $Prob\left(VI = 1 \mid \bar{\mathbf{x}}_1, \overline{PF}\right)$, is 3.17 percent; the conditional probability $Prob\left(VI = 1 \mid DE = 1, PF = 0, \bar{\mathbf{x}}\right)$ is equal to 27.02 percent. The marginal effects in Table 2.9 are divided by these probabilities. Therefore, they represent changes in the outcome probability as a proportion to the predicted probability for the representative observation.

The bottom of Table 2.8 presents estimates for the correlation coefficients $\rho_{12}$ and $\rho_{13}$. In practice, the inverse hyperbolic tangent of these parameters are estimated and transformed back to their original values.[60] $\rho_{12}$ is estimated to be significantly positive with a large absolute value, indicating that $\varepsilon_1$ and $\varepsilon_2$, the error terms in the vertical integration and downstream entry equations, are strongly correlated. Thus, firms with a higher unobserved propensity for downstream entry tend to have higher unobserved returns from vertical integration. On the other hand, the estimate for $\rho_{13}$ is negative with a smaller absolute value and a lower significance level. The negative correlation between $\varepsilon_1$ and $\varepsilon_3$ suggests that firms' unobserved returns from vertical integration are somewhat lower in markets that are more likely, in unobserved ways, to be the target of paragraph

[58] In other words, the values of these two variables for the representative firm-level observation are constructed *without* excluding the firm from the calculation. The reason for doing so is that when calculating their marginal effects on the conditional outcome probability, these variables need to move together inside the three equations.

[59] During parameter estimation, the coefficient on a therapeutic class dummy is defined to measure the difference between that category and the "Other Therapeutic Class" category.

[60] The inverse hyperbolic tangent of $\rho$, also known as Fisher's $z$ transformation, is defined as $\text{arctanh}(\rho) = \frac{1}{2}\ln\frac{1+\rho}{1-\rho}$. This transformation has the benefit of lying on the real number line while $\rho$ is confined to the interval $[-1,1]$. As a result, the transformation is simpler to estimate than $\rho$ itself and its standard error is more easily obtained. Standard errors for the $\rho$ parameters can be obtained from the standard errors of their transformations using the delta method. However, the practice is not advisable because the standard errors thus obtained may imply confidence intervals that go outside the $[-1,1]$ interval.

Table 2.8: Parameter Estimates of Trivariate Probit Model

| | Dependent Variable | | |
|---|---|---|---|
| | Paragraph IV (*PF*) | Downstream Entry (*DE*) | Vertical Integration (*VI*) |
| User Population | 0.011 | 0.047 * | 0.077 *** |
| | (0.069) | (0.027) | (0.018) |
| Per-User Expenditure | 0.134 | 0.019 | 0.005 |
| | (0.235) | (0.049) | (0.051) |
| Anti-infective [a] | −0.528 | 0.115 | 0.03 |
| | (1.567) | (0.291) | (0.301) |
| Cardiovascular | 0.07 | 0.398 | 0.147 |
| | (1.479) | (0.301) | (0.298) |
| Central Nervous System | 1.176 | 0.730 ** | 0.052 |
| | (1.507) | (0.318) | (0.305) |
| Gastrointestinal / Endocrine-Metabolic | 0.14 | 0.405 | −0.425 |
| | (1.498) | (0.284) | (0.320) |
| Oncology | 2.783 | 0.530 * | 0.104 |
| | (1.979) | (0.307) | (0.478) |
| Injectable [b] | −0.858 | −0.508 | 1.173 *** |
| | (2.070) | (0.379) | (0.308) |
| Topical | | −0.371 | 1.175 |
| | | (0.676) | (0.976) |
| Paragraph IV (*PF*) | | | 0.424 ** |
| | | | (0.214) |
| Upstream Originator Patents | −0.153 | −0.003 | |
| | (0.133) | (0.018) | |
| Downstream Originator Patents | 0.268 ** | −0.013 | |
| | (0.108) | (0.016) | |
| Own Upstream Experience | | 0.004 | 0.045 *** |
| | | (0.004) | (0.006) |
| Own Downstream Experience | | 0.056 *** | |
| | | (0.004) | |
| Potential Downstream Entrants' (Rivals') Mean Upstream Experience [c] | 0.032 | 0.04 | 0.171 * |
| | (0.396) | (0.068 ) | (0.094) |

*(Table continued on next page.)*

*(Continued from previous page.)*

| | Dependent Variable | | |
|---|---|---|---|
| | Paragraph IV (*PF*) | Downstream Entry (*DE*) | Vertical Integration (*VI*) |
| # Potential Upstream-Only Entrants | 0.29 | 0.086 * | −0.091 *** |
| | (0.192) | (0.046) | (0.034) |
| # Potential Downstream Entrants (Rivals) [d] | 0.203 | 0.006 | |
| | (0.157) | (0.033) | |
| Year 2000 [e] | −0.482 | 0.089 | 0.441 |
| | (0.915) | (0.307) | (0.671) |
| Year 2001 | −1.614 | −0.167 | 1.287 * |
| | (1.110) | (0.291) | (0.672) |
| Year 2002 | −1.065 | −0.247 | 0.883 |
| | (0.982) | (0.291) | (0.655) |
| Year 2003 | −0.960 | −0.474 | 0.467 |
| | (1.057) | (0.319) | (0.596) |
| Year 2004 | −0.025 | −0.084 | −0.181 |
| | (1.602) | (0.408) | (0.619) |
| Year 2005 | −0.168 | 0.269 | −0.694 |
| | (2.458) | (0.626) | (0.875) |
| Constant | −0.630 | −2.190 *** | −4.931 *** |
| | (3.395) | (0.640) | (1.032) |
| $\rho_{12}$ [f] | 0.886 *** | | |
| $\rho_{13}$ [g] | −0.300 * | | |
| Number of observations | 2,539 | (85 markets) | |

Notes:

***, **, and * represent significance at the one percent, five percent, and ten percent levels, respectively. The cluster-adjusted asymptotic standard errors are in parentheses.

[a] The baseline therapeutic class is "Other".

[b] The baseline dose form is "Oral Solid".

[c] For the vertical integration and downstream entry equations, the variable measures the mean upstream experience among rivals (potential downstream entrants other than the firm in question).

[d] For the vertical integration and downstream entry equations, the number of rivals is used.

[e] The baseline year is 1999.

[f] The inverse hyperbolic tangent of $\rho_{12}$ is estimated as 1.404 with a standard error of 0.380.

[g] The inverse hyperbolic tangent of $\rho_{13}$ is estimated as -0.310 with a standard error of 0.171.

Table 2.9: Marginal Effects of Trivariate Probit Model

| | Marginal Effect on: | |
|---|---|---|
| | $Prob(VI = 1)$ | $Prob(VI = 1 \mid DE = 1, PF = 0)$ |
| User Population | 0.164 *** | 0.082 * |
| | (0.049) | (0.046) |
| Per-User Expenditure | 0.011 | 0.014 |
| | (0.112) | (0.099) |
| Anti-infective [a] | −0.224 | 0.063 |
| | (0.376) | (0.449) |
| Central Nervous System | −0.186 | −0.233 |
| | (0.352) | (0.352) |
| Gastrointestinal / Endocrine-Metabolic | −0.743 *** | −0.715 *** |
| | (0.171) | (0.165) |
| Oncology | −0.089 | 0.927 |
| | (0.847) | (1.413) |
| Other Therapeutic Class | −0.275 | 0.272 |
| | (0.789) | (1.093) |
| Injectable [b] | 5.816 * | 1.878 |
| | (3.310) | (1.468) |
| Topical | 5.835 | 1.881 |
| | (9.679) | (1.493) |
| Paragraph IV (*PF*) | 1.279 | |
| | (0.986) | |
| Upstream Originator Patents | | −0.029 |
| | | (0.031) |
| Downstream Originator Patents | | 0.075 |
| | | (0.051) |
| Own Upstream Experience | 0.096 *** | 0.077 *** |
| | (0.014) | (0.018) |
| Own Downstream Experience | | −0.072 *** |
| | | (0.019) |
| Potential Downstream Entrants' Mean Upstream Experience | 0.364 ** | 0.267 * |
| | (0.182) | (0.148) |

*(Table continued on next page.)*

*(Continued from previous page.)*

| | Marginal Effect on: | |
|---|---|---|
| | $Prob(VI = 1)$ | $Prob(VI = 1\|DE = 1, PF = 01)$ |
| # Potential Upstream-Only Entrants | −0.193 ** | −0.214 ** |
| | (0.079) | (0.097) |
| # Potential Downstream Entrants | | 0.036 |
| | | (0.070) |
| Year 1999 [c] | −0.892 *** | −0.871 *** |
| | (0.220) | (0.235) |
| Year 2000 | −0.642 *** | −0.711 *** |
| | (0.225) | (0.152) |
| Year 2001 | 1.199 | 0.576 |
| | (1.186) | (0.556) |
| Year 2003 | −0.617 ** | −0.393 |
| | (0.259) | (0.404) |
| Year 2004 | −0.937 *** | −0.915 *** |
| | (0.075) | (0.094) |
| Year 2005 | −0.989 *** | −0.991 *** |
| | (0.037) | (0.029) |

Notes:

Formulas for the marginal effects of continuous variables are given in (2.13) and (A.1). The marginal effects of dichotomous variables are calculated as the change in the outcome probability as the variable changes from zero to one. The marginal effects are evaluated at representative values of the covariates whose choice is explained in the beginning of Section 2.6. Each marginal effect is divided by the predicted probability for the representative observation so that the figures represent changes in the outcome probability as a proportion of the base probability. The asymptotic standard errors, in parentheses, are obtained by the delta method. ***, **, and * represent significance at the one percent, five percent, and ten percent levels, respectively.

[a] The baseline therapeutic class is “Cardiovascular”.

[b] The baseline dose form is “Oral Solid”.

[c] The baseline year is 2002.

IV certification. This may be because such markets tend to attract a greater number of unintegrated upstream suppliers.

### 2.6.1 Paragraph IV and Downstream Entry Equations

Before turning to the vertical integration equation which is of primary interest, let us consider the other two equations. In the paragraph IV equation, the coefficient on the downstream originator patents variable is significantly positive. Thus, the observation by Grabowski (2004) and Hemphill and Sampat (2010) that patents on new formulations and new uses are more vulnerable to challenge by generic entrants is supported. This finding has interesting implications regarding the effectiveness of such patents as entry barriers. To the extent that formulation patents and new use patents induce more aggressive entry behavior by generic firms – in the form of paragraph IV ANDA filings – they may be ineffective at delaying generic entry. In fact, the existence of vulnerable secondary patents might make a drug market more attractive in the eyes of potential generic entrants because it creates an opportunity for 180-day exclusivity, and may induce more of them to enter.

In the downstream entry equation, the user population variable has a significantly positive coefficient, which agrees with the intuition that larger downstream markets attract more entrants. On the other hand, the coefficient on per-user expenditure is not significantly different from zero. This suggests that downstream generic entrants are attracted more by market size than by the willingness-to-pay of patients and other payers. Two therapeutic classes – central nervous system agents and oncology drugs – have a significantly positive coefficient, which implies that drugs in these classes tend to attract more generic entry than those in the "Other Therapeutic Class" category.

The coefficient on the firm's own downstream experience is positive and highly significant, confirming earlier results by Scott Morton (1999) and Gallant et al. (2008) that past downstream entry experience reduces firms' entry costs in current markets. On the other hand, the coefficient on the own upstream experience variable is not significantly different from zero, which suggests that the effect of upstream experience on downstream entry costs is small.

The number of potential upstream-only entrants has a significantly positive coefficient in the downstream entry equation. This implies that in markets where the number of potential unintegrated API suppliers is large, downstream entrants expect to earn higher payoffs. It may not be obvious why the number of *potential* entrants in the unintegrated upstream category, as opposed to the number of *actual* entrants, affects the expected payoffs of potential downstream entrants. A likely explanation is that when there are many potential unintegrated upstream entrants, downstream firms expect the equilibrium market structure to be characterized by a greater presence of unintegrated upstream suppliers – in other words, a lower degree of vertical integration. The payoffs of downstream entrants would be higher in markets with less vertical integration if such markets have lower API prices – in other words, if foreclosure effects exist.

Meanwhile, the coefficient on the number of potential downstream rivals is not significantly different from zero. Keeping the size and other characteristics of the market fixed, one would

expect an individual firm's entry probability to fall with the number of rivals vying to enter the same market, because the equilibrium number of entrants is not expected to change. Therefore, it comes as somewhat of a surprise that this coefficient is not significantly negative.

### 2.6.2 Vertical Integration Equation

#### Effect of Market Characteristics

In the vertical integration equation, the user population variable has a significantly positive coefficient and its marginal effect on the probability of vertical integration is also positive and significant. An increase in the number of users by one million raises a potential downstream entrant's marginal probability of vertical integration by 16.4 percent. Conditional on the firm entering the downstream segment and on the market not being subject to paragraph IV certification, the same increase in user population raises the probability of vertical integration by 8.2 percent. The finding of a positive relationship between market size and vertical integration, which runs counter to Stigler's (1951) hypothesis that vertical integration is less prevalent in larger markets, is somewhat puzzling. One possible explanation is that unintegrated upstream firms (whose behavior is not the subject of analysis here) are more efficient in the manufacture of APIs than vertically integrated firms. If the equilibrium selection process for the entry game is such that the more efficient API manufacturers are given higher priority in entry, then we are likely to see a higher share of the upstream market being taken up by unintegrated entrants in smaller markets.

Of the therapeutic class dummy variables, the one for gastrointestinal and endocrine-metabolic agents has a significantly negative marginal effect on the probability of vertical integration. This may be because for some drugs belonging to this class (e.g., antacids), tighter control over the upstream manufacturing process through vertical integration is not as important as it is for cardiovascular drugs, the baseline category. The dummy variable for injectable formulations has a significantly positive coefficient, and its marginal effect on the marginal probability of vertical integration is also positive and significant. This is consistent with the notion that control over manufacturing processes is more important for injectables than for oral solids, and that vertical integration enables firms to have better control.[61]

The coefficient on the paragraph IV indicator variable is estimated to be significantly positive. This lends some support to the hypothesis that vertical integration is motivated by the need for early API development when pursuing a paragraph IV patent challenge. However, the marginal effect of the paragraph IV variable on the marginal probability of vertical integration is not significantly different from zero. The inability to detect a significant marginal effect may be due to the following: while paragraph IV patent challenges change the format of the entry game from a

[61]The estimated coefficients on the dummy variables for injectables and topicals have extremely high absolute values. For instance, the marginal probability of vertical integration is shown to be 581.6 percent higher for injectables than for oral solids. This is an artifact of the low predicted probabilities at representative values of the covariates; the marginal probability of vertical integration for a representative oral solid observation is 3.17 percent, while that for an injectable observation is 21.61 percent.

simultaneous-move one to a race to be first, not all firms want to participate in the race. The change in game format raises the benefit of being vertically integrated only for those firms who participate in the race – i.e., those who intend to file paragraph IV certifications. Thus, in order to more accurately capture the impact of the paragraph IV indicator on vertical integration probabilities, one should examine the effect among firms that actually participate in the race. Unfortunately, the data required for such an analysis – firm-level observations on paragraph IV certification – are not available to me at this time.

**Effect of Own and Rival Firm Characteristics**

A firm's past experience at entering the upstream segment of a market has a significantly positive impact on its probability of vertical integration. One additional upstream entry event during the previous year raises the marginal probability of vertical integration by 9.6 percent and increases the conditional probability of vertical integration by 7.7 percent. This finding indicates that the past upstream experience of a potential downstream entrant lowers its cost of vertical integration – that is, past entry experience has a cost-lowering effect in the upstream API segment just as it does in the downstream finished formulation segment.

Meanwhile, the downstream entry experience of a firm has a significantly negative marginal effect on the conditional probability of vertical integration. Since the downstream experience variable appears only in the downstream entry equation (the selection equation), this is entirely attributable to an indirect effect (Greene, 2008, p.822). As the downstream experience variable rises, firms having a low value of $\varepsilon_1$ become more likely to enter the downstream segment (i.e., to be included in the selected group). Because such firms tend to have low values of $\varepsilon_2$ due to the positive correlation between the two error terms, their inclusion into the selected group lowers the conditional probability of vertical integration.

The mean upstream experience among potential downstream entrants (i.e., rivals) has a significantly positive coefficient in the vertical integration equation, and its marginal effect on the probability of vertical integration is also positive and significant. When the mean upstream experience of rivals increases by one unit (equivalent to one upstream entry event during the previous year), the representative firm's marginal probability of vertical integration rises by 36.4 percent and its conditional probability of integration increases by 26.7 percent.

Combined with the earlier result that a downstream entrant's own upstream experience increases its probability of vertical integration by lowering its cost of vertical integration, this finding implies the following: lower vertical integration costs among rivals raises a firm's incentive to vertically integrate. According to the model presented in Section 2.3.1, this implies that firms' vertical integration decisions are strategic complements, which, in the context of a simultaneous-move vertical integration game, is equivalent to the existence of bandwagon effects. Interestingly, increasing the mean upstream experience of rivals by one unit raises a firm's vertical integration probability by more than three times the amount caused by increasing the firm's own upstream experience by one unit. This suggests that the magnitude of bandwagon effects in the generics industry is quite substantial.

The number of potential upstream-only entrants, which was found to affect downstream payoffs positively, has a significantly negative coefficient in the vertical integration equation. The estimated marginal effects also indicate that increasing the number of potential upstream suppliers significantly lowers a firm's probability of vertically integrating. This finding can be interpreted as follows: when the number of potential unintegrated upstream entrants is large so that a lower degree of vertical integration is expected to hold in equilibrium, each downstream entrant has a lower incentive to vertically integrate. This provides additional support to the view that firms' vertical integration decisions are strategic complements.

**Possible Sources of Bandwagon Effects**

The main finding from the econometric analysis is that vertical integration decisions in the generics industry exhibit bandwagon effects: a firm's incentive to vertically integrate is higher if it expects a greater prevalence of vertical integration among its rivals. What could be the cause of such strategic complementarity? One possible explanation is that the strategic complementarity of vertical integration is caused by foreclosure effects in the post-entry market. Imagine a market where the foreclosure effects of vertical integration are severe relative to its efficiency effects. In such a market, an unintegrated downstream entrant earns a low profit when many of its rivals are vertically integrated, but it gains a high incremental profit by choosing to vertically integrate. On the other hand, when few of its rivals are vertically integrated, the firm's incremental profit from integrating is likely to be small. By comparison, when foreclosure effects are weak relative to efficiency effects, the firm's incremental profit from vertical integration is likely to be larger when fewer of its rivals are integrated (Buehler and Schmutzler, 2005).

Another possibility is that firms in the industry learn from others about the benefits of vertical integration, as suggested by Rosengren and Meehan (1994). The performance of a vertical integrated entrant in one market may inform others in the industry about the hitherto unknown benefits of vertical integration, and influence their actions in future markets. The existence of such learning spillovers would cause vertically integrated entry to become more prevalent over time; it would also create correlation between individual firms' probability of vertical integration and their rivals' upstream experience levels. However, while such inter-firm learning effects cannot be ruled out entirely, they are unlikely to be driving the estimated positive impact that rivals' mean upstream experience has on the probability of vertical integration. This is because the year dummy variables in the vertical integration equation are expected to pick up any learning spillover effects that exist. Turning to the marginal effects of the year dummies, we find that the probability of vertical integration was significantly higher in 2001 and 2002. The rising trend during the first half of the observation period is consistent with the existence of learning spillovers. Somewhat puzzling is the decreasing trend during the second half. One possible explanation is that some of the vertically integrated entries in the former period were caused by fad behavior, which declined in importance during the latter period.

## 2.7 Conclusion

The US generic pharmaceutical industry has experienced a wave of vertical integration since the late 1990s. Industry reports suggest that this pattern may be associated with the increase in paragraph IV patent challenges that followed key court decisions in 1998. The 180-day market exclusivity given to the first generic entrant to file a patent challenge has turned the entry process in some generic drug markets into a race to be first; vertical integration may provide an advantage to the participants of the race by promoting investments aimed at the early development of active pharmaceutical ingredients (APIs). Another cause of the vertical merger wave suggested by industry reports is the existence of bandwagon effects: the rising degree of vertical integration in newly opening markets may have motivated firms to become vertically integrated themselves.

This paper employs simple theoretical models to demonstrate the validity of these two explanations and to derive empirical tests. In the context of a simultaneous-move vertical integration game such as the one seen generally in the generics industry, the existence of bandwagon effects is equivalent to the strategic complementarity of vertical integration decisions. The theoretical model in Section 2.3.1 shows that under strategic complementarity, a firm's probability of vertical integration increases as its rivals' cost of integration decreases. This result leads naturally to a simple test of bandwagon effects. The other model, presented in Section 2.3.2, shows that vertical integration enables firms to develop their APIs early during a patent challenge, increasing their chances of winning first-to-file status, when API supply contracts are incomplete and payment terms are determined through *ex post* bargaining. This prediction can be tested by seeing if markets characterized by paragraph IV certification are more likely to attract vertically integrated entrants.

The two tests are applied to data on 85 generic drug markets that opened up during 1999-2005, using a trivariate probit model that accounts for selection and endogeneity. The coefficient estimate for the paragraph IV indicator variable shows that vertical integration probabilities are higher in paragraph IV markets as the theory suggests, but the marginal effect evaluated at representative values of the covariates is not significantly different from zero. Thus, the hypothesis that vertical integration facilitates relationship-specific non-contractible investments is only partially supported by the data.

The past upstream entry experience of a downstream entrant is found to have a significantly positive impact on its probability of vertical integration. This suggests that upstream experience lowers the cost of vertical integration. We also find that the mean upstream experience of rivals has a significantly positive effect on a firm's vertical integration probability. These two results combined indicate that vertical integration decisions are strategic complements – in other words, bandwagon effects are likely to exist.

There are several possible sources of bandwagon effects. One possibility is that vertical integration generates foreclosure effects in the post-entry market, which, according to Buehler and Schmutzler (2005), give rise to the strategic complementarity of vertical integration decisions. There is some empirical evidence to support the existence of foreclosure effects: the number of potential unintegrated upstream entrants has a positive effect on downstream payoffs but its effect on the returns to vertical integration is negative, which suggests that unintegrated downstream en-

trants are better off if the market is less vertically integrated. Another candidate for the source of bandwagon effects is inter-firm learning about the benefits of vertical integration. The marginal effects of the year dummy variables provide some indication of inter-firm informational spillovers. However, learning effects are unlikely to be behind the estimated positive relationship between a firm's probability of vertical integration and its rivals' upstream experience levels.

# Chapter 3

# Inferring the Effects of Vertical Integration from Entry Games

## 3.1 Introduction

The effect of vertical integration on market outcomes such as prices and product quality in the final goods market can be either positive or negative. For instance, an increase in the level of vertical integration can lead to higher prices or lower prices in the downstream market, depending on the underlying demand and cost function parameters (Salinger, 1988; Hendricks and McAfee, 2010). This is because vertical integration has countervailing effects. One is to decrease the integrating firm's costs – for instance, through the elimination of double marginalization or the facilitation of non-contractible investments. Such efficiency effects tend to lead to lower final good prices or higher product quality. Another effect is to foreclose unintegrated rivals' access to upstream suppliers or downstream buyers. Such foreclosure practices often lead to higher prices or lower quality for the final good. Finally, vertical integration can deter or facilitate entry by unintegrated firms, or induce them to become vertically integrated themselves. In other words, vertical integration can affect market outcomes by influencing the market structure formation process. As this discussion suggests, the link between vertical integration and market outcomes is quite complicated. For this reason, modern analyses on the effects of vertical integration tend to be conducted on an industry-by-industry basis.

This paper presents a novel method for empirically examining vertical integration in an individual industry. It is based on a game theoretic model of simultaneous entry into an oligopolistic market consisting of an upstream segment and a downstream segment. The players of the game are potential entrants who can enter into one of the vertical segments or both. After they make entry and investment decisions, competition occurs within the post-entry market structure and profits are realized. Firms' entry decisions are based on their expectations of post-entry profits, which in turn are affected by the entry decisions of others. Put another way, potential entrants form profit expectations according to the vertical market structure they expect in the entry equilibrium, as well as the position they foresee for themselves within that market structure. It is assumed that po-

tential entrants are heterogeneous in observable ways and that the entry game is one of complete information.

The econometric model is designed for application to a dataset consisting of multiple markets where vertical entry patterns are observed. The entry patterns are interpreted as outcomes of the vertical entry game. The object of estimation is the set of firm-level post-entry payoff equations corresponding to three different categories of entry: downstream-only, upstream-only, and vertically integrated. Potential entrants choose the entry category, or action, that yields the highest profit net of entry costs. Each payoff equation contains as arguments variables that describe the actions of other potential entrants. Estimates for the coefficients on these variables provide the main results of the study. They represent rival effects – the effect of upstream, downstream, and vertically integrated rival entry on profits. While such estimates provide direct measures of inter-firm effects, they can also be used as indirect evidence on the effect of vertical integration on market outcomes.

Like Chapter 2, the dataset used in this chapter comes from the US generic pharmaceutical industry. It covers multiple markets, each defined by a distinct pharmaceutical product. The upstream segment of each market supplies the active pharmaceutical ingredient (API) while the downstream segment processes the API into finished formulations such as tablets and injectables. For each market, we observe multiple firms entering the two vertical segments – some of them into both segments – when patents and other exclusivities that protect the original product expire and generic entry becomes possible.

From the estimated parameters of the vertical entry game, I find that vertical integration between a pair of firms has a significantly positive effect on independent downstream rivals. This suggests that vertical integration has a substantial efficiency effect that spills over to other firms in the downstream segment. Another finding is that in markets containing two upstream units and one downstream unit, backward integration by the downstream monopolist significantly reduces the profit of the unintegrated upstream firm. This is consistent with the existence of efficiency effects due to vertical integration; the independent upstream firm's profit falls if it must contend with a tougher rival.

The parameter estimates are used to simulate the effect of a hypothetical policy that bans vertically integrated entry. I find that while the ban tends to increase the number of upstream entrants, it tends to reduce the number of downstream entrants. Even though competition in the upstream segment is increased as a result, the lower efficiency of unintegrated suppliers or the existence of double marginalization problems leads to less entry in the downstream segment. This suggests that vertical integration has an entry-promoting effect in the generic drug industry. We cannot observe the effect of the policy on other market outcomes such as prices. However, the finding that vertical integration has significant efficiency effects as well as entry-promoting effects leads us to conclude that banning vertically integrated entry has an adverse effect on market performance.

The remainder of the chapter is structured as follows. Section 3.2 explains how this study fits into the empirical industrial organization literature on vertical integration and that on market entry. To my knowledge, this is the first empirical paper to exploit an entry game structure in

order to analyze the effects of vertical integration. In Section 3.3, I describe the process of vertical market structure formation in the generic pharmaceutical industry. The section also discusses the possible motives for, and effects of, vertical integration in this industry. Section 3.4 presents the econometric specification, followed by a description of the estimation strategy in Section 3.5. After describing the data in Section 3.6, I present the estimation results, including the findings from policy simulation, in Section 3.7. A concluding section follows.

## 3.2 Relationship to Previous Studies

### 3.2.1 Competitive Effects of Vertical Integration

This subsection briefly reviews the literature on the competitive effects of vertical integration. After providing separate discussions for the three types of competitive effects identified in the Introduction – efficiency effects, foreclosure effects, and market structure effects – the current study's contribution to this literature is discussed.

#### Efficiency Effects

Vertical integration has three main types of efficiency effects. The first one is the elimination of double marginalization (Spengler, 1950). Double marginalization occurs when oligopolistic markups are charged in both the upstream and downstream segments. By eliminating the markup in the upstream segment, a vertically integrated firm enjoys a cost advantage over its unintegrated downstream rivals. Chipty (2001) finds evidence from the cable TV industry that is consistent with the elimination of double marginalization by vertically integrated firms.

The second type of efficiency effect arises from the ability of vertically integrated firms to carry out higher levels of non-contractible relation-specific investments (Williamson, 1971; Klein et al., 1979; Grossman and Hart, 1986). There is an abundance of empirical research – recent examples of which include Woodruff (2002) and Ciliberto (2006) – indicating the existence of such investment-facilitation effects.

The third type of efficiency effect relates to the ability of vertically integrated firms to secure the supply of an intermediate good or more generally, to improve coordination in logistics. Theoretical models that explore this aspect of vertical integration – namely, Carlton (1979) and Bolton and Whinston (1993) – find that the overall effect on market outcomes is indeterminate. Meanwhile, Hortaçsu and Syverson's (2007) empirical analysis of the cement and ready-made concrete industries finds that vertical integration motivated by logistical concerns has a price-lowering effect.[1]

One issue that has not been addressed in the literature is the possibility that the positive efficiency effects of vertical integration may spill over to other firms that are not vertically integrated.

[1] Hortaçsu and Syverson (2007) find that vertical integration allows firms to allocate the intermediate good more efficiently among downstream units.

Such efficiency spillovers would reinforce the price-lowering or quality-enhancing effects of vertical integration.

**Foreclosure Effects**

Foreclosure typically occurs when a vertically integrated firm restricts supply of the intermediate good with an aim to raise the final good price. The significance of such practices has been the subject of continuing debate; Riordan (2008) summarizes the notable theoretical models that have shaped the discussion. The models are roughly divided into two groups: (i) models where vertical integration raises downstream rivals' costs by dampening competition in the upstream market (e.g., Salinger, 1988; Ordover et al., 1990; Chen, 2001), and (ii) models where vertical integration allows upstream units to restrict the supply of the intermediate good and restore monopoly power (Hart and Tirole, 1990; Rey and Tirole, 2007).

Most of the empirical analysis on vertical foreclosure looks directly at the effect of vertical integration on market outcomes.[2] A general conclusion from this literature is that the effect of vertical integration varies across industries; higher final good prices (or lower product quality) due to vertical integration is found in some industries but not in others. This may be because foreclosure effects exist only in certain industries. It is also possible that in many industries, any foreclosure effects that do exist are offset by the efficiency effects of vertical integration.

Useful experimental evidence on vertical foreclosure exists. Normann (2007) finds that vertically integrated players often employ strategies that raise their rivals' costs, as in Ordover et al. (1990). Similarly, Martin et al. (2001) demonstrate that the monopoly restoration model of Rey and Tirole (2007) is partially supported by experimental data. Thus, the experimental literature provides support for vertical foreclosure theory, not least because efficiency effects are absent by design.

Rosengren and Meehan (1994) and Snyder (1995) look at the effect of vertical integration on rival profits to make inferences about vertical foreclosure. Both papers focus on the effect of a vertical merger announcement on the stock prices of unintegrated rivals. Rosengren and Meehan (1994) do not find that vertical mergers have a significant effect on independent downstream rivals. Thus, they find no support for foreclosure theory. Christopher Snyder's study of the British beer industry, described in Snyder (1995), finds that an independent upstream brewery was harmed by vertical integration between rival breweries and downstream pubs. He interprets this as support for foreclosure theory.[3]

A common feature of the existing empirical work on foreclosure is that they assume exogenous changes in market structure. A defining feature of recent studies such as Hastings and Gilbert (2005) and Suzuki (2008) has been to design dataset construction and estimation methods so that

[2]Recent examples include Hastings and Gilbert (2005), Hortaçsu and Syverson (2007), Ciliberto and Dranove (2006), and Suzuki (2008).

[3]It should be noted that in Ordover et al.'s (1990) model, independent upstream firms *benefit* from rival vertical integration when foreclosure effects are present.

the exogeneity assumption can be made plausible.[4]

### Market Structure Effects

Hart and Tirole's (1990) theoretical paper contains some analysis on the effect of vertical integration on market structure formation. Essentially, the changes in profits brought about by vertical integration may induce unintegrated firms to become integrated themselves or to exit the market. Ordover et al. (1990) also investigates the possibility that vertical integration by one firm may lead another to become vertically integrated.

The possibility that vertical integration can affect the market structure formation process – in other words, that vertical integration exhibits "market structure effects" – is an area that has only recently begun to receive attention from empirical economists. The leading example is Hortaçsu and Syverson (2007). They find that in the cement and ready-made concrete industries, unintegrated upstream firms had higher exit probabilities in markets where a higher proportion of entrants were vertically integrated. This was apparently caused by higher productivity levels among vertically integrated firms. In other words, the efficiency effects of vertical integration may have led unintegrated upstream firms to exit.

### Characteristics of the Vertical Entry Model

The vertical entry model presented in this chapter is designed to estimate the effect of rival actions, including vertical integration, on firm payoffs. For instance, the estimated parameters can be used to calculate how an unintegrated firm's payoff changes when a rival pair consisting of one upstream firm and one downstream firm is replaced by a vertically integrated one. In this sense, the model is closest in spirit to the event studies of Rosengren and Meehan (1994) and Snyder (1995) that look at the effect of rival vertical integration on firm value. The weakness of the Rosengren and Meehan (1994) and Snyder (1995) studies – that the impact of vertical integration on market outcomes is not directly observed – is thus shared by the current model. A major concern is that foreclosure effects and efficiency effects often affect unintegrated firms' profits in the same manner, and thus tend to be indistinguishable (Rey and Tirole, 2007). For example, if an unintegrated downstream firm's profit decreases as a result of rival vertical integration, it could be due to a foreclosure effect, an efficiency effect, or both. Therefore, even if a significant payoff impact is found, one may not be able to conclude anything about the existence of either of these effects.

The advantage of my model is that different types of payoffs can be observed. For a few of the payoff functions, the direction of foreclosure effects is different from that of efficiency effects, so that one is distinguishable from the other. For example, if we find that unintegrated upstream profits

[4]Hastings and Gilbert (2005) use a dataset in which the degree of vertical integration varies exogenously across geographical markets. They exploit a gasoline refinery's acquisition of a downstream retailing chain. The change in the degree of vertical integration caused by the acquisition varies across markets according to the pre-existing market share of the acquired firm. Suzuki (2008) argues for exogeneity by utilizing a propensity score matching estimator.

increase in response to rival vertical integration, the existence of foreclosure effects is implied. This is because efficiency effects can only have a negative impact on an unintegrated upstream firm's payoff. Similarly, if unintegrated downstream profits increases in response to vertical integration, it must be due to the positive spillover of efficiency effects, because any foreclosure effect would affect unintegrated downstream profits negatively.

Another characteristic of the vertical entry model is that, unlike in existing studies such as Hastings and Gilbert (2005) and Suzuki (2008), one need not assume that firms' vertical integration decisions are exogenous. In fact, entire market structures, including the vertical integration status of individual firms, are modeled as endogenous outcomes. This implies that the data requirements for the current model may, in some sense, not be as demanding as that of existing methods. There is, however, a rather stringent requirement that the dataset contain observations from multiple markets where complete vertical market structures are observed.

An additional strength of the vertical entry model lies in its ability to examine how vertical integration influences market structure formation. In addition to asking what happens to an unintegrated firm's payoff when a rival pair becomes integrated, one can ask whether the payoff impact is so large that the firm's entry decision changes. In this connection, a useful application of the model is to evaluate the effect of a policy that bans vertically integrated entry. How does such a ban affect the number of entrants in the upstream and downstream segments? While the answer is not clear *a priori*, the model and parameter estimates can be used to obtain one by simulation.

### 3.2.2 Empirical Analysis of Market Entry

The vertical entry model builds on the large and growing empirical literature on static market entry. This field has been motivated by the technical challenge of how to handle the number of rival entrants – a variable that is clearly endogenous – as a key argument of the firm's payoff function. The earliest studies are Bresnahan and Reiss (1991a,b) and Berry (1992). Building on the pioneering work of Bjorn and Vuong (1984), their econometric models explicitly allow market structure outcomes to be equilibria of entry games. For example, Berry's (1992) model contains an equilibrium finding algorithm that is run at each iteration of the parameter search. These early papers focus exclusively on horizontal competition among firms that produce a homogeneous good. Coefficients on the number-of-rivals variables represent rival effects; from them, information on the degree of competitiveness in the market can be inferred.

More recent papers such as Mazzeo (2002b), Seim (2006), and Orhun (2005) expand the entry model framework to allow for product and spatial differentiation. For instance, in Mazzeo's (2002b) study of motel markets, potential entrants choose between entering the low-quality segment or the high-quality one. His results provide insight not only into the degree of competition within and across different market segments, but also into the process of market structure formation. For instance, the estimated parameters are used to predict how the product-differentiated market structure changes in response to increases in population and traffic. Another group of papers uses the entry model framework to investigate the existence of complementarities in firm actions (Vitorino, 2010; Ellickson and Misra, 2008). For example, Vitorino (2010) finds that the

existence of agglomeration effects allows stores to profit from co-locating inside shopping centers.

The present model examines the formation of vertical market structures in which suppliers and buyers trade and compete. As in the papers on horizontal entry, the estimates provide information on the degree of competition within each vertical segment. In addition, the complementarity between upstream entry and downstream entry can be examined. Finally, and most interestingly, the model should provide evidence on the competitive role of vertically integrated firms. Do vertically integrated firms hurt upstream rivals more than they harm downstream ones? Can some firms benefit from facing a vertically integrated competitor instead of an unintegrated pair of firms? Such questions are empirical in nature, and answering them is the subject of this chapter.

## 3.3 Vertical Market Structure Formation in the Generic Pharmaceutical Industry

This section describes the process of vertical market structure formation in the generic pharmaceutical industry to motivate the econometric model. As described in Chapter 2, drug markets open up to generic competition when patents and data exclusivities that cover the drug expire. In each market, generic drug manufacturers make entry decisions a few years before the market opening date. If an upstream unit decides to enter, it develops the active pharmaceutical ingredient (API) and submits a dossier, called the Drug Master File (DMF), to the Food and Drug Administration (FDA). A downstream entrant, on the other hand, procures the API – either from an outside supplier or from its own production – and develops the finished formulation. It then conducts bioequivalence tests using the finished product and files an Abbreviated New Drug Application (ANDA) to the FDA.

Two peculiar aspects of the generic entry process need to be addressed before providing a stylized description. The first is the possibility of patent challenge by generic entrants. As described at length in Chapter 2, the regulatory rules governing generic entry incentivize generic entrants to challenge the ability of originator patents to block entry, by way of a 180-day generic exclusivity awarded to the first-to-file paragraph IV ANDA applicant.[5] The existence of such incentives pushes firms into a race to be first whenever a paragraph IV patent challenge is involved. The economics of such a race is very different from that of a conventional entry game where firms move simultaneously. For this reason, this chapter focuses only on markets that are not subject to a paragraph IV patent challenge.

The second peculiarity of the entry process is the existence of a matching stage where upstream entrants and downstream entrants form vertical relationships and decide on API trade. While in principle such a stage can be built into the econometric model, the incremental benefit from doing so is unlikely to compensate for the computational difficulty that it entails. Therefore, I assume that matching in the API market takes the following simplified form: every upstream entrant, including

[5]A paragraph IV ANDA is one where the applicant includes a "paragraph IV certification" – a claim that the originator firm's patent is invalid or not infringed.

the upstream plants of vertically integrated entrants, is matched with every downstream entrant including units belonging to vertically integrated firms. In other words, every downstream unit is entitled – from a regulatory point of view – to use the API produced by any of the upstream units.[6]

Once we ignore patent challenges and allow every upstream entrant to be matched with every downstream entrant, the generic entry process can be characterized as a simple simultaneous-move game with discrete actions. Potential entrants choose their actions and receive payoffs according to the oligopoly game played out in the resulting vertical market structure. I do not explicitly model how firms trade the intermediate good and compete against each other in the post-entry vertical oligopoly; the effects of such interactions among firms are summarized into reduced-form payoff equations that have the actions of rival potential entrants as arguments.

Each potential entrant in a market faces an action set consisting of two or four elements. For a firm that is a potential entrant in the downstream segment but not in the upstream segment, the action set is {Not Enter, Unintegrated Downstream Entry}. Similarly, a firm that is a potential upstream entrant but not a potential downstream entrant has the action set {Not Enter, Unintegrated Upstream Entry}. Finally, the choice set of a firm that is a potential entrant in both segments is {Not Enter, Unintegrated Downstream Entry, Unintegrated Upstream Entry, Vertically Integrated Entry}. It is assumed that only pure strategy Nash equilibria of the entry game are played. Therefore, market structure outcomes that are not pure strategy Nash equilibria – for example, one unintegrated downstream entry and no other entries, or one independent upstream entrant and no other entrants – are ruled out. Elberfeld (2002) shows that vertical entry games are characterized by multiple equilibria even if entry decisions in one vertical segment are made prior to those in the other segment. This implies that the existence of multiple equilibria is an unavoidable aspect of simultaneous-move vertical entry games.

Before moving on to the empirical analysis, let us briefly consider the motives for, and effects of, vertical integration in the generics industry. A former executive at Sandoz, one of the largest firms, mentions lower API costs, earlier access to APIs, and stability of supply as the advantages of vertical integration (Stafford, 2006). Others have mentioned the possibility that vertical integration allows better control over the information flow between segments as well as better risk-sharing (Erdei, 2004; Hoffman, 2004), which would presumably lead to higher levels of productive investment. These point to the existence of efficiency effects generated through vertical integration. Such effects are likely to benefit final consumers through lower prices, more reliable supply of drugs, or higher product quality.

On the other hand, recent antitrust cases suggest that vertical integration can generate anticompetitive foreclosure effects. In *FTC v. Mylan et al.* (D.D.C., 1999), the Federal Trade Commission (FTC) claimed that an exclusive dealing contract, signed between a finished formulation manufacturer and its upstream supplier, regarding the APIs for lorazepam and clorazepate tablets (both anti-anxiety agents) contributed to price increases of between 1,900 and 3,200 percent for the

[6] According to the regulatory rules, downstream entrants must state the source of APIs in their ANDA filings. This is done by referencing the serial numbers of their API suppliers' DMFs. The assumption stated here is equivalent to saying that every ANDA references every DMF in a given market.

downstream product.[7]

Unlike exclusive dealing contracts, vertically integrated entry in the generic drug industry is not subject to antitrust scrutiny. As a result, such anecdotal evidence on foreclosure effects is not readily available in the case of vertical integration. There is, however, no reason to assume that vertical integration cannot have similar anticompetitive effects.[8]

## 3.4 Econometric Specification

The basic econometric framework follows that of Berry (1992). Each firm in the dataset is provided with a set of payoff equations corresponding to its possible actions. Given values for the explanatory variables, parameters, and error terms, the payoff equations implicitly define each firm's best response given the actions of the other firms. For every market in the dataset, the system of best responses can be solved for to yield the pure strategy Nash equilibrium of the vertical entry game. The objective is to find the parameter values such that the predicted Nash equilibria of the entry games are as close as possible to observed entry patterns.

### 3.4.1 Post-Entry Payoff Equations

Let $m = 1,2,...,M$ index drug markets as well as the products sold in those markets. Firms are indexed by $i = 1,2,...,I$, and each chooses its strategy $a_{mi} \in \mathcal{J}_{mi} \subseteq \{0,D,U,V\}$ where 0, $D$, $U$ and $V$ are shorthand for "no entry", " independent downstream entry", "independent upstream entry", and "vertically integrated entry", respectively. Market structures are characterized by the number of entrants in each entry category, represented as $\mathbf{N} \equiv \begin{bmatrix} N_D & N_U & N_V \end{bmatrix}^{\mathsf{T}}$, where $N_j = \sum_{i=1}^{I} \mathbf{1}(a_i = j)$, $j \in \{0,D,U,V\}$, and $\mathbf{1}(\cdot)$ is the indicator function.

The object of estimation is the set of firm-level post-entry payoff equations:

$$\pi_j\left(\mathbf{x}_{mi}, \mathbf{N}^{(j)}, \varepsilon_{mj}, \beta, \zeta_m, \alpha_m\right) = h_j(\mathbf{x}_{mi}, \beta, \zeta_m) + g_j\left(\mathbf{N}^{(j)}, \alpha_m\right) + \varepsilon_{mj}, \qquad j \in \mathcal{J}_{mi}. \tag{3.1}$$

[7] See Amended Complaint for Injunctive and Other Equitable Relief, *FTC v. Mylan Laboratories, Inc., et al.* (D.D.C., 1999). For another case of vertical foreclosure through exclusive dealing, see *Geneva and Apothecon v. Barr et al.* (2d Cir., 2004), available at `http://cisgw3.law.pace.edu/cases/020510u1.html`.

[8] Unfortunately, exclusive dealing contracts are rarely observable and my econometric analysis does not take them into account. There is some reason to think that exclusive dealing contracts may have become more costly to implement, relative to vertical integration, after *FTC v Mylan*. As part of a settlement with the FTC, the finished formulation manufacturer Mylan and its API suppliers (Cambrex, Profarmaco, and Gyma Laboratories) agreed, for a five-year period starting in 2000, to notify the FTC before entering into any exclusive dealing agreement with any other firm. The firms were also prohibited from taking part in any exclusive dealing contract whose effect is to "unreasonably restrain trade" and "create an unlawful monopoly". See Order and Stipulated Permanent Injunction, *FTC v. Mylan et al.* (D.D.C., 2000), available at `http://www.ftc.gov/os/2000/11/mylanordandstip.htm`.

$\mathbf{x}_{mi}$ is a row vector of market and firm characteristics (including an intercept) and $\mathbf{N}^{(j)}$ is defined as the realized market structure $\mathbf{N}$ minus firm $i$'s own entry into category $j$. Thus, for example, $\mathbf{N}^{(D)} = \begin{bmatrix} N_D - 1 & N_U & N_V \end{bmatrix}^{\mathsf{T}}$. $h_j(\cdot)$ is a function whose value depends on market and firm characteristics and $g_j(\cdot)$ describes the competitive effects of rival entrants for a firm in category $j$. I assume additive separability between $\mathbf{x}_{mi}$ and $\mathbf{N}^{(j)}$, and the payoff for no entry is normalized to zero.

The random error term $\varepsilon_{mj}$ varies across markets and entry categories, but its value is assumed to be common to all firms within the same market.[9] The reason for making this somewhat restrictive assumption is that the estimation method I employ, called "importance sampling with change-of-variables", does not perform well when the dimensionality of the error term is high.[10] I assume that $\varepsilon_m = (\varepsilon_{m\mathrm{D}}, \varepsilon_{m\mathrm{U}}, \varepsilon_{m\mathrm{V}})$ is normally distributed with covariance matrix

$$\Sigma = \begin{bmatrix} 1 & \rho_{\mathrm{DU}}\sigma_{\mathrm{U}} & \rho_{\mathrm{DV}}\sigma_{\mathrm{V}} \\ \rho_{\mathrm{DU}}\sigma_{\mathrm{U}} & \sigma_{\mathrm{U}}^2 & \rho_{\mathrm{UV}}\sigma_{\mathrm{U}}\sigma_{\mathrm{V}} \\ \rho_{\mathrm{DV}}\sigma_{\mathrm{V}} & \rho_{\mathrm{UV}}\sigma_{\mathrm{U}}\sigma_{\mathrm{V}} & \sigma_{\mathrm{V}}^2 \end{bmatrix}.$$

$\beta$, $\zeta_m$, and $\alpha_m$ are parameter vectors. I allow the latter two to vary randomly across markets. Variability in $\zeta_m$ allows the random component of payoffs to vary across firms. Similarly, by allowing $\alpha_m$ to be random, I allow the random component of payoffs to vary across market structures. Letting the random component of payoffs vary along these dimensions is a requirement of the importance sampling with change-of-variables method.[11]

The covariates in the firm-level payoff equations, $\mathbf{x}_{mi}$, include market and firm characteristics. The market characteristics include the following: market size measured by the number of users of a drug; the willingness-to-pay of patients, insurers, and other payers proxied by average annual per-user expenditure; a dummy variable for drugs belonging to the gastrointestinal and endocrine-metabolic classes; a dummy variable for injectable products; and a dummy variable for drug markets that opened up to generic competition in year 2001 or later.

The first two variables capture the market's revenue potential as perceive by potential generic entrants. The user population variable is constructed as the estimated number of users of each drug in the year before the market opening date – in other words, the user population for the originator product just before generic entry. Since this variable is measured before actual generic entry takes

[9] A similar assumption is made in Mazzeo's (2002b) study of entry in the motel industry.

[10] In the estimation, I use markets as the unit of observation. If each firm is allowed to have a separate value for the $\varepsilon$ term, the dimensionality of integration becomes very large when computing the likelihood contribution of each observation. While such high-dimensional integration can be handled by simulation-based methods (Train, 2009), the use of importance sampling with change-of-variables introduces additional problems. While details of the method and rationale for its use are presented in Section 3.5 and Appendix B.1, a key step in the method involves evaluating the probability density of a simulated draw for the error term vector and taking its ratio to another density value. When the error term vector has high dimensionality, the ratio of densities tends to have an excessively high variance across draws. This renders the estimator unstable and unreliable.

[11] Bajari et al. (2010), who apply the same method, employ an error term that varies across firms and market structures. I employ the random coefficients specification because in my model, an error term that varies across firms and market structures would have unacceptably high dimensionality.

place, it is considered to be exogenous to the entry decisions of generic firms.[12] The per-user expenditure variable is constructed by multiplying the per-user annual consumed quantity of the drug by its per-unit price in the year before generic market opening. Per-user consumed quantity is determined by pharmacological factors, rather than economic ones, and can be considered exogenous to the generic entry decision. The price of the originator drug prior to generic entry is also considered to be exogenous, as it is set in response to market conditions that exist before generics appear.[13]

The three dummy variables reflect the market's revenue potential as well as generic firms' cost of entry and vertical integration. In Chapter 2, we found that the gastrointestinal/endocrine-metabolic variable has a significantly negative impact on a firm's propensity to vertically integrate. The injectables dummy, on the other hand, was found to have a significantly positive impact on the vertical integration decision (see Table 2.9). Inclusion of the post-2000 dummy variable is motivated by the observation, made in Chapter 2, that vertical integration in the generic drug industry increased dramatically after around 2001 (see Figure 2.1).

Following Scott Morton (1999) and Gallant et al. (2008), the firm characteristics that we use relate to the firm's past experience at entering the upstream and downstream segments of other generic drug markets. These variables are expected to capture cross-firm variation in entry costs. I allow a firm's unintegrated downstream payoffs to depend on its downstream experience but not its upstream experience. Similarly, the independent upstream payoff equation has upstream experience as an argument, but not downstream experience. Vertically integrated payoffs are allowed to depend on both downstream and upstream experience.

I allow the random component of payoffs to vary across firms by assuming that the coefficients on the experience variables contained in $\zeta_m$ vary randomly across markets. The $h_j(\cdot)$ component of the payoff equation is then specified as follows:

$$h_j(\mathbf{x}_{mi}, \beta, \zeta_m) = \beta_j' \mathbf{w}_m + \zeta_{mj\mathrm{D}}\, DownExp_{mi} + \zeta_{mj\mathrm{U}}\, UpExp_{mi}, \quad j \in \{\mathrm{D}, \mathrm{U}, \mathrm{V}\}.$$

where $\mathbf{w}_m$ contains the market characteristics, and $DownExp_{mi}$ and $UpExp_{mi}$ are firm $i$'s downstream and upstream experience levels, respectively, when it makes its entry decision in market $m$. It is assumed that $\zeta_{m\mathrm{DU}} = \zeta_{m\mathrm{UD}} = 0$ while the other elements of $\zeta_m$ are independently and normally distributed across markets with a common variance. They can be written out as

$$\zeta_{ms} = \zeta_s + \sigma_\zeta\, \iota_{ms}, \quad s \in \{\mathrm{DD}, \mathrm{UU}, \mathrm{VD}, \mathrm{VU}\}, \tag{3.2}$$

[12] One might argue that unobserved factors that influence user population prior to generic competition, such as the level of advertisement by the originator firm, may be correlated with the unobserved portion of generic firm payoffs. The peculiar nature of pharmaceutical demand allows us to assume away such correlations. In the pharmaceutical industry, consumption decisions are made by prescribing physicians who make their choices based primarily on patients' therapeutic needs. In addition, there are strict restrictions against product-specific advertisements targeted directly at consumers (Schweitzer, 2007).

[13] Originators might use pricing strategies to influence generic entry behavior; setting low prices to discourage entry, for example. Ellison and Ellison (2007) do not find significant evidence of such behavior, however.

where $\iota_{ms}$ is an i.i.d. standard normal random variable.

### 3.4.2 Rival Effects

Estimates for $g_j(\mathbf{N}^{(j)}, \alpha_m)$, the payoff components containing the rival effects, form the main results. Ideally, $g_j(\cdot)$ would have a flexible specification such as the multi-dimensional step function employed by Mazzeo (2002b) and Cohen and Mazzeo (2007). This would allow us to capture, for instance, the difference between the competitive impact of the first rival entrant and that of subsequent ones. Moreover, in models of vertical oligopoly, as in models of product differentiation (e.g., Mazzeo, 2002b), the impact of rival entry into one category is likely to vary according to the presence of entrants in the other categories.

While the step function specification has distinct advantages, it has a stringent data requirement: the market structure outcomes observed in the data must cover the domain of the step function. The number of observations contained in my data is insufficient to provide such variation in market structure outcomes. As a result, trying to estimate the model under a step function specification for $g_j(\cdot)$ results in poor identification (Wooldridge, 2002, pp.345-346): over large regions of the parameter space, the value of the objective function used in estimation fails to respond to changes in some of the step function parameters. I therefore employ the following specification for $g_j(\cdot)$ that is mostly linear in the elements of $\mathbf{N}^{(j)}$:

$$\begin{aligned} g_{\mathrm{D}}\left(\mathbf{N}^{(D)}, \alpha_m\right) = {} & \alpha_{m,\mathrm{DD}} \times \text{number of downstream rivals} \\ & + \alpha_{m,\mathrm{DU}} \times \text{number of upstream entrants beyond first one} \\ & + \alpha_{m,\mathrm{DV1}} \times \text{presence of first vertical entrant (no upstream entrants)} \\ & + \alpha_{m,\mathrm{DV2}} \times \text{number of additional vertical entrants.} \end{aligned} \tag{3.3}$$

$$\begin{aligned} g_{\mathrm{U}}\left(\mathbf{N}^{(U)}, \alpha_m\right) = {} & \alpha_{m,\mathrm{UD}} \times \text{number of downstream entrants beyond first one} \\ & + \alpha_{m,\mathrm{UU}} \times \text{number of upstream rivals} \\ & + \alpha_{m,\mathrm{UV1}} \times \text{presence of first vertical entrant (no downstream entrants)} \\ & + \alpha_{m,\mathrm{UV2}} \times \text{number of additional vertical entrants.} \end{aligned} \tag{3.4}$$

$$\begin{aligned} g_{\mathrm{V}}\left(\mathbf{N}^{(V)}, \alpha_m\right) = {} & \alpha_{m,\mathrm{VD}} \times \text{number of downstream entrants} \\ & + \alpha_{m,\mathrm{VU}} \times \text{number of upstream entrants} \\ & + \alpha_{m,\mathrm{VV}} \times \text{number of vertical rivals.} \end{aligned} \tag{3.5}$$

The nature of the vertical entry process necessitates the slightly peculiar specification. Recall from Section 3.3 that a firm would never enter a market as an independent downstream firm if it expected no upstream suppliers to be present; neither would it choose independent upstream entry in a market with no downstream buyers. That such market structures are ruled out matters for the specification of rival effects. In particular, it affects the choice of the baseline market structure against which other market structures are compared.

For the downstream payoff equation, $\mathbf{N} = (N_D, N_U, N_V) = (1,1,0)$ is chosen as the baseline market structure. This means that in $g_{\mathrm{D}}(\cdot)$, the impact of the first upstream entrant is assigned a value of zero. The coefficients $\alpha_{m\mathrm{DD}}$ and $\alpha_{m\mathrm{DU}}$ represent the incremental impact of additional entrants in the downstream-only and upstream-only categories, respectively. $\alpha_{m\mathrm{DV1}}$ represents the impact of the first vertically integrated entrant when there is no independent upstream supplier. It measures the difference between downstream payoffs under market structure $\mathbf{N} = (1,0,1)$ and that under the baseline market structure. $\alpha_{m\mathrm{DV2}}$ represents the impact of the first vertically integrated entrant when one or more upstream suppliers are present. It also measures the incremental effect of the second and subsequent vertically integrated entrants.

The baseline market structure for the upstream payoff equation is also $\mathbf{N} = (1,1,0)$ so that the impact of the first downstream entrant has a value of zero in $g_{\mathrm{U}}(\cdot)$. $\alpha_{m\mathrm{UD}}$ and $\alpha_{m\mathrm{UU}}$ represent the incremental impact of additional downstream and upstream entrants, respectively. $\alpha_{m\mathrm{UV1}}$ measures the difference between unintegrated upstream payoffs under $\mathbf{N} = (0,1,1)$ and that under the baseline market structure. $\alpha_{m\mathrm{UV2}}$ represents the incremental effect of the second and subsequent vertically integrated rivals, as well as the impact of the first one when one or more unintegrated downstream buyers are present.

For the vertically integrated payoff equation, $\mathbf{N} = (0,0,1)$ is chosen as the baseline market structure; unlike unintegrated entrants, a vertically integrated entrant can earn positive profits as the sole entrant. The coefficients $\alpha_{m\mathrm{VD}}$, $\alpha_{m\mathrm{VU}}$, and $\alpha_{m\mathrm{VV}}$ measure the incremental effects of independent downstream entrants, independent upstream entrants, and vertically integrated rivals, respectively.

The rival effect coefficients are assumed to be normally and independently distributed. I also assume that coefficients belonging to the same equation share a common variance. Thus, they can be expressed as follows:

$$\begin{aligned}
\alpha_{ms} &= \alpha_s + \sigma_{\alpha\mathrm{D}}\,\nu_{ms}, \quad && s \in \{\mathrm{DD},\ \mathrm{DU},\ \mathrm{DV1},\ \mathrm{DV2}\},\\
\alpha_{ms} &= \alpha_s + \sigma_{\alpha\mathrm{U}}\,\nu_{ms}, \quad && s \in \{\mathrm{UD},\ \mathrm{UU},\ \mathrm{UV1},\ \mathrm{UV2}\},\\
\alpha_{ms} &= \alpha_s + \sigma_{\alpha\mathrm{V}}\,\nu_{ms}, \quad && s \in \{\mathrm{VD},\ \mathrm{VU},\ \mathrm{VV}\},
\end{aligned} \tag{3.6}$$

where $\nu_{ms}$ is an i.i.d. standard normal random variable.

**Tests Based on Rival Effects**

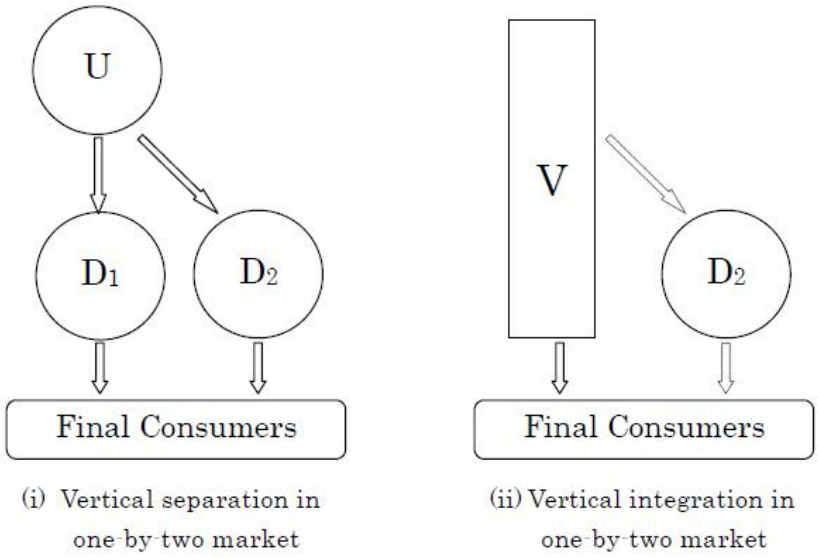
Estimates of the rival effects can be used to test hypotheses regarding the competitive effects of vertical integration. The basic idea is to consider what happens to a firm's payoffs when a pair of other firms in the same market – one that consists of an independent upstream entrant and an independent downstream entrant – is replaced by a vertically integrated firm. In effect, this is equivalent to the impact of a vertical merger within the post-entry market structure. Mergers are common in the generic drug industry and many of them involve a vertical component; some mergers involve firms that are present in the same drug markets. The method discussed here offers a way to assess how such a merger might affect rival profits and outcomes in those markets where both merging parties are present.[14] A more immediate use of this method, however, is as a source of information on how potential entrants perceive the competitive effects of vertical integration – in particular, whether vertical integration generates significant foreclosure effects or efficiency effects.

The first set of tests involves the payoff impact of rival vertical integration on independent downstream firms. In market structures that involve two or more manufacturing units in both vertical segments – which I shall term "large market structures" – this effect is captured by the linear combination $\alpha_{DV2} - (\alpha_{DD} + \alpha_{DU})$. A negative sign on this expression implies that vertical integration by a rival pair is associated with lower independent downstream payoffs. Such an association would be observed if vertical integration had foreclosure effects; higher prices for the intermediate good, caused by vertical foreclosure, would lead to lower unintegrated downstream profits. It is also consistent with certain types of efficiency effects, such as the elimination of double margins; a more efficient rival would cause the unintegrated downstream firm's profit to fall. Therefore, the foreclosure and efficiency effects are indistinguishable in this case.

On the other hand, if $\alpha_{DV2} - (\alpha_{DD} + \alpha_{DU})$ is positive, one can infer that vertical integration has a positive efficiency effect that spills over to other downstream firms who buy the intermediate good from the integrated firm, overwhelming any foreclosure effects that may exist. One example of such an efficiency effect is an increase in the quality of the intermediate good, possibly through higher investment incentives provided to the upstream unit of the vertically integrated entrant. Another example is the stability of supply that may be afforded by vertical integration.

In a market structure involving only one upstream unit and two downstream units – called the "one-by-two market structure" (see Figure 3.1) – the impact of vertical integration on the independent downstream firm can be expressed as $\alpha_{DV1} - \alpha_{DD}$. A negative sign on this impact is consistent with the existence of foreclosure effects of the type described by Rey and Tirole (2007). By vertically integrating, the upstream monopolist is able to reduce the quantity of the intermediate good supplied to the independent downstream firm, thereby raising the market price of the final

[14] When Teva, the industry leader, acquired IVAX, another large generic drug manufacturer, in 2006, the Federal Trade Commission reviewed the possible impact of the merger in multiple markets where the two firms were already present. The review only considered horizontal aspects of the merger (Federal Trade Commission, 2006; Silber, 2007). As Villas-Boas (2007), Hendricks and McAfee (2010), and Manuszak (2010) illustrate, the incorporation of a vertical perspective into the analysis may have yielded additional relevant information.




Notes:

(a) In panel (1), $D_1$ and $D_2$ purchase the intermediate good from the upstream monopolist $U_1$.

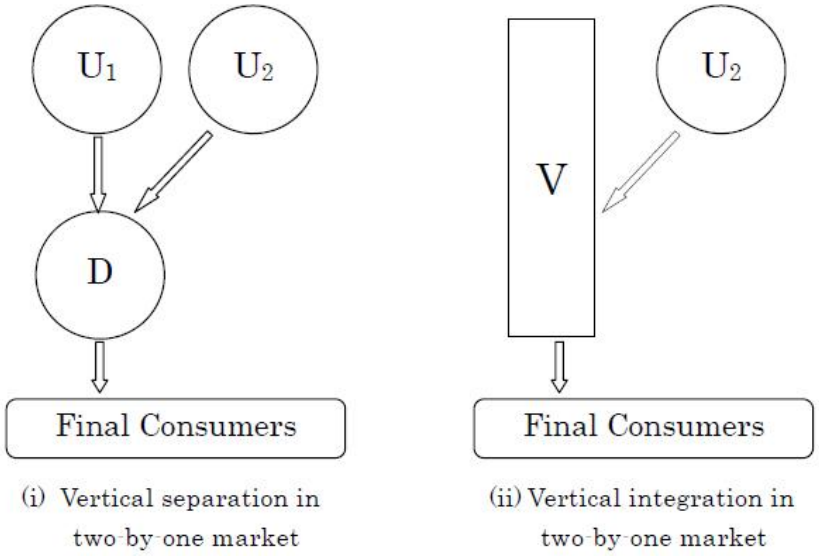
(b) In panel (2), $V$ may or may not sell the intermediate good to $D_2$.

Figure 3.1: One-by-Two Market Structure

good – in other words, the firm is able to "restore monopoly".[15] At the same time, the independent downstream firm can be adversely affected by efficiency effects which make the integrated firm a tougher competitor. Therefore, this test does not allow us to distinguish between foreclosure and efficiency effects (Rey and Tirole, 2007, p.2166). As in the large market structure case, efficiency spillovers are implied if independent downstream firms gain from rival vertical integration under the one-by-two market structure.

The second set of tests concerns the impact of vertical integration on unintegrated upstream payoffs. The effect in large market structures is represented by $\alpha_{UV2} - (\alpha_{UD} + \alpha_{UU})$. A negative sign for this expression, which implies that upstream profits tend to fall in response to rival vertical integration, implies the existence of efficiency effects that make the vertically integrated firm a tougher competitor in the upstream segment. Alternatively, it could reflect a strategic decision on the part of the vertically integrated entrant not to buy the intermediate good from independent

[15] While this test is suggested by Rey and Tirole (2007, p.2165), in their model the entire profit of independent downstream firms is extracted by the upstream firm through nonlinear tariffs and take-it-or-leave-it (TIOLI) bargaining. Therefore, forward integration by the upstream monopolist does not affect independent downstream profits. As Rey and Tirole (2007) suggest, however, relaxing the assumption of TIOLI bargaining is likely to lead to the predictions mentioned here.




Notes:

(a) In panel (1), $U_1$ and $U_2$ sell the intermediate good to the downstream monopsonist $D$.

(b) In panel (2), $V$ may or may not purchase the intermediate good from $U_2$.

Figure 3.2: Two-by-One Market Structure

suppliers. While such an action is unlikely to increase the integrated firm's profit in a static sense, the threat of such action, if credible, may discourage entry into the upstream segment. If, on the other hand, independent upstream profits rise in response to vertical integration, the existence of foreclosure effects is implied. As noted by Ordover et al. (1990), foreclosure by the vertically integrated firm creates more room for unintegrated upstream firms to raise the prices they charge to unintegrated buyers. Thus, it allows upstream firms to gain at the expense of their customers. None of the efficiency effects due to vertical integration are likely to have a positive impact on unintegrated upstream payoffs.

Finally, let us consider the effect of vertical integration on independent upstream profits when the market structure involves two upstream units and one downstream unit (the "two-by-one market structure"; see Figure 3.2). A negative value for $\alpha_{UV1} - \alpha_{UU}$ implies that the upstream firm is adversely affected by its rival's integration. One possible cause of such an impact is the existence of efficiency effects; the upstream unit of a vertically integrated firm may pose a tougher competitor than an unintegrated one. Another possibility is the integrated firm's strategic decision not to buy from the unintegrated supplier. While such an action is unlikely to increase the vertically integrated firm's profit, the threat of such action may be effective at deterring upstream entry. A positive value for $\alpha_{UV1} - \alpha_{UU}$ would present a puzzle; one is hard-pressed to imagine how an unintegrated upstream entrant would be better off if it faced a vertically integrated rival instead of a pair of

unintegrated firms in the two-by-one market structure.

### 3.4.3 Equilibrium Conditions

Given values for $\mathbf{x}_{mi}$, $\varepsilon_{mj}$, $\mathbf{a}_{m,-i}$ (the vector of strategies in market $m$ for firms other than $i$) and the parameters, firm $i$ can figure out its optimal strategy. By the time $i$ observes the strategies of its rivals, however, it is too late for $i$ to choose its own.[16] To allow for simultaneous and optimal decision-making by all potential entrants, I assume that firms play an entry game of complete information where the equilibrium is Nash in pure strategies, and run an equilibrium finding routine within the estimation process. This involves finding equilibrium strategy profiles that satisfy the conditions of Nash equilibrium in each market. Finding all pure strategy Nash equilibrium in generic drug markets is computationally costly, however, because of the large number of players and the sheer size of the strategy profile space $\mathcal{J}_m \equiv \prod_i \mathcal{J}_{mi}$.[17] Thus, instead of searching over strategy profiles, I look for market structures that satisfy the Nash equilibrium conditions. This is the approach taken by Bresnahan and Reiss (1991a), Berry (1992), and others. The estimation process involves finding, for each market, all market structure vectors that are supported in a pure strategy Nash equilibrium of the entry game.

The equilibrium conditions are defined as inequalities in the values of the payoff equations. For illustration, consider the case where we want to check if $\mathbf{N}$ is an equilibrium in market $m$. The set of equilibrium conditions involving firm $i$ is defined as follows, where $\pi_{mij}$ is shorthand for $\pi_j\left(\mathbf{x}_{mi}, \mathbf{N}^{(j)}, \varepsilon_{mj}, \beta, \zeta_m, \alpha_m\right)$:

- If $\mathbf{a}_{m,-i}$ can be summarized as $\mathbf{N}^{(D)} = \begin{bmatrix} N_D - 1 & N_U & N_V \end{bmatrix}^{\mathsf{T}}$, then it must be that $\max_j \pi_{mij} = \pi_{mi\mathrm{D}} > 0$.
- If $\mathbf{a}_{m,-i}$ can be summarized as $\mathbf{N}^{(U)} = \begin{bmatrix} N_D & N_U - 1 & N_V \end{bmatrix}^{\mathsf{T}}$, then it must be that $\max_j \pi_{mij} = \pi_{mi\mathrm{U}} > 0$.

[16] There is a delay between a downstream firm's entry decision and its actual entry in the form of an ANDA approval. This is due to the time required for product development as well as for the FDA's ANDA review. Given that market prices for generic drugs fall with the number of entrants, downstream entrants lose substantially by delaying their entry decisions until rivals' ANDA approvals are observed. We should therefore expect downstream entry decisions to be made more or less simultaneously (Scott Morton, 1999). On the other hand, one cannot rule out the possibility that some potential entrants make their downstream entry decisions after observing the upstream entry actions of their rivals, some of whom may be vertically integrated. This is because DMF submissions, which signify upstream entry, are publicized by the FDA soon after they are received. In this chapter, however, I make the simplifying assumption that downstream and upstream entry decisions are made simultaneously.

[17] Turocy (2008) provides an indication of the lengthy time required for finding all Nash equilibria of large finite games such as the one considered here. His Table 2 shows that the Gambit program (used in econometric applications by Bajari et al., 2010, and others) takes approximately 5.5 minutes to find all equilibria (including pure and mixed strategies Nash equilibria as well as equilibria other than Nash) of a static game consisting of four players whose strategy space consists of three possible actions.

- If $\mathbf{a}_{m,-i}$ can be summarized as $\mathbf{N}^{(V)} = \begin{bmatrix} N_D & N_U & N_V - 1 \end{bmatrix}^{\mathsf{T}}$, then it must be that $\max_j \pi_{mij} = \pi_{mi\mathrm{V}} > 0$.
- If $\mathbf{a}_{m,-i}$ can be summarized as $\mathbf{N}$, then it must be that $\max_j \pi_{mij} \leq 0$.

These equilibrium conditions are checked for all potential entrants by an algorithm embedded into the estimation routine.

## 3.5 Estimation Strategy

Estimation of the model is based on a maximum likelihood framework. At each iteration of the parameter search, I calculate the probability that the market structure observed in each of the sample markets is an equilibrium of the entry game. Loosely speaking, the likelihood contribution of a market is calculated by integrating over the region of the error term vector where the observed market structure is predicted as an equilibrium. Thus, the first tool we need is a mapping from the error term vector to equilibrium market structures. I implement the mapping by programming an equilibrium finding algorithm.

The error term vector in my model has fairly high dimensionality and the region of integration is expected to have a complex shape. As a result, we can expect no closed-form solution to exist for the market structure probability. I therefore approximate the integral using a simulation-based method in which draws of the error term vector are taken (Train, 2009). For each draw, the equilibrium finding algorithm is run to see if the market structure observed in the data is predicted as an equilibrium (i.e., whether the error term vector falls inside the region of integration). The market structure probability is approximated by calculating the proportion of the draws for which the observed market structure is an equilibrium.

In general, vertical entry games are characterized by multiple equilibria (Elberfeld, 2002). Therefore, the equilibrium finding algorithm often indicates more than one market structure as a possible outcome of the entry game in a given market. When such equilibrium multiplicity occurs, an additional assumption is required in order to assign a unique value to the probability that the observed market structure is generated by the model.[18] One possibility is to assume that the equilibrium with the highest joint profits is always realized.(e.g., Jia, 2008). Alternatively, one can specify an equilibrium selection function whose arguments consist of the characteristics of each equilibrium, such as whether or not it is Pareto dominant (Bajari et al., 2010). Following Bjorn and Vuong (1984), I employ the simpler rule that when there are multiple pure strategy equilibria, each market structure has the same probability of occurring.

The entry game for each generic drug market involves a large number of heterogeneous players, which makes the equilibrium finding algorithm time-consuming. Even if one has a fast algorithm, the large number of evaluations that are required during estimation suggests the need to economize

[18] Assigning a unique value to the probability of the observed market structure is not necessary for estimation. In the methods recently developed by Ciliberto and Tamer (2009), Andrews et al. (2004), and Pakes et al. (2006) to estimate discrete games under equilibrium multiplicity, this probability is assigned a range rather than a specific value.

on the number of runs of the algorithm. To this end, I employ the method of importance sampling with change-of-variables (Ackerberg, 2009). The advantage of this method is that the equilibrium finding algorithm needs to be run only once during estimation. In the remainder of this section, I describe the equilibrium finding algorithm, the simulation-based approximation of the market structure probability, and the importance sampling with change-of-variables method.

### 3.5.1 Equilibrium Finding Algorithm

I restrict attention to pure strategy Nash equilibria when searching for the equilibrium market structures. The equilibrium finding algorithm involves plugging in various candidate market structures into each potential entrant's set of payoff equations defined in (3.1), along with the parameter candidate values and a set of values for the error terms, and seeing whether or not the resulting payoffs satisfy the equilibrium conditions spelled out in Subsection 3.4.3.

#### Comparing the Fitted Values for Firm Payoffs

For each candidate market structure vector $\mathbf{N}^{(0)} \equiv \begin{bmatrix} N_D & N_U & N_V \end{bmatrix}^{\mathsf{T}}$, three additional vectors are generated:

$$\mathbf{N}^{(D)} = \begin{bmatrix} N_D - 1 \\ N_U \\ N_V \end{bmatrix}, \mathbf{N}^{(U)} = \begin{bmatrix} N_D \\ N_U - 1 \\ N_V \end{bmatrix}, \mathbf{N}^{(V)} = \begin{bmatrix} N_D \\ N_U \\ N_V - 1 \end{bmatrix}.$$

These four vectors are plugged into each potential entrant's payoff equations. The fitted values are then compared with each other to predict each firm's strategy.

As a demonstration, firm $i$'s fitted payoff values in market $m$ under $\mathbf{N}^{(D)}$, for a particular set of parameter candidate values and a particular set of values for the error terms, are:

$$\pi_j\left(\mathbf{x}_{mi}, \mathbf{N}^{(D)}, \varepsilon_{mj}, \beta, \zeta_m, \alpha_m\right) = h_j\left(\mathbf{x}_{mi}, \beta, \zeta_m\right) + g_j\left(\mathbf{N}^{(D)}, \alpha_m\right) + \varepsilon_{mj}, \qquad j \in \mathcal{J}_{mi}.$$

For brevity, I use the shorthand $\pi_{mij}^{(\kappa)} = \pi_j\left(\mathbf{x}_{mi}, \mathbf{N}^{(\kappa)}, \varepsilon_{mj}, \beta, \zeta_m, \alpha_m\right)$, $\kappa \in \{0, D, U, V\}$. Given the fitted payoff values, firm $i$ finds it profit maximal to be an independent downstream entrant in market $m$ if $\pi_{mi\mathrm{D}}^{(D)} > 0$ and $\pi_{mi\mathrm{D}}^{(D)} = \max_j \pi_{mij}^{(D)}$. Firm $i$ can also maximize profit by deciding to enter as an independent upstream supplier if $\max_j \pi_{mij}^{(U)} = \pi_{mi\mathrm{U}}^{(U)} > 0$. Likewise, vertically integrated entry may be firm $i$'s optimal choice if $\max_j \pi_{mij}^{(V)} = \pi_{mi\mathrm{V}}^{(V)} > 0$. In fact, it is possible that *all* entry categories are profit maximal choices for firm $i$ in an equilibrium characterized by market structure $\mathbf{N}^{(0)}$. In addition, "not entering" is a profit maximal choice for firm $i$ if $\pi_{mij}^{(0)} \leq 0, \ \forall j \in \mathcal{J}_{mi}$.

#### Checking If a Candidate Market Structure Satisfies Equilibrium Conditions

For each market $m$, the following four steps are followed to find out if the candidate market structure $\mathbf{N}^{(0)}$ is an equilibrium.

1. Generate four matrices corresponding to the four vectors $\mathbf{N}^{(\kappa)}, \quad \kappa \in \{0, D, U, V\}$:

$$\Pi_m^{(\kappa)} = \begin{bmatrix} \pi_{m1\mathrm{D}}^{(\kappa)} & \pi_{m1\mathrm{U}}^{(\kappa)} & \pi_{m1\mathrm{V}}^{(\kappa)} \\ \pi_{m2\mathrm{D}}^{(\kappa)} & \pi_{m2\mathrm{U}}^{(\kappa)} & \pi_{m2\mathrm{V}}^{(\kappa)} \\ \vdots & \vdots & \vdots \\ \pi_{mI\mathrm{D}}^{(\kappa)} & \pi_{mI\mathrm{U}}^{(\kappa)} & \pi_{mI\mathrm{V}}^{(\kappa)} \end{bmatrix},$$

where firms are indexed by $i = 1, 2, ..., I$. If firm $i$ is not a potential entrant into category $j$ of market $m$, then $\pi_{mij}^{(\kappa)}$ is automatically set to zero for all $\kappa$.

2. From each of the three matrices $\Pi_m^{(\kappa)}, \ \kappa \in \{D, U, V\}$, create a vector of length $I$ called $\psi_m^{(\kappa)}$. This vector is generated from the $\tau(\kappa)$th column of $\Pi_m^{(\kappa)}$, where

$$\tau(\kappa) = \begin{cases} 1 & \text{if} \quad \kappa = D \\ 2 & \text{if} \quad \kappa = U \\ 3 & \text{if} \quad \kappa = V. \end{cases}$$

The $i$th element of $\psi_m^{(\kappa)}$ is set to zero if $\pi_{mi\kappa}^{(\kappa)} \neq \max_j \pi_{mij}^{(\kappa)}$ or if $\pi_{mi\kappa}^{(\kappa)} < 0$.

3. Merge the three vectors $\psi_m^{(\kappa)}$ column-wise to form the $I \times 3$ matrix $\Psi_m$. It would look something like this:

$$\Psi_m = \begin{bmatrix} \pi_{m1\mathrm{D}}^{(D)} & 0 & 0 \\ 0 & \pi_{m2\mathrm{U}}^{(U)} & \pi_{m2\mathrm{V}}^{(V)} \\ \vdots & \vdots & \vdots \\ 0 & \pi_{mI\mathrm{U}}^{(U)} & 0 \end{bmatrix}$$

This particular example shows that $D$ is a profit maximal action for firm 1, firm 2's optimal actions include $U$ and $V$, and an optimal action for firm $I$ is $U$.

4. The candidate market structure $\mathbf{N}^{(0)}$ is an equilibrium for market $m$ if and only if there exists a strategy profile $\mathbf{a}_m$ with ordered elements $a_{mi}, \ i = 1, 2, ..., I$, such that the following four conditions are satisfied:

(a) $\sum_i \mathbf{1}(a_{mi} = j) = N_j, \quad \forall j \in \{D, U, V\}$.

(b) $\sum_i \mathbf{1}(a_{mi} = 0) = I - \sum_j N_j$.

(c) $\Psi_m(i, \tau(j)) \neq 0$ if $a_{mi} = j, \quad \forall j \in \{D, U, V\}$.

(d) $\Pi_m^{(0)}(i, \tau(j)) \leq 0$ if $a_{mi} = 0, \quad \forall j \in \{D, U, V\}$.

Conditions (a) and (b) simply say that the market structure implied by $\mathbf{a}_m$ must be $\mathbf{N}^{(0)}$. Condition (c) says the following: if $\mathbf{a}_m$ indicates that firm $i$ enters category $j$, then it must be the case that $j$ is the profit maximal action for $i$, given that the actions of its rivals is summarized by $\mathbf{N}^{(j)}$. This condition guarantees that all entrants make positive profits and that they are profit maximizing. Finally, condition (d) ensures that non-entering firms would make a loss if they were to enter. In essence, $\mathbf{a}_m$ is an equilibrium strategy profile characterized by market structure $\mathbf{N}^{(0)}$.

Several different algorithms can be devised for finding $\mathbf{a}_m$. The one I employ chooses equilibrium actions one firm at a time, starting with the element in $\Psi_m$ with the highest value, and moving towards the element with the lowest value. Because I am only interested in checking if candidate market structures are supported in pure strategy Nash equilibrium, it suffices to look for one equilibrium strategy profile for each market structure candidate.

### 3.5.2 Simulation-Based Approximation of Market Structure Probabilities

To find all pure strategy Nash equilibrium market structures, the equilibrium finding algorithm must be run for every possible market structure. I define the market structure space as $\mathfrak{N} = \prod_j \mathfrak{N}_j$, with $\mathfrak{N}_j$ denoting the set of the possible number of entrants in segment $j$. I set the boundaries of $\mathfrak{N}$ based on the market structures observed in the data. Specifically, $\mathfrak{N}_D = \mathfrak{N}_U = \{0, 1, ..., 16\}$ and $\mathfrak{N}_V = \{0, 1, ..., 9\}$.[19]

In most markets, the algorithm identifies multiple market structures as pure strategy Nash equilibria of the entry game. Following Bjorn and Vuong (1984), I assume that each predicted equilibrium has an equal probability of occurring. To apply this method, I must first identify the set of all market structures that are supported as pure strategy equilibria in market $m$ given the data, parameter values, and error term values. This set is defined by the correspondence $\mathscr{E}(\mathbf{x}_m, \varepsilon_m, \beta, \zeta_m, \alpha_m)$, which is manifested in the equilibrium finding algorithm described above. Each element $\mathbf{N} \in \mathscr{E}(\mathbf{x}_m, \varepsilon_m, \beta, \zeta_m, \alpha_m)$ can have multiple equilibrium strategy profiles associated with it. However, because my equilibrium finding algorithm looks for only one equilibrium strategy profile per market structure, I assume that $\mathbf{N}$ maps one-to-one to $\mathbf{a}_m$, the equilibrium strategy profile found by the algorithm. With these assumptions, the predicted probability for the observed market structure in market $m$ is can be written as

$$\begin{aligned} P_m(\theta) = \iint \cdots \int & \lambda(\mathbf{N}_m^o;\ \mathbf{x}_m, \varepsilon_m, \beta, \zeta_m, \alpha_m) \\ & \times f_\varepsilon(\varepsilon_m;\ \sigma_\varepsilon, \rho_\varepsilon)\, f_\zeta(\zeta_m; \zeta, \sigma_\zeta)\, f_\alpha(\alpha_m; \alpha, \sigma_\alpha)\, d\varepsilon_m d\zeta_m d\alpha_m, \end{aligned} \tag{3.7}$$

$$\lambda(\mathbf{N}_m^o;\ \mathbf{x}_m, \varepsilon_m, \beta, \zeta_m, \alpha_m) = \frac{\mathbf{1}\left[\mathbf{N}_m^o \in \mathscr{E}(\mathbf{x}_m, \varepsilon_m, \beta, \zeta_m, \alpha_m)\right]}{\#\left[\mathscr{E}(\mathbf{x}_m, \varepsilon_m, \beta, \zeta_m, \alpha_m)\right]}, \tag{3.8}$$

[19] The maximum number of entrants observed in the data are 13, 11 and 6 for independent downstream entry, independent upstream entry, and vertically integrated entry, respectively (see Table 3.1).

where $\mathbf{N}_m^o$ is the observed market structure in market $m$ and $\#(\cdot)$ counts the number of elements in a set. $\theta$ is the vector of parameters consisting of $\beta$ and the parameters that define the distributions of $\zeta_m$, $\alpha_m$, and $\varepsilon_m$. $f_\varepsilon(\cdot)$ is the trivariate normal density of $\varepsilon_m$ with zero mean vector, and variance parameters $\sigma_\varepsilon = \begin{bmatrix}\sigma_U & \sigma_V\end{bmatrix}$ and $\rho_\varepsilon = \begin{bmatrix}\rho_{DU} & \rho_{DV} & \rho_{UV}\end{bmatrix}$. $f_\zeta(\cdot)$ and $f_\alpha(\cdot)$ are products of normal densities:

$$f_\zeta(\zeta_m;\zeta,\sigma_\zeta) = \prod_{s\in\{DD,UU,VD,VU\}} \frac{1}{\sigma_\zeta}\phi\left(\frac{\zeta_{ms}-\zeta_s}{\sigma_\zeta}\right).$$

$$f_\alpha(\alpha_m;\alpha,\sigma_\alpha) = \prod_{s\in\mathcal{S}_D} \frac{1}{\sigma_{\alpha D}}\phi\left(\frac{\alpha_{ms}-\alpha_s}{\sigma_{\alpha D}}\right) \prod_{s\in\mathcal{S}_U} \frac{1}{\sigma_{\alpha U}}\phi\left(\frac{\alpha_{ms}-\alpha_s}{\sigma_{\alpha U}}\right) \prod_{s\in\mathcal{S}_V} \frac{1}{\sigma_{\alpha V}}\phi\left(\frac{\alpha_{ms}-\alpha_s}{\sigma_{\alpha V}}\right),$$

$$\mathcal{S}_D = \{DD, DU, DV1, DV2\},$$
$$\mathcal{S}_U = \{UD, UU, UV1, UV2\},$$
$$\mathcal{S}_V = \{VD, VU, VV\},$$

where $\phi(\cdot)$ is the standard normal density.

There is no convenient closed-form expression for the integral in (3.7), but the market structure probability can be approximated by using a large number of draws for $\varepsilon_m$, $\zeta_m$, and $\alpha_m$ as follows:

$$P_m^S(\theta) = \frac{1}{R}\sum_{r=1}^{R} \lambda(\mathbf{N}_m^o;\ \mathbf{x}_m, \varepsilon_m^r, \beta, \zeta_m^r, \alpha_m^r). \tag{3.9}$$

The $r$ superscript on $\varepsilon_m$, $\zeta_m$, and $\alpha_m$ indexes draws and $R$ is the number of draws used to approximate the integral.

### 3.5.3 Use of Importance Sampling with Change-of-Variables

The equilibrium finding algorithm described earlier is quite fast; it takes only around 30 minutes to calculate all pure strategy equilibrium market structures for 50 simulated draws pertaining to 85 markets, where each market involves on average 30.64 potential downstream entrants and 75.85 potential upstream entrants (see Table 3.1). Nevertheless, it is prohibitively time-consuming to run the algorithm for every iteration of the parameter search. To deal with this problem, I apply the method of importance sampling with change-of-variables. Introduced by Ackerberg (2009), this technique modifies the econometric model so that the equilibria of the entry game need to be solved only once during estimation, rather than at every iteration of the parameter search. Appendix B.1 provides a concise description of the method.

To apply it to my model, I change the variables of integration from $\{\varepsilon_m, \zeta_m, \alpha_m\}$ to $\{\vartheta_m, \zeta_m, \alpha_m\}$, where the elements of $\vartheta_m$ are defined as follows:

$$\vartheta_{mj} = \beta_j' \mathbf{w}_m + \varepsilon_{mj}, \qquad j \in \{D, U, V\}.$$

$\vartheta_{mj}$ is normally distributed due to the inclusion of $\varepsilon_{mj}$. With this specification, the $r$th draw for the payoff of firm $i$ in segment $j$ of market $m$ under market structure $\mathbf{N}^{(j)}$ is re-written as

$$\check{\pi}_j(DownExp_{mi}, UpExp_{mi}, \mathbf{N}^{(j)}, \vartheta_m^r, \zeta_m^r, \alpha_m^r) = \vartheta_{mj}^r + \zeta_{mjD}^r DownExp_{mi} + \zeta_{mjU}^r UpExp_{mi} + g_j\left(\mathbf{N}^{(j)}, \alpha_m^r\right).$$

These modified payoffs are plugged into the equilibrium finding algorithm to generate the set of equilibrium market structures for the $r$th draw of market $m$. This correspondence is represented by $\check{\mathscr{E}}(DownExp_m, UpExp_m, \vartheta_m^r, \zeta_m^r, \alpha_m^r)$, where $DownExp_m$ and $UpExp_m$ are vectors that contain the experience levels of all potential entrants in market $m$.

Under importance sampling with change-of-variables, market structure probabilities are re-formulated as

$$\begin{aligned} P_m^{\text{IS}}(\theta) = \frac{1}{R}\sum_{r=1}^{R} \; & \check{\lambda}(\mathbf{N}_m^o;\, DownExp_m, UpExp_m, \vartheta_m^r, \zeta_m^r, \alpha_m^r) \\ & \times \frac{f_\vartheta(\vartheta_m^r; \mathbf{w}_m, \beta, \sigma_\varepsilon, \rho_\varepsilon)\;\; f_\zeta(\zeta_m^r; \zeta, \sigma_\zeta)\;\; f_\alpha(\alpha_m^r; \alpha, \sigma_\alpha)}{f^{\text{IS}}(\vartheta_m^r, \zeta_m^r, \alpha_m^r)} \\ & \times \left(\sum_{r=1}^{R} \frac{f_\vartheta(\vartheta_m^r; \mathbf{w}_m, \beta, \sigma_\varepsilon, \rho_\varepsilon)\;\; f_\zeta(\zeta_m^r; \zeta, \sigma_\zeta)\;\; f_\alpha(\alpha_m^r; \alpha, \sigma_\alpha)}{f^{\text{IS}}(\vartheta_m^r, \zeta_m^r, \alpha_m^r)}\right)^{-1}, \end{aligned} \tag{3.10}$$

where $\check{\lambda}(\cdot)$ is the market structure selection function (3.8), modified to take the vectors of firm experience levels, $\vartheta_m$, $\zeta_m$, and $\alpha_m$ as arguments. $f^{\text{IS}}(\cdot)$ is the density for the importance sampling distribution. This distribution is specifically defined *not* to depend on the model parameters. The multiplicative term in the second row is called the importance sampling weight and the last multiplicative term is called the self-normalization factor. The latter term normalizes the importance sampling weights to sum to one across draws.[20]

Estimation by importance sampling involves the following steps. First, $R$ sets of draws for $\vartheta_m$, $\zeta_m$, and $\alpha_m$ are taken from the importance sampling distribution. Next, the equilibrium finding algorithm is run to calculate the market structure selection probability under each of these draws. The third step is the parameter search, which looks for parameter values that maximize

[20] An estimator based on self-normalized importance sampling weights typically has a smaller mean square error than one based on unnormalized weights. This allows the former to perform better (Robert and Casella, 2004). While self-normalization introduces a small bias into the estimates, consistency of the estimator is maintained.

an objective function based on the market structure probabilities (3.10). The values of the draws remain fixed during the search, because the importance sampling distribution does not depend on $\theta$. An important implication is that the equilibrium finding algorithm does not need to be re-run during the parameter search. Intuitively, what happens during the parameter search is that changes in the parameters are "absorbed" by changes in the implied values of the error terms $\varepsilon_m$, $\iota_m$, and $\nu_m$. This is accompanied by changes in the values of the numerator density $f_\vartheta(\vartheta_m^r; \mathbf{w}_m, \beta, \sigma_\varepsilon, \rho_\varepsilon) f_\zeta(\zeta_m^r; \zeta, \sigma_\zeta) f_\alpha(\alpha_m^r; \alpha, \sigma_\alpha)$. Therefore, the importance sampling weights need to be recalculated at every iteration of the parameter search.

The necessity of the random coefficient assumption – namely, that $\zeta_m$ and $\alpha_m$ vary randomly across markets – can be explained at this point. To illustrate, let us suppose that the coefficients on the firm-level experience variables are fixed at $\zeta$ while maintaining the assumption that the value of $\vartheta_m$ is common to all firms. When the value of $\zeta$ is changed during the parameter search, its effect on payoff values cannot be absorbed by simply changing the implied value of $\varepsilon_m$. This is because the effect of changing $\zeta$ differs across firms according to the value of their experience variables. Since the effect of the parameter change cannot be absorbed, the values of the payoff draws must change and the equilibrium finding algorithm must be re-run. By contrast, if $\zeta_m$ is allowed to vary randomly across markets, changes in $\zeta$ can be fully absorbed by implied changes in $\iota_m$ as suggested by (3.2). Similarly, if $\alpha_m$ is allowed to vary randomly across markets, implied changes in $\nu_m$ can absorb changes in $\alpha$ (see (3.6)). Therefore, the random coefficient assumption allows us to keep the payoff draws fixed during the parameter search while maintaining the assumption that $\vartheta_m$ is common to all firms.

The combined number of dimensions in $\{\vartheta_m, \zeta_m, \alpha_m\}$ is $3 + 4 + 11 = 18$, which is not large considering the complexity of the model. As described in Appendix B.1, the small dimensionality of the variable of integration is necessary to make importance sampling with change-of-variables work. Using random coefficients, as suggested by Ackerberg (2009), is a convenient way to keep the dimensionality of the variable of integration low.[21]

A natural choice for the importance sampling distribution is the joint distribution of $\{\vartheta_m, \zeta_m, \alpha_m\}$ when $\theta$ is fixed at some arbitrary value (Ackerberg, 2009). In practice, initial values for the parameters are plugged into the density functions of $\vartheta_m$, $\zeta_m$, and $\alpha_m$ to define the density for the importance sampling distribution as follows:

$$f^{\text{IS}}(\vartheta_m, \zeta_m, \alpha_m) = f_\vartheta(\vartheta_m; \mathbf{w}_m, \beta^\diamond, \sigma_\varepsilon^\diamond, \rho_\varepsilon^\diamond) \;\; f_\zeta(\zeta_m; \chi^\diamond, \sigma_\chi^\diamond) \;\; f_\alpha(\alpha_m; \alpha^\diamond, \sigma_\alpha^\diamond),$$

with the $\diamond$ superscript indicating intial values.

[21] Suppose that, instead of employing the random coefficient structure, I allow the additive error term $\varepsilon$ to vary across markets, firms, entry segments, and market structures as in Bajari et al. (2010). Then, the dimensionality of the variable of integration for a single market observation would be in the hundreds of thousands because of the large number of potential entrants in each market and the abundance of possible market structures. Such a high dimensionality is unacceptable when implementing importance sampling with change-of-variables.

### 3.5.4 Allowing for Misclassification of Market Structure Outcomes

One final issue that needs to be addressed before defining the estimator is the zero likelihood problem. This refers to the possibility that in some markets, none of the simulated draws yield the observed market structure as an equilibrium. The likelihood contribution of such markets is zero, which means that the entire sample likelihood is driven down to zero regardless of the value of other markets' likelihood contributions. The problem is especially severe in regions of the parameter space that are far away from the maximizer of the likelihood, but zero likelihood can occur even in regions near the final estimate (El-Gamal et al., 1993).

A strategy for resolving this issue is to assume that market structure outcomes are subject to misclassification.[22] By doing so, I can avoid the zero likelihood problem even if, for some market $m$, none of the simulated draws yield the observed market structure as a predicted equilibrium. This is because the likelihood contribution of market $m$ is calculated from the equilibrium selection probabilities not only of the observed market structure, but also of other market structures.

Market-level misclassification can be introduced by applying the method proposed by Abrevaya and Hausman (1999). I assume that market structure outcomes are misclassified with positive probability, and that the rate of misclassification depends on the distance, defined on the market structure space $\mathfrak{N}$, between the true market structure and the misclassified market structure. Specifically, I introduce a scalar misclassification parameter $\mu$ that represents the probability that a market structure outcome is misclassified by one entrant.[23] With this assumption, the simulated likelihood contribution of market $m$ subject to misclassification can be defined as

[22] The data generating process suggests three ways by which observed entry patterns may diverge from actual entry patterns, and their common cause is the use of regulatory data from the Food and Drug Administration (FDA). As described in Section 3.6, I define downstream entry as the grant of marketing approval for a finished formulation by the FDA, and upstream entry as the registration of an active pharmaceutical ingredient (API) at the FDA. The first type of misclassification arises when a firm decides not to enter even though it has received downstream approval from the FDA or registered its upstream product with the agency. The second type of misclassification happens when a single marketing approval is shared by multiple downstream firms. The third type of misclassification occurs when an upstream firm sells its API without registering it with the FDA. I assume that, unlike the econometrician, firms in the industry do not rely solely on the FDA's regulatory data to track or predict the activities of their rivals.

[23] This implies, for instance, that

$$Pr\left[\mathbf{N} \quad \text{is misclassified as} \quad \left(\mathbf{N} - \begin{bmatrix}1 & 0 & 0\end{bmatrix}^{\mathsf{T}}\right)\right] = \mu.$$

I also assume that two-entrant misclassifications, such as $\mathbf{N}$ being misclassified as $\left(\mathbf{N} - \begin{bmatrix}2 & 0 & 0\end{bmatrix}^{\mathsf{T}}\right)$ or as $\left(\mathbf{N} + \begin{bmatrix}1 & 1 & 0\end{bmatrix}^{\mathsf{T}}\right)$, each occur with a probability of $\mu^2$. I make analogous assumptions regarding three-entrant misclassifications and beyond, so that $n$-entrant misclassifications have probability $\mu^n$ of occurring. These assumptions imply that while one-entrant misclassifications occur with a constant probability, they can never cancel each other out.

$$P_m^{\text{ISM}}(\theta,\mu) = \frac{1}{R}\sum_{r=1}^{R} \sum_{\mathbf{N}\in\check{\mathscr{E}}(m,r)} \xi(\mathbf{N};\, \mathbf{N}_m^o,\mu)\, \check{\lambda}(\mathbf{N};\, DownExp_m, UpExp_m, \vartheta_m^r, \zeta_m^r, \alpha_m^r)$$

$$\times \frac{f_\vartheta(\vartheta_m^r;\mathbf{w}_m,\beta,\sigma_\varepsilon,\rho_\varepsilon)\;\; f_\zeta(\zeta_m^r;\zeta,\sigma_\zeta)\;\; f_\alpha(\alpha_m^r;\alpha,\sigma_\alpha)}{f^{\text{IS}}(\vartheta_m^r,\zeta_m^r,\alpha_m^r)} \tag{3.11}$$

$$\times \left(\sum_{r=1}^{R} \frac{f_\vartheta(\vartheta_m^r;\mathbf{w}_m,\beta,\sigma_\varepsilon,\rho_\varepsilon)\;\; f_\zeta(\zeta_m^r;\zeta,\sigma_\zeta)\;\; f_\alpha(\alpha_m^r;\alpha,\sigma_\alpha)}{f^{\text{IS}}(\vartheta_m^r,\zeta_m^r,\alpha_m^r)}\right)^{-1},$$

$$\xi(\mathbf{N};\, \mathbf{N}_m^o,\mu) = \begin{cases} \mu^{\left(\sum_j |N_j - N_{mj}^o|\right)} & \text{if} \quad \mathbf{N} \neq \mathbf{N}_m^o \\ 1 - \sum\limits_{\mathbf{N}\in\{\mathfrak{N}\setminus\mathbf{N}_m^o\}} \mu^{\left(\sum_j |N_j - N_{mj}^o|\right)} & \text{if} \quad \mathbf{N} = \mathbf{N}_m^o. \end{cases} \tag{3.12}$$

where $\check{\mathscr{E}}(m,r)$ is shorthand for $\check{\mathscr{E}}(DownExp_m, UpExp_m, \vartheta_m^r, \zeta_m^r, \alpha_m^r)$ and the function $\xi(\cdot)$ represents the probabilities of market structures being correctly or incorrectly classified. The exponential term in brackets in (3.12) represents the size of misclassification. The specification of $\xi(\cdot)$ guarantees that the (mis)classification probabilities add up to one.

In principle, $\mu$ can be estimated along with the other model parameters. In practice, however, I find the parameter search to converge to a high value of $\mu$ (around 0.15, to be specific). Such a high misclassification rate implies a negative value for the probability of correct classification. To avoid this problem, I fix $\mu$ at the arbitrary value of 0.05.[24]

### 3.5.5 Additional Implementation Issues

One problem with the importance sampling with change-of-variables method is that useful expressions for the estimated covariance matrix of the parameters have not yet been developed (Ackerberg, 2009). Moreover, the large amount of time required to generate one set of estimates precludes the calculation of standard errors by bootstrapping. I therefore follow Bajari et al. (2010) in employing Bayesian methodology. Specifically, I define a prior distribution for the parameters and use Markov Chain Monte Carlo (MCMC) to simulate draws from the posterior distribution. While the concepts of prior and posterior distributions are specific to the Bayesian approach, the output of Bayesian methods – namely, the MCMC realizations from the posterior distribution – can

[24] If I am willing to assume that $\xi(\mathbf{N};\, \mathbf{N}_m^o,\mu) = 0$ when $\mathbf{N}$ is outside some small neighborhood of $\mathbf{N}_m^o$, then $\mu$ is free to have a large value and it can be estimated. By thus excluding large regions of the market structure space, however, I weaken the effectiveness of the smoothing strategy.

be useful even in a frequentist setting (Train, 2009; Chernozhukov and Hong, 2003).[25] In particular, draws from the posterior distribution can be used to construct highest posterior density intervals for the parameters which are comparable to confidence intervals in frequentist econometrics.

I implement the method by assuming a flat prior distribution whose density is denoted by $p(\theta)$. In the prior distribution, the elements of $\beta$, $\zeta$, and $\alpha$ are distributed uniformly on the real number line. The variance parameters $\sigma_\varepsilon$, $\sigma_\zeta$, and $\sigma_\alpha$ are distributed uniformly on $\mathbb{R}_+$ and the correlation coefficients are uniformly distributed on the interval $[-1,1]$.[26] The posterior distribution, from which MCMC draws are taken, is proportional to the product of the simulated likelihood of the model and the prior density:

$$p(\theta|\mathbf{Y}) \propto \mathrm{SL}(\theta)p(\theta),$$

$$\mathrm{SL}(\theta) = \prod_m^M P_m^{\mathrm{ISM}}(\theta),$$

where $\mathbf{Y}$ denotes the data including market and firm characteristics as well as market structure outcomes. The procedure for taking MCMC draws from the posterior distribution is explained in Bayesian econometrics textbooks such as Lancaster (2004).

As a final note, the importance sampling with change-of-variables method requires a careful choice of initial values for the parameters. This is because the same set of predicted equilibrium market structures are used throughout the parameter search. During the search, new parameter values do not add or remove market structures from the set of predicted equilibria; they merely assign new importance sampling weights to the simulated draws. The initial values should be chosen so that to begin with, the model predicts the market structures observed in the sample markets to occur as equilibrium outcomes with sufficiently high probability.

To obtain such initial values, I look for a parameter vector that allows the market structure observed in each market to be included in the set of predicted market structures with high frequency. This is obtained through the following maximization problem:

$$\theta^\diamond = \underset{\theta}{\mathrm{argmax}} \prod_m \frac{1}{R}\sum_{r=1}^R \mathbf{1}\{\mathbf{N}_m^o \in \mathscr{E}(\mathbf{x}_m, \varepsilon_m^r, \beta, \zeta_m^r, \alpha_m^r)\}. \tag{3.13}$$

[25] When estimation involves a difficult objective function, Bayesian methods can be faster and easier to implement than optimization methods. This is because numerical Bayesian estimation does not involve maximization or minimization of an objective function and one does not have to worry about getting stuck in one local extremum after another. On the other hand, the MCMC algorithm does require a large number of realizations before its output converges to the posterior distribution.

[26] Even though the prior distribution is improper in the sense that it does not integrate to one, the resulting posterior distribution is proper.

The computational burden of this maximization problem is relatively low, because it is not necessary to find all equilibria of the entry game; one only needs to check if $\mathbf{N}_m^o$ is included in $\mathscr{E}(\mathbf{x}_m, \varepsilon_m^r, \beta, \zeta_m^r, \alpha_m^r)$ for every draw of every market.

## 3.6 Data

The data used in this chapter come from the US generic pharmaceutical industry. The dataset consists of observations from 85 drug markets that opened up to generic competition during the period 1993-2005. This section begins by describing how markets are selected for inclusion in the sample. Descriptions of data source and variable construction are kept brief as most of the information is presented in Chapter 2.

### 3.6.1 Selection of Markets

As mentioned at the beginning of Section 2.5, there are 128 generic drug markets that satisfy the following criteria: (i) the market opened up to generic competition during 1993-2005; (ii) the downstream product is the first one, among all single-ingredient products using the same API, to become generic; (iii) the downstream product is an oral solid, injectable, or topical formulation; and (iv) data on market characteristics are available for the product. These are the markets where we are likely to see upstream and downstream entry decisions being made at around the same time.

43 of the 128 markets are subject to a patent challenge by one or more of the generic entrants. These are identified by the Food and Drug Administration's list of drug markets where one or more Abbreviated New Drug Applications (ANDAs) containing a paragraph IV certification have been filed. As discussed in Section 3.3, the existence of a patent challenge changes the market structure formation process from a simultaneous-move entry game to a race to be the first-to-file entrant. Meanwhile, my econometric model is only designed to estimate the parameters of a simultaneous entry game. Therefore, the 43 markets with paragraph IV certification are dropped from the analysis. This leaves a sample of 85 markets that are not subject to patent challenge by any of the entrants.

The exclusion of paragraph IV markets raises concerns of sample selection. If the incidence of paragraph IV certification is correlated with any of the error terms of the model, the removal of paragraph IV markets from the sample may lead to biased estimates. While acknowledging the importance of such concerns, the estimation conducted in this chapter does not take them into account. The main reason is the difficulty of incorporating selection into the econometric model. As in Chapter 2, sample selection can be modeled by specifying a dichotomous choice process for the determination of paragraph IV status at the market level, and allowing the error term in the paragraph IV equation to be correlated with the remaining error terms of the model.[27] In practice,

[27] In Chapter 2, firm-level choice probabilities are computed by evaluating a low-dimensional multivariate normal cumulative distribution function. Sample selection can be incorporated by increasing the dimensionality of the multivariate normal density by one. Here, we compute market structure selection probabilities by taking draws of a

however, I find that the parameter estimates fail to converge when joint estimation of the vertical entry model and the paragraph IV equation is attempted.[28] Another factor that may allow us to ignore the sample selection problem is that, according to the results in Chapter 2, the error term of the paragraph IV equation is not likely to be strongly correlated with the error terms of the firm-level payoff equations.[29]

### 3.6.2 Variable Construction

#### Entry Indicators and Potential Entrant Status

The definition of downstream and upstream market entry follows that in Chapter 2. Downstream entry into market $m$ is observed when a potential entrant's ANDA for that market is approved by the FDA. Upstream entry is observed when the potential entrant's submission of a DMF for that market is publicized by the FDA. Vertical entry occurs when the potential entrant receives ANDA approval and submits a DMF in the same market.

The definition of potential entrant status also follows from Chapter 2. A firm is a potential downstream entrant of market $m$ if its previous entry into the downstream segment of another market was less than five years before the market opening date of market $m$.[30] Similarly, a firm is a potential upstream entrant if its previous upstream entry in another market was not more than seven years before the market opening date. If a firm is both a potential downstream entrant and a potential upstream entrant, it is considered to be a potential vertically integrated entrant.

Table 3.1 shows summary statistics for the number of potential entrants and the number of actual entrants in each market. The number of potential upstream entrants is greater than the number of downstream entrants. On average, there are 75.859 potential upstream entrants in a market while the average number of potential downstream entrants is 30.635. 18.929 of those entrants, on average, are potential vertically integrated entrants. The actual number of entrants is much smaller: an average market structure contains 3.365 unintegrated downstream entrants, 3.565 unintegrated upstream entrants, and 0.647 vertically integrated entrants.

---

high-dimensional error term vector. In principle, sample selection can be modeled by appending the error term in the paragraph IV equation to the other error terms and taking draws from their joint distribution.

[28] An alternative is to run a two-step procedure where the paragraph IV equation is estimated first. The first-step estimates can be used to adjust the distribution of the error term vector for the vertical entry model when it is estimated in the second step. The two-step method is useful when one can obtain consistent point estimates and methods such as Murphy and Topel (1985) to adjust standard errors are applicable. Unfortunately, I am not aware of any method to correct the highest posterior density intervals obtained from MCMC realizations to account for two-step estimation.

[29] See Table 2.8. The correlation coefficient between the error term in the paragraph IV equation and that in the vertical integration equation, $\varepsilon_{13}$, is found to be significantly negative only at the ten percent level and its absolute value is not high.

[30] The market opening date is defined as the first generic approval date or the first generic marketing date, whichever is later.

Table 3.1: Summary Statistics for the Number of Potential and Actual Entrants

| | Independent downstream | Independent upstream | Vertical |
|---|---|---|---|
| *Number of potential entrants per market* | | | |
| Mean | 30.635 | 75.859 | 18.929 |
| Minimum | 6 | 71 | 4 |
| Maximum | 42 | 80 | 32 |
| *Number of actual entries per market* | | | |
| Mean | 3.365 | 3.565 | 0.647 |
| Minimum | 0 | 0 | 0 |
| Maximum | 13 | 11 | 6 |

Notes:

Counts are based on a sample of 85 pharmaceutical markets that opened up to generic competition between 1993 and 2005, and that were not subject to a paragraph IV patent challenge.

**Covariates**

Table 3.2 presents summary statistics for the covariates in the payoff equations. User population is a measure of market size; it is constructed as the estimated number of users for each drug during the year before generic competition.[31] The average number of users for the drug markets in the sample is 2.12 million. Per-user expenditure is a proxy for the willingness-to-pay of users and other payers.[32] The average annual expenditure for the sample drugs is 642 dollars per year. Following Mazzeo (2002b), these two variables are transformed in the following manner before using in estimation:

$$\tilde{x}_m = \ln\left[\frac{x_m}{(\sum_{n=1}^{M} x_n)/M}\right],$$

[31]This variable is constructed from results of the National Ambulatory Medical Care and National Hospital Ambulatory Medical Care Surveys.

[32]This variable is constructed by multiplying per-user annual usage, estimated from Medical Expenditure Panel Survey data, with per-unit drug prices from the *Red Book*.

Table 3.2: Summary Statistics for Covariates

| Variable Name | Unit | Mean | Min. | Max. |
|---|---|---|---|---|
| *Market Characteristics* | | | | |
| User Population [a] | 1 million people | 2.120 | 0.022 | 18.127 |
| Per-User Expenditure [a] | 1,000 US dollars | 0.642 | 0.018 | 9.511 |
| Gastrointestinal/Endocrine Metabolic | Dummy | 0.153 | | |
| Injectable | Dummy | 0.106 | | |
| Post-2000 | Dummy | 0.447 | | |
| *Firm Characteristics* | | | | |
| Downstream Experience [b] | Count (depreciated) | 2.042 | 0 | 55.162 |
| Upstream Experience [b] | Count (depreciated) | 5.777 | 0 | 71.423 |

Notes:
The data consist of observations from 85 markets that opened up to generic competition between 1993 and 2005, and that were not subject to a paragraph IV patent challenge..

[a] Transformed values of these variables are used in estimation. Let $x_m$ be the level of the untransformed variable for market $m$. Then, the transformed value is $\ln\left[\frac{x_m}{(\sum_{n=1}^{M} x_n)/M}\right]$.

[b] Transformed values of these variables are used in estimation. Let $x_{mi}$ be the level of the untransformed variable for firm $i$ in market $m$. Then, the transformed value is $\ln\left[\frac{x_{mi}}{(\sum_{n=1}^{M}\sum_{l=1}^{I_n} x_{nl})/\sum_{n=1}^{M} I_n}\right]$, where $I_n$ is the number of potential entrants in market $n$.

where $x_m$ is the level of the untransformed variable for market $m$ and $M$ is the number of markets in the sample.[33] 15.3 percent of the drugs in the sample belong to the category of gastrointestinal and endocrine-metabolic agents. Injectable formulations make up 10.6 percent of the downstream markets in the sample, and 44.7 percent of the markets opened up to generic competition after the year 2000.

The experience variables are constructed by adding up depreciated counts of each firm's past ANDA approvals and DMF submissions, as described in Chapter 2. Past entry experience is assumed to depreciate by a factor of 0.886 each year, as implied by the estimates of Gallant et al. (2008). The experience variables are transformed in the following manner:

$$\tilde{x}_{mi} = \ln\left[\frac{x_{mi}}{(\sum_{n=1}^{M}\sum_{l=1}^{I_n} x_{nl})/\sum_{n=1}^{M} I_n}\right],$$

where $x_{mi}$ is the level of the untransformed variable for firm $i$ in market $m$ and $I_n$ is the number of potential entrants in market $n$.

[33] According to Mazzeo (2002b), parameter search is facilitated by thus transforming the variables, both of which are distributed with highly positive skewness.

## 3.7 Results

I generate a Markov chain of 25,000 realizations from the posterior distribution of the parameters. The first 10,000 realizations are discarded because the chain may not have converged to the posterior distribution during the earlier realizations; the last 15,000 are used for analysis. The simulation-based likelihood contribution of each market, represented by (3.11), is calculated by taking 50 draws of the error term vector.[34]

Table 3.3 presents, for each parameter, the 95 percent highest posterior density interval (HPDIs) of the marginal posterior distributions as well as their modes. Each HPDI is constructed by drawing a kernel smoothed density for the marginal posterior distribution, and finding a cutoff level such that the set of points with a density higher than the cutoff constitute 95 percent of the posterior distribution (Lancaster, 2004).

Estimates for the standard deviations $\sigma_U$ and $\sigma_V$ indicate that payoffs are more variable in the unintegrated upstream and vertically integrated categories than in the unintegrated downstream one. It is also found, from the significantly negative estimate for $\rho_{DV}$, that $\varepsilon_{mD}$ and $\varepsilon_{mV}$ are negatively correlated.

### 3.7.1 Exogenous Covariates

The estimated coefficients on the two continuous market characteristics, User Population and Per-User Expenditure, are largely within expectations. User Population has a significantly positive coefficient in the unintegrated downstream payoff equation, while in the other equations its coefficient is not significantly different from zero. The Per-User Expenditure variable is not significant in the unintegrated downstream and vertically integrated payoff equations, but it has a significantly positive coefficient in the unintegrated upstream payoff equation. These results support the notion that larger market size and higher willingness-to-pay raise firms' entry incentives.

The dichotomous market characteristic variables have contrasting effects in the three equations. The Gastrointestinal/Endocrine-Metabolic dummy variable has a significantly positive impact on unintegrated downstream payoffs but its effect on vertically integrated payoffs is significantly negative. This conforms to the finding in Chapter 2 that vertical integration probabilities are lower in markets belonging to the gastrointestinal and endocrine-metabolic classes. It may be because the tighter control over upstream manufacturing processes afforded by vertical integration is less important for such drugs than for other drugs. Another finding that agrees with the results in Chapter 2 is that the Injectable dummy variable has a significantly negative coefficient in the unintegrated downstream and upstream equations, while its coefficient in the vertically integrated equation is significantly positive. This confirms the intuition that tighter manufacturing controls through vertical integration are more important for injectables than for other dosage forms.

[34] Following Train (2009), the draws are generated using a deterministic Halton sequence. In comparison to draws produced by pseudo-random generators, draws taken from a Halton sequence provide better coverage of the support of the error component to be simulated. As a result, fewer draws are needed to achieve a given level of accuracy (Train, 2009).

The Post-2000 dummy variable follows the same pattern as the Injectable dummy: it is significantly negative in the unintegrated downstream and upstream equations while being significantly positive in the vertically integrated payoff equation. One possible explanation comes from the finding in Chapter 2 that vertically integrated entry became more common after the year 2000 (see Figure 2.1). Potential entrants in markets that opened up in the years after 2000 tend to have higher levels of experience at vertically integrated entry, which is likely to lower their costs of entering as an integrated firm. If the effect of past vertical entry experience is not fully captured by the Downstream Experience and Upstream Experience variables – which would be the case if vertically integrated entry generates cost-lowering effects that are not gained from separately entering the downstream and upstream segments of different markets – it is possible for the Post-2000 dummy variable to pick up the unexplained portion of the effect.

The coefficients on the two firm-level experience variables all have the same expected sign. Past downstream entry experience has a significantly positive impact on unintegrated downstream and vertically integrated payoffs. Past upstream entry experience similarly affects the payoffs of unintegrated upstream entrants and vertically integrated entrants in a significantly positive manner.

### 3.7.2 Rival Effects

In all three equations, the effect of incremental rival entry in the same category – represented by $\alpha_{DD}$, $\alpha_{UU}$, and $\alpha_{VV}$ – is significantly negative, as expected; same-type rival entry has negative payoff effects. Another expected result is that unintegrated downstream entrants have a significantly positive effect on independent upstream payoffs, as seen by the HPDI for $\alpha_{UD}$. The presence of more buyers leads to higher profits for independent suppliers. On the other hand, the significantly negative estimate for $\alpha_{DU}$ is an unexpected result for which an explanation cannot readily be found.

The change in unintegrated downstream payoffs as the vertical market structure changes from $\mathbf{N} = (N_D, N_U, N_V) = (1,1,0)$ to $\mathbf{N} = (1,0,1)$, represented by $\alpha_{DV1}$, is significantly negative. This is an expected result. Under the latter structure, the independent downstream firm faces a downstream competitor, whereas it enjoys a downstream monopoly under the former structure. Similarly, the significantly negative estimate for $\alpha_{UV1}$ indicates that unintegrated upstream payoffs decrease when the market structure changes from $(1,1,0)$ to $(0,1,1)$, as expected.

For the remaining rival effect parameters, there is no *a priori* reason to expect a particular sign. We find that unintegrated downstream payoffs increase in response to incremental entry by vertically integrated firms, as implied by the significantly positive estimate for $\alpha_{DV2}$. This is a somewhat unexpected result, but not one without an explanation. As discussed in 3.4.2, it is consistent with the efficiency effects of vertical integration spilling over to benefit unintegrated downstream buyers of the intermediate good. Meanwhile, the significantly negative estimate for $\alpha_{VU}$ suggests that vertically integrated entrants compete with unintegrated entrants in the upstream segment. The remaining rival effects, $\alpha_{UV2}$ and $\alpha_{VD}$, are not significantly different from zero.

Table 3.3: Parameter Estimates

| | | Mode of Marginal Posterior Distribution | 95% Highest Posterior Density Interval |
|---|---|---|---|
| *Independent Downstream Payoff Eq.* | | | |
| Constant | | 1.685 | [ 1.356, 2.222] |
| User Population | | 0.545 | [ 0.247, 0.763] |
| Per-User Expenditure | | −0.025 | [ −0.366, 0.226] |
| Gastrointestinal/Endocrine-Metabolic | | 0.825 | [ 0.519, 1.856] |
| Injectable | | −1.075 | [ −1.815, −0.526] |
| Post-2000 | | −1.655 | [ −2.304, −1.317] |
| Downstream Experience | | 0.975 | [ 0.619, 1.302] |
| *Rival Effects:* | $\alpha_{DD}$ | −1.185 | [ −1.578, −0.871] |
| | $\alpha_{DU}$ | −0.425 | [ −0.841, −0.105] |
| | $\alpha_{DV1}$ | −3.225 | [ −4.211, −0.889] |
| | $\alpha_{DV2}$ | 2.355 | [ 1.755, 3.944] |
| *Independent Upstream Payoff Eq.* | | | |
| Constant | | 3.225 | [ 2.548, 3.638] |
| User Population | | 0.265 | [ −0.034, 0.602] |
| Per-User Expenditure | | 0.775 | [ 0.437, 1.153] |
| Gastrointestinal/Endocrine-Metabolic | | -0.475 | [ −1.991, 0.834] |
| Injectable | | −2.695 | [ −4.117, −1.999] |
| Post-2000 | | −2.635 | [ −3.266, −1.443] |
| Upstream Experience | | 1.335 | [ 1.010, 1.640] |
| *Rival Effects:* | $\alpha_{UD}$ | 0.655 | [ 0.308, 0.901] |
| | $\alpha_{UU}$ | −0.565 | [ −0.924, −0.295] |
| | $\alpha_{UV1}$ | −2.315 | [ −5.538, −1.708] |
| | $\alpha_{UV2}$ | 1.735 | [ −0.806, 2.196] |

*(Table continued on next page.)*

*(Continued from previous page.)*

| | Mode of Marginal Posterior Distribution | 95% Highest Posterior Density Interval |
|---|---|---|
| *Vertical Payoff Eq.* | | |
| Constant | −6.265 | [ −7.680, −5.573] |
| User Population | 0.075 | [ −0.305, 0.671] |
| Per-User Expenditure | 0.325 | [ −0.425, 0.932] |
| Gastrointestinal/Endocrine-Metabolic | −3.765 | [ −6.596, −3.151] |
| Injectable | 2.285 | [ 1.685, 3.650] |
| Post-2000 | 5.045 | [ 3.910, 6.397] |
| Downstream Experience | 0.675 | [ 0.372, 1.065] |
| Upstream Experience | 0.975 | [ 0.728, 1.490] |
| *Rival Effects:* $\alpha_{VD}$ | 0.085 | [ −0.617, 1.418] |
| $\alpha_{VU}$ | −1.085 | [ −2.184, −0.161] |
| $\alpha_{VV}$ | −5.265 | [ −7.135, −4.452] |
| *Variance Parameters* | | |
| $\sigma_U$ | 1.555 | [ 1.488, 1.922] |
| $\sigma_V$ | 2.195 | [ 2.145, 2.367] |
| $\rho_{DU}$ | 0.035 | [ −0.106, 0.129] |
| $\rho_{DV}$ | −0.305 | [ −0.463, −0.162] |
| $\rho_{UV}$ | −0.095 | [ −0.150, 0.316] |
| $\sigma_\zeta$ | 0.505 | [ 0.436, 0.575] |
| $\sigma_{\alpha D}$ | 0.225 | [ 0.117, 0.292] |
| $\sigma_{\alpha U}$ | 0.225 | [ 0.152, 0.297] |
| $\sigma_{\alpha V}$ | 0.365 | [ 0.251, 0.450] |

Notes:
The $\alpha$ parameters representing rival effects are defined in equations (3.3), (3.4), and (3.5).
The posterior modes are found using a grid with steps of 0.01. It is for this reason that the third decimal place is always 5.

### 3.7.3 Competitive Effects of Vertical Integration

Table 3.4 and Figures 3.3 through 3.6 present the HPDIs for specific linear combinations of the rival parameters, as motivated in Section 3.4.2. They describe how the payoffs of unintegrated downstream and upstream entrants are affected when a pair of firms in the same market becomes vertically integrated.

The first row of Table 3.4 and Figure 3.3 show the significantly positive impact that rival verti-

cal integration has on unintegrated downstream payoffs in a large market structure. As discussed in Section 3.4.2, vertical integration has a beneficial effect on independent downstream firms only if positive efficiency effects are present and they overwhelm any foreclosure effects that may exist. The efficiency effects must be quite large in order to generate spillover effects that benefit unintegrated downstream entrants. Therefore, this result strongly supports the existence of efficiency effects through vertical integration. Meanwhile, the second row of 3.4 and Figure 3.4 hint at the possibility of vertical integration having a significantly negative effect on unintegrated downstream payoffs in the one-by-two market structure. The effect, however, is not significantly different from zero.

Rows three and four of Table 3.4 present the effect of rival vertical integration on independent upstream profits. While the effect is not significant in large market structures (Figure 3.5), it is significantly negative in the two-by-one market structure (Figure 3.6). The latter finding is consistent with the existence of efficiency effects; the independent upstream firm's profit falls if it must contend with a tougher rival. Another possible explanation is that the vertically integrated firm forecloses the unintegrated supplier from access to final consumers. The strategy of not buying from the unintegrated supplier – or buying less from it – would be profitable if it deters upstream entry, because independent upstream entry has a negative effect on vertically integrated payoffs. However, it is not clear how the vertically integrated entrant can credibly commit not to buy from the unintegrated supplier.[35] It is more likely that the evidence presented in Figure 3.6 is an indication of efficiency effects.

### 3.7.4 Simulating a Ban on Vertically Integrated Entry

The preceding results indicate that firms' entry actions are consistent with the existence of efficiency effects and the absence of foreclosure effects. This implies that vertical integration in this industry is likely to be procompetitive from a static point of view. Nevertheless, vertical integration can still have an anticompetitive market structure effect. For instance, independent upstream firms may be deterred from entering when they anticipate tough competition from vertically integrated rivals. This, in turn, may reduce the expected profits of independent downstream entrants and lead to fewer entrants in the downstream segment. Given the stylized fact that the prices of generic drugs fall monotonically with the number of downstream entrants (Frank and Salkever, 1997; Reiffen and Ward, 2005), competition authorities are likely to be concerned if vertical integration tends to reduce the equilibrium number of entrants.[36]

On the other hand, it is possible for vertical integration to increase the equilibrium number of

[35] If we explicitly consider post-entry market competition as a separate stage game, we can see that an equilibrium where the vertically integrated entrant tries to deter the entry of an unintegrated upstream supplier by not buying from it is not subgame perfect.

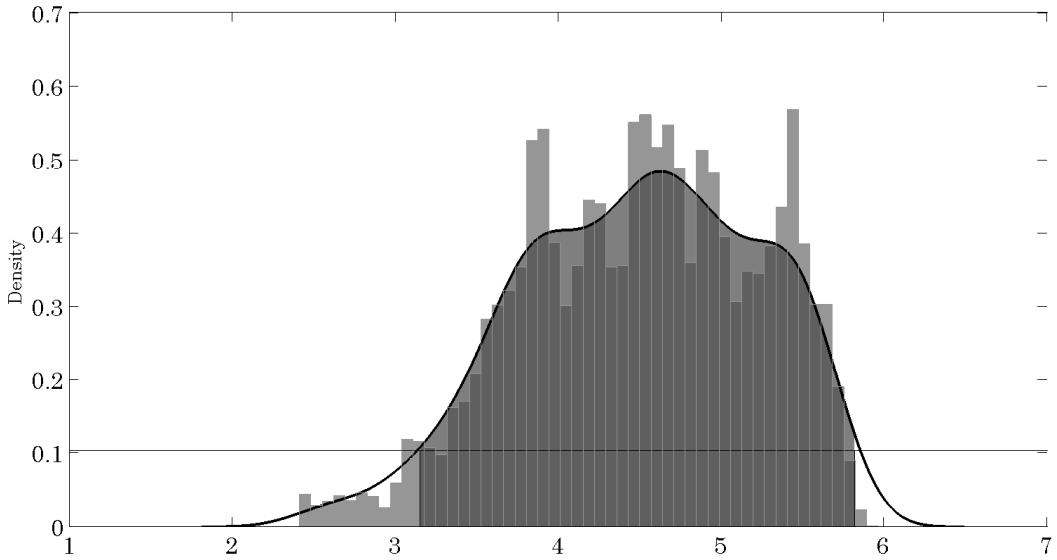
[36] Even if vertical integration leads to fewer downstream entrants, it will not necessarily cause downstream prices to be higher, given its efficiency effects. Therefore, the equilibrium number of entrants *per se* should not be the primary concern of policymakers. In practice, however, any entry-reducing effect is likely to attract the attention of antitrust authorities.

Table 3.4: Payoff Impact of Vertical Integration by a Rival Pair

| | Mode of Marginal Posterior Distribution | 95% Highest Posterior Density Interval |
|---|---|---|
| Effect on Independent Downstream Payoff in Large Market Structure: [a] | | |
| $\alpha_{DV2} - (\alpha_{DD} + \alpha_{DU})$ | 4.485 | [ 3.150, 5.818] |
| Effect on Independent Downstream Payoff in One-by-Two Market Structure: | | |
| $\alpha_{DV1} - \alpha_{DD}$ | $-1.965$ | [ $-2.984$, 0.145] |
| Effect on Independent Upstream Payoff in Large Market Structure: [a] | | |
| $\alpha_{UV2} - (\alpha_{UD} + \alpha_{UU})$ | 1.605 | [ $-0.846$, 2.274] |
| Effect on Independent Upstream Payoff in Two-by-One Market Structure: | | |
| $\alpha_{UV1} - \alpha_{UU}$ | $-1.635$ | [ $-4.879$, $-0.990$] |

Notes:
The $\alpha$ parameters are defined in equations (3.3), (3.4), and (3.5).

[a] "Large Market Structures" are those that have two or more entrants (including units of vertically integrated entrants) in both the upstream and downstream segments.



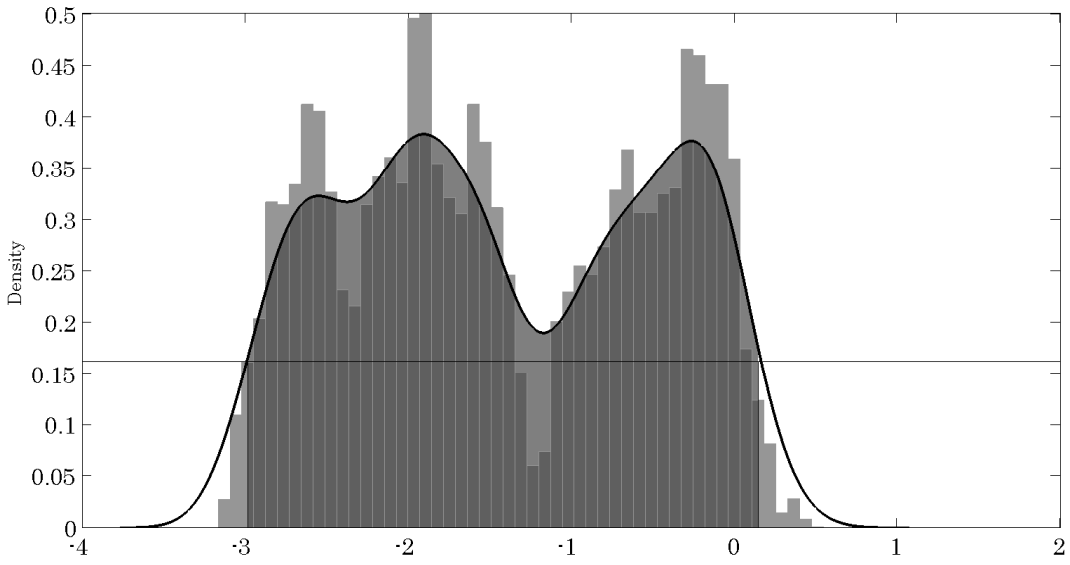

Notes:

(a) The bars form a density histogram of MCMC realizations for the impact of rival vertical integration on independent downstream payoffs in market structures with two or more entrants in both the upstream and downstream segments.

(b) The smoothed density is generated using a Gaussian kernel with bandwidth of 0.15 and 1,000 points of support.

(c) The shaded area represents the 95% highest posterior density interval defined by the smoothed density.

Figure 3.3: Impact of Vertical Integration on Independent Downstream Payoffs in Large Market Structures



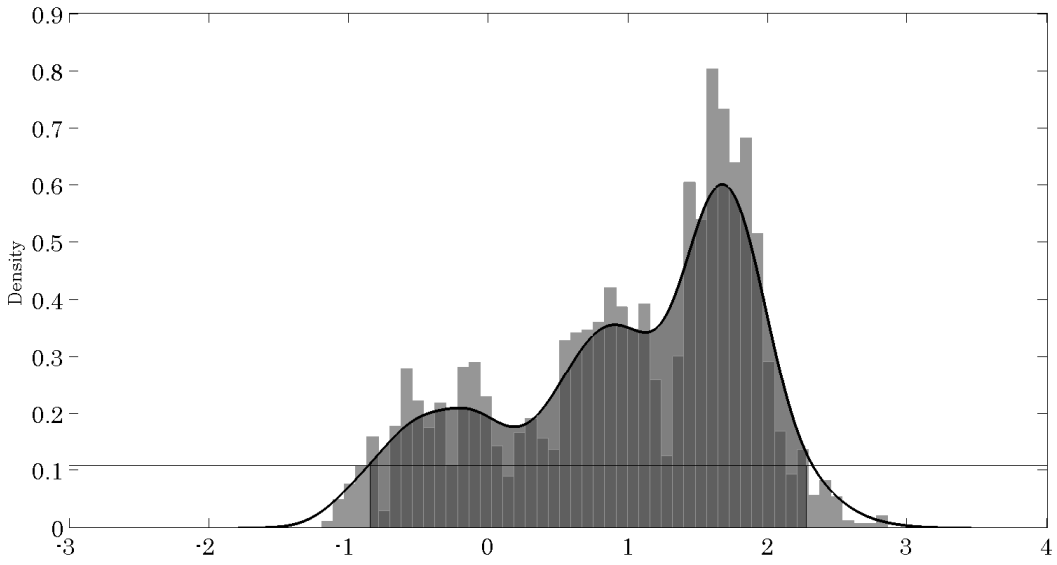

Notes:

(a) The bars form a density histogram of MCMC realizations for the impact of rival vertical integration on independent downstream payoffs in the 1-by-2 market structure.

(b) The smoothed density is generated using a Gaussian kernel with bandwidth of 0.15 and 1,000 points of support.

(c) The shaded area represents the 95% highest posterior density interval defined by the smoothed density.

Figure 3.4: Impact of Vertical Integration on Independent Downstream Payoffs in One-by-Two Market Structure



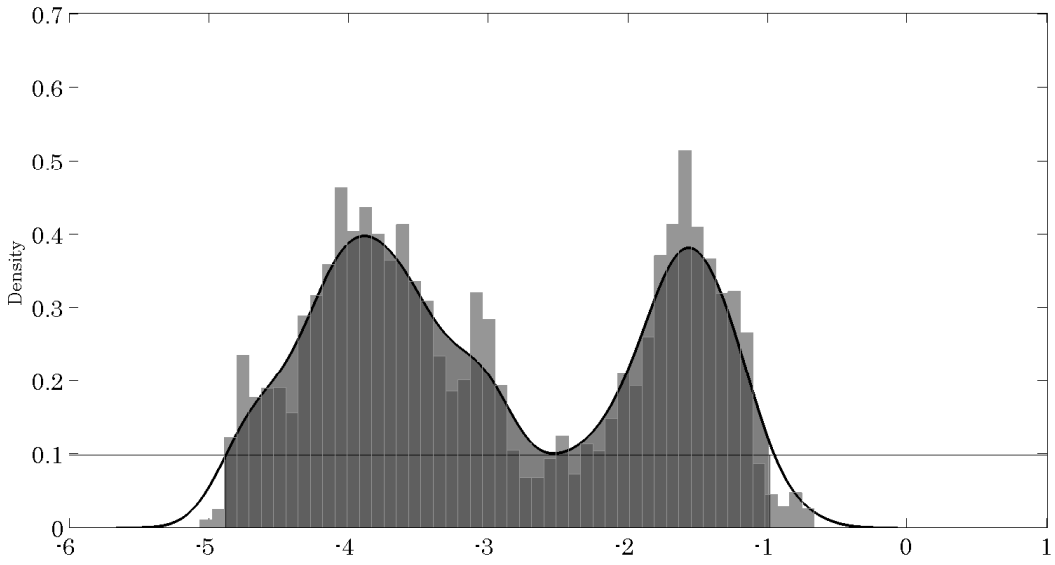

Notes:

(a) The bars form a density histogram of MCMC realizations for the impact of rival vertical integration on independent upstream payoffs in market structures with two or more entrants in both the upstream and downstream segments.

(b) The smoothed density is generated using a Gaussian kernel with bandwidth of 0.15 and 1,000 points of support.

(c) The shaded area represents the 95% highest posterior density interval defined by the smoothed density.

Figure 3.5: Impact of Vertical Integration on Independent Upstream Payoffs in Large Market Structures




Notes:

(a) The bars form a density histogram of MCMC realizations for the impact of rival vertical integration on independent upstream payoffs in the 2-by-1 market structure.

(b) The smoothed density is generated using a Gaussian kernel with bandwidth of 0.15 and 1,000 points of support.

(c) The shaded area represents the 95% highest posterior density interval defined by the smoothed density.

Figure 3.6: Impact of Vertical Integration on Independent Upstream Payoffs in Two-by-One Market Structure

downstream entrants, because its efficiency effects may benefit unintegrated downstream entrants through positive spillovers. If vertical integration is found to promote entry into the downstream segment, we can conclude that vertical integration has a procompetitive overall effect. This is because the combination of greater downstream entry and significant efficiency effects unambiguously implies lower prices and/or higher quality for the final product.

To examine how vertical integration affects market structure formation, I conduct a policy simulation. Specifically, I simulate the effect of a hypothetical policy that bans any firm from entering both vertical segments of the same market. To my knowledge, no such policy has yet been contemplated for the generic pharmaceutical industry. However, it is similar in spirit to vertical separation regulations found, for example, in the electric utility industry. Recent antitrust cases such as *FTC v. Mylan et al.* (D.D.C., 1999), involving exclusive dealing contracts between API manufacturers and finished formulation firms, have shown that vertical practices can have highly anticompetitive effects in the generic drug industry. This suggests that vertically integrated entry might come under stronger antitrust scrutiny in the future.

If vertical integration has an entry-reducing effect, the ban on vertically integrated entry should increase the equilibrium number of downstream entrants relative to the status quo where vertically integrated entry is allowed. Conversely, the ban would reduce the number of downstream entrants in equilibrium if vertical integration has an entry-promoting effect.

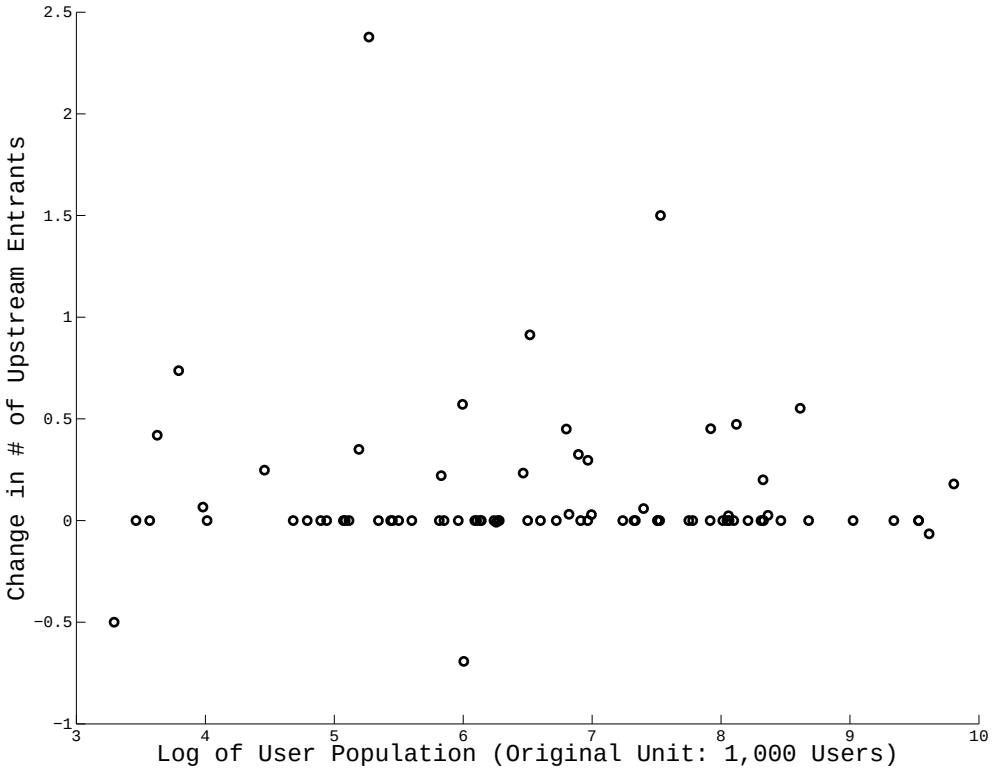
To simulate the effect of the policy, I run two sets of predictions on equilibrium market structures. In the first set, firms are allowed to enter as a vertically integrated entity. In the second set, I simulate the ban by removing "vertically integrated entry" from the choice set of every potential entrant. For both sets, I make predictions for 50 draws of the error term vector (the same draws that are used in simulation-based estimation of the parameters), and for each draw I compute all pure strategy Nash equilibria of the entry game. The parameter values that I use are the modes of the marginal posterior distributions.[37]

Figures 3.7 and 3.8 present the results of policy simulation. To compare the predicted equilibrium market structures with and without the ban on vertically integrated entry, I count the number of entrants that are predicted to be present in each vertical segment. A vertically integrated entrant is counted as both an upstream entrant and a downstream entrant. For the 85 markets in the dataset, I calculate the mean number of entrants in each segment, averaging over draws as well as over multiple equilibria. Each dot in Figures 3.7 and 3.8 represents a sample market, with the horizontal axis measuring the market's user population and the vertical axis measuring the change in number of entrants caused by the vertical entry ban.

In Figure 3.7, we see that the number of upstream entrants tends to be greater when vertically integrated entry is banned. Of the 85 sample markets, 24 (28.24 percesnt) experience an increase in the predicted number of entrants, only four (4.71 percent) experience a decrease, and 52 (61.18 percent) experience no change.[38] The magnitude of impact tends to be greater for pos-

[37]This implies that the distribution of the parameters is not taken into account. The results can be made to reflect the distribution of parameters by running the simulation separately for a large number of draws from the posterior distribution.

[38]For five markets (5.88 percent of the total), no pure strategy equilibrium could be found for the vertical entry




Notes:

(a) Each dot represents a market.

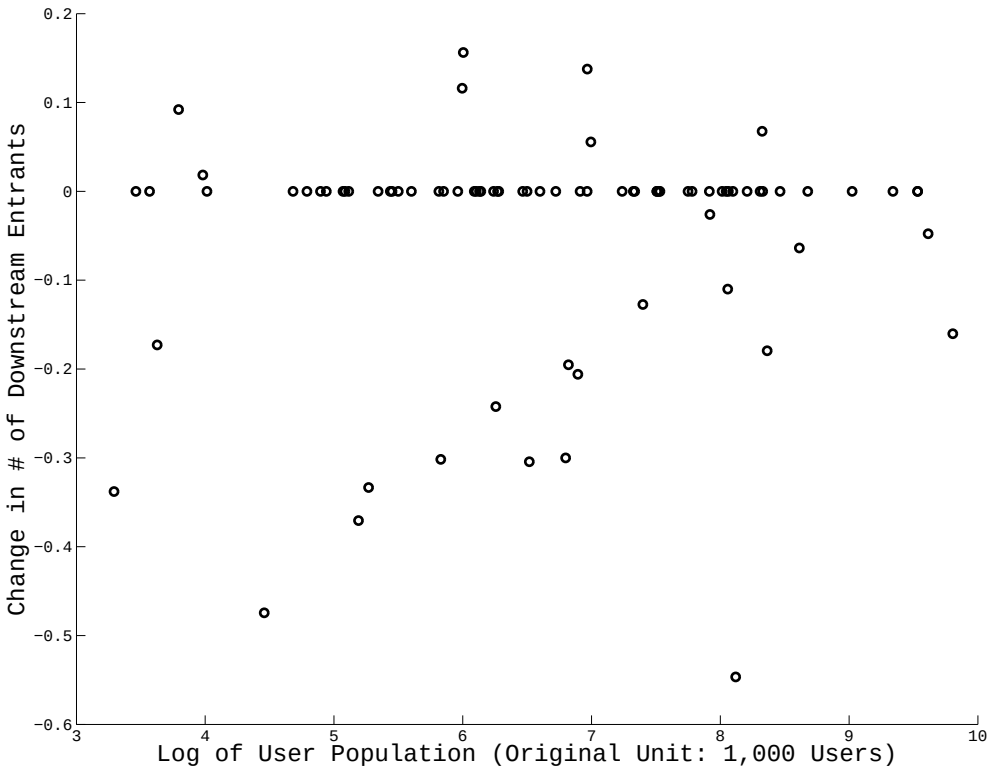
(b) The marginal posterior modes are used as parameter values to predict the equilibrium market structure for each market under both “Vertical Entry Allowed” and “Vertical Entry Prohibited”.

(c) The vertical axis measures the difference between the predicted number of upstream entrants under “Vertical Entry Prohibited” and that under “Vertical Entry Allowed”.

(d) For comparability, each vertically integrated entry in the “Vertical Entry Allowed” scenario is recounted as a pair consisting of one independent upstream and one independent downstream entry.

(e) The predicted market structures are averaged, separately for each vertical segment, over multiple equilibria as well as over draws of the random error vector.

Figure 3.7: Simulated Effect of Vertical Entry Ban on Number of Upstream Entrants




Notes:

(a) Each dot represents a market.

(b) The marginal posterior modes are used as parameter values to predict the equilibrium market structure for each market under both "Vertical Entry Allowed" and "Vertical Entry Prohibited".

(c) The vertical axis measures the difference between the predicted number of downstream entrants under "Vertical Entry Prohibited" and that under "Vertical Entry Allowed".

(d) For comparability, each vertically integrated entry in the "Vertical Entry Allowed" scenario is recounted as a pair consisting of one independent upstream and one independent downstream entry.

(e) The predicted market structures are averaged, separately for each vertical segment, over multiple equilibria as well as over draws of the random error vector.

Figure 3.8: Simulated Effect of Vertical Entry Ban on Number of Downstream Entrants

itive changes: among those markets where the expected number of upstream entrants increases, the average change is +0.4472, whereas among the markets experiencing a decrease the average change is -0.3167. The efficiency effect of vertical integration appears to be sufficiently strong to deter some firms from entering as an unintegrated upstream supplier. By removing this entry deterrent effect, the vertical entry ban succeeds in promoting upstream entry.

On the other hand, Figure 3.8 shows that the number of downstream entrants tends to decrease in response to the vertical entry ban. The number of downstream entrants decreases in nineteen markets (22.35 percent), while it increases in seven (8.24 percent) and remains unchanged in 54 (63.53 percent). The absolute value of the change is larger in markets experiencing a decrease: while the average positive change is +0.0919, the average negative change is -0.2368. It thus appears that the efficiency spillovers from vertical integration are so large that, despite decreased upstream entry, more downstream entry occurs when vertically integrated entry is allowed. It is also possible that the problem of double marginalization is more severe under the vertical entry ban, leading to fewer downstream entrants despite greater entry in the upstream segment. The policy of banning vertically integrated entry is therefore counterproductive. By decreasing the number of downstream entrants and depriving the opportunities for efficiency enhancement through vertical integration, the ban is likely to have a negative impact on market outcomes.

## 3.8 Conclusion

The econometric model presented in this chapter offers a novel way to make inferences about the competitive effects of vertical integration based on an entry game framework. The model requires observations on multiple markets where entry into both the upstream and downstream segments are recorded. While such data may not be available for most industries, in some sense the data requirements for the method are lighter than for other methods used for investigating the effects of vertical integration: one need not directly observe market outcomes and one need not assume that vertical market structures are exogenously given.[39]

Application of the vertical entry model to the US generic pharmaceutical industry yields the following conclusion: vertical integration in this industry is characterized by significant efficiency effects that spill over to benefit unintegrated downstream firms. This follows from the finding that unintegrated downstream firms gain significantly from rival vertical integration. The finding that unintegrated upstream profits are lowered by rival vertical integration in the two-by-one market structure is also consistent with the existence of efficiency effects.

The parameter estimates are used to simulate a hypothetical policy that bans vertically integrated entry. While the number of upstream entrants increases in response to the ban, the number of downstream entrants decreases. Both movements can be explained by the existence of large efficiency effects. Because vertically integrated entrants are tough competitors in the upstream

game under the ban.

[39] As demonstrated by Berry and Waldfogel (1999) and Mazzeo (2002a), the range of questions that can be answered in an entry game framework is enhanced if data on prices and other market outcomes are also available.

segment, banning vertical integrated entry allows a greater number of unintegrated upstream entrants – who are likely to be less efficient than their vertically integrated counterparts, but also likely to have lower entry costs – to come in. Despite the greater competition caused by increased entry,, the number downstream entrants decreases. This could be due to the lower efficiency of unintegrated upstream units, or to double marginalization problems under complete vertical separation. The combination of reduced downstream entry and reduced upstream efficiency implies that overall market performance falls as a result of the vertical entry ban.

The policy simulation described above demonstrates the advantage of the present model over existing methods that analyze the effect of vertical integration. Because the model treats market structure formation as an endogenous process, it can be used to examine how market structures are affected by the act of vertical integration.

# Bibliography


Abrevaya, Jason and Jerry A. Hausman (1999) "Semiparametric Estimation with Mismeasured Dependent Variables: An Application to Duration Models for Unemployment Spells." *Annales d'Économie et de Statistique* 55/56, 243–275.

Ackerberg, Daniel A. (2009) "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation." *Quantitative Marketing and Economics* 7(4), 343–376.

Andrews, Donald W. K., Steven Berry and Panle Jia (2004) "Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Store Location." Mimeograph.

Aydemir, Zava and Stefan Buehler (2003) "Estimating Vertical Foreclosure in US Gasoline Supply." Paper Presented at the Royal Economic Society Annual Conference.

Bain, Joe S. (1956) *Barriers to New Competition: Their Character and Consequences in Manufacturing Industries* (Cambridge, MA: Harvard University Press).

Bajari, Patrick, C. Lanier Benkard and Jonathan Levin (2007) "Estimating Dynamic Models of Imperfect Competition." *Econometrica* 75(5), 1331–1370.

Bajari, Patrick, Han Hong and Stephen Ryan (2010) "Identification and Estimation of a Discrete Game of Complete Information." *Econometrica* 78(5), 1529–1568.

Barnes, Kirsty (2006) "Watson Sees Future in India." *In-PharmaTechnologist.com.* April 18.

Berndt, Ernst R. and Murray L. Aitken (2010) "Brand Loyalty, Generic Entry and Price Competition in Pharmaceuticals in the Quarter Century After the 1984 Waxman-Hatch Legislation." NBER Working Paper No. 16431.

Berry, Steven T. (1992) "Estimation of a Model of Entry in the Airline Industry." *Econometrica* 60(4), 889–917.

Berry, Steven T. and Joel Waldfogel (1999) "Free Entry and Social Inefficiency in Radio Broadcasting." *Rand Journal of Economics* 30(3), 397–420.

Bjorn, Paul A. and Quang H. Vuong (1984) "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation." California Institute of Technology Social Science Working Paper No.537.

Bolton, Patrick and Michael Whinston (1993) "Incomplete Contracts, Vertical Integration, and Supply Assurance." *Review of Economic Studies* 60(1), 121–148.

Bresnahan, Timothy F. and Peter C. Reiss (1991a) "Empirical Models of Discrete Games." *Journal of Econometrics* 48, 57–81.

_ (1991b) "Entry and Competition in Concentrated Markets." *Journal of Political Economy* 99(5), 977–1009.

Bryant, Rob (2004) *Pharmaceutical Fine Chemicals: Global Perspectives 2000* (Orpington: Brychem Business Consulting).

Buehler, Gary (2006) "FDA Efforts to Expedite Generic Drug Approvals." Statement before the Senate Special Committee on Aging.

Buehler, Stefan and Armin Schmutzler (2005) "Asymmetric Vertical Integration." *Advances in Theoretical Economics.* 5(1), article 1.

Bulow, Jeremy (2004) "The Gaming of Pharmaceutical Patents." *Innovation Policy and the Economy* 4, 145–187.

Burck, Benjamin (2010) "Preparing for Paragraph IV Patent Challenges." Scrip Intelligence.

Butler, J. S. (1996) "Estimating the Correlation in Censored Probit Models." *Review of Economics and Statistics* 78(2), 356–358.

Carlton, Dennis W. (1979) "Vertical Integration in Competitive Markets Under Uncertainty." *Journal of Industrial Economics* 27(3), 189–209.

Case, Anne C. and Lawrence F. Katz (1991) "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." NBER Working Paper No. 3705.

Caves, Richard E. and Ralph M. Bradburd (1988) "The Empirical Determinants of Vertical Integration." *Journal of Economic Behavior and Organization* 9, 265–279.

Caves, Richard, Michael Whinston and Mark Hurwitz (1991) "Patent Expiration, Entry, and Competition in the US Pharmaceutical Industry." *Brookings Papers on Economic Activity (Microeconomics)* pp. 1–48.

Chaudhuri, Sudip (2005) *The WTO and India's Pharmaceuticals Industry: Patent Protection, TRIPS, and Developing Countries* (New Delhi, India: Oxford University Press).

Chen, Yongmin (2001) "On Vertical Mergers and their Competitive Effects." *Rand Journal of Economics* 32(4), 667–685.

Chernozhukov, Victor and Han Hong (2003) "An MCMC Approach to Classical Estimation." *Journal of Econometrics* 115, 293–346.

Chipty, Tasneem (2001) "Vertical Integration, Market Foreclosure, and Consumer Welfare in the Cable Television Industry." *American Economic Review* 91(3), 428–453.

Ciliberto, Federico (2006) "Does Organizational Form Affect Investment Decisions?" *Journal of Industrial Economics* 54(1), 63–93.

Ciliberto, Federico and David Dranove (2006) "The Effect of Physician-Hospital Affiliations on Hospital Prices in California." *Journal of Health Economics* 25(1), 29–38.

Ciliberto, Federico and Elie Tamer (2009) "Market Structure and Multiple Equilibria in Airline Markets." *Econometrica* 77(6), 1791–1828.

Cohen, Andrew M. and Michael J. Mazzeo (2007) "Market Structure and Competition Among Retail Depository Institutions." *Review of Economics and Statistics* 89(1), 60–74.

de Fontenay, Catherine C. and Joshua S. Gans (2005) "Vertical Integration in the Presence of Upstream Competition." *Rand Journal of Economics* 36(3), 544–572.

Deutsch, Claudia H. (2003) "Market Place; FedEx Moves to Expand with Purchase of Kinko's." *New York Times* p. C1. December 31.

Dufeu, Ivan (2004) "Market Size and Vertical Equilibrium in the Context of Successive Cournot Oligopolies." *Topics in Theoretical Economics.* 4(1), article 2.

El-Gamal, Mahmoud A., Richard D. McKelvey and Thomas R. Palfrey (1993) "A Bayesian Sequential Experimental Study of Learning in Games." *Journal of the American Statistical Association* 88(422), 428–435.

Elberfeld, Walter (2002) "Market Size and Vertical Integration: Stigler's Hypothesis Reconsidered." *Journal of Industrial Economics* 50(1), 23–42.

Ellickson, Paul B. and Sanjog Misra (2008) "Supermarket Pricing Strategies." *Marketing Science* 27(5), 811–828.

Ellison, Glenn and Sara Fisher Ellison (2007) "Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbents Prior to Patent Expiration." NBER Working Paper No. 13069.

Erdei, Tommy (2004) "M&A Strategy Within the Generic Pharmaceuticals Sector: Vertical Integration into Active Pharmaceutical Ingredients." *Journal of Generic Medicines* 2(1), 18–22.

Evans, William N., Wallace E. Oates and Robert M. Schwab (1992) "Measuring Peer Group Effects: A Study of Teenage Behavior." *Journal of Political Economy* 100(5), 966–991.

Farrell, Joseph (1987) "Cheap Talk, Coordination, and Entry." *Rand Journal of Economics* 18(1), 34–39.

Federal Trade Commission (1966) *Economic Report on Mergers and Vertical Integration in the Cement Industry* (Washington, D.C.: Federal Trade Commission).

_ (2002) *Generic Drug Entry Prior to Patent Expiration: An FTC Study* (Washington, D.C.: Federal Trade Commission).

_ (2006) "In the Matter of Teva Pharmaceutical Industries Ltd. and IVAX Corporation." Analysis of Agreement Containing Consent Orders to Aid Public Comment, File No. 051-0214, Docket No. C-4155.

Filson, Darren and Ahmed Oweis (2010) "The Impacts of the Rise of Paragraph IV Challenges on Startup Alliance Formation and Firm Value in the Pharmaceutical Industry." *Journal of Health Economics* 29, 575–584.

Food and Drug Administration (1989) "Guidline for Drug Master Files." Procedural guidance.

Frank, Richard G. and David Salkever (1997) "Generic Entry and the Pricing of Pharmaceuticals." *Journal of Economics and Management Strategy* 6, 75–90.

Gallant, A. Ronald, Han Hong and Ahmed Khwaja (2008) "Estimating a Dynamic Oligopolistic Game with Serially Correlated Unobserved Production Costs." Mimeograph.

_ (2010) "Dynamic Entry with Cross-Product Spillovers: An Application to the Generic Drug Industry." Economic Research Initiatives at Duke Working Paper Series No. 57.

Garnett, Leah R. (2000) "Prozac Revisited: As Drug Gets Remade: Concerns About Suicides Surface." *Boston Globe* p. A1. May 7.

Grabowski, Henry G. (2004) "Are the Economics of Pharmaceutical Research and Development Changing? Productivity, Patents, and Political Pressures." *Pharmacoeconomics* 22(Suppl. 2), 15–24.

Grabowski, Henry G. and Margaret Kyle (2007) "Generic Competition and Market Exclusivity Periods in Pharmaceuticals." *Managerial and Decision Economics* 28, 491–502.

Greene, William H. (1996) "Marginal Effects in the Bivariate Probit Model." NYU Working Paper EC-96-11.

_ (2008) *Econometric Analysis,* 6 ed. (Upper Saddle River, NJ: Pearson Prentice Hall).

Grossman, Sanford J. and Oliver D. Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94(4), 691–719.

Harrison, Charlotte (2008) "Ranbaxy Loses in Caduet Dispute." *Nature Reviews Drug Discovery* 7, 113.

Hart, Oliver and Jean Tirole (1990) "Vertical Integration and Market Foreclosure." *Brookings Papers on Economic Activity (Microeconomics)* pp. 205–285.

Hastings, Justine and Richard Gilbert (2005) "Market Power, Vertical Integration, and the Wholesale Price of Gasoline." *Journal of Industrial Economics* 53(4), 469–492.

Hemphill, C. Scott and Bhaven N. Sampat (2010) "When Do Generics Challenge Drug Patents." Columbia University Center for Law and Economic Studies Working Paper No. 379.

Hendricks, Kenneth and R. Preston McAfee (2010) "A Theory of Bilateral Oligopoly." *Economic Inquiry* 48(2), 391–414.

Higgins, Matthew J. and Stuart J.H. Graham (2006) "Balancing Innovation and Access: Patent Challenges Tip the Scales." *Science* 326, 370–371.

Hoffman, Jean (2004) "Global Generics Profitability: Increasing Need for Business Development and Licensing." *Journal of Generic Medicines* 2(1), 9–17.

Hortaçsu, Ali and Chad Syverson (2007) "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." *Journal of Political Economy* 115(2), 250–301.

Hsiao, Chun-Ju (2010) "Understanding and Using NAMCS and NHAMCS Data: Data Tools and Basic Programming Techniques." Presentation Slides for the 2010 National Conference on Health Statistics.

Hutchins, Mike (2003) "Extending the Monopoly: How 'Secondary Patents' Can Be Used to Delay or Prevent Generic Competition Upon Expiry of the Basic Product Patent." *Journal of Generic Medicines* 1(1), 57–71.

International Federation of Pharmaceutical Manufacturers and Associations (2005) *A Review of Existing Data Exclusivity Legislation in Selected Countries*, 5 ed. (Geneva, Switzerland: International Federation of Pharmaceutical Manufacturers and Associations).

Jia, Panle (2008) "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry." *Econometrica* 76(6), 1263–1316.

Karst, Kurt R. (2010) "OGD's ANDA Backlog and Median ANDA Approval Times are Up - WAY UP! 'The Solution Lies in Resources,' Says FDA Commissioner Hamburg." FDA Law Blog.

Karwal, Vijay (2006) "The Changing Competitive Landscape in the Global Generics Market: Threat or Opportunity?" *Journal of Generic Medicines* 3(4), 269–279.

Klein, Benjamin, Robert G. Crawford and Armen A. Alchian (1979) "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *Journal of Law and Economics* 21(2), 297–326.

Korn, David E., Erika Lietzan and Shaw W. Scott (2009) "A New History and Discussion of 180-Day Exclusivity." *Food and Drug Law Journal* 64(2), 335–390.

Lafontaine, Francine and Margaret Slade (2007) "Vertical Integration and Firm Boundaries: The Evidence." *Journal of Economic Literature* 45(3), 629–685.

Lancaster, Tony (2004) *An Introduction to Modern Bayesian Econometrics* (Malden, MA: Blackwell Publishing).

Lanjouw, Jean O. (1998) "The Introduction of Pharmaceutical Product Patents in India: 'Heartless Exploitation of the Poor and Suffering'?" NBER Working Paper No. 6366.

Lei, Zhen and Brian D. Wright (2009) "Why Weak Patents? Rational Ignorance or Pro-"customer" Tilt?" Paper presented at the 2009 Annual Meeting of the Agricultural and Applied Economics Association.

Lemley, Mark and Carl Shapiro (2005) "Probabilistic Patents." *Journal of Economic Perspectives* 19(2), 75–98.

Lieberman, Marvin B. (1991) "Determinants of Vertical Integration: An Empirical Test." *Journal of Industrial Economics* 39(5), 451–466.

Lietzan, Erika King (2004a) "2004 Update: 180-Day Exclusivity Under the Hatch-Waxman Amendments to the Federal Food, Drug, and Cosmetic Act." *Food and Drug Law Journal* 59, 459–463.

_ (2004b) "A Brief History of 180-Day Exclusivity Under the Hatch-Waxman Amendments to the Federal Food, Drug, and Cosmetic Act." *Food and Drug Law Journal* 59, 287–323.

Loury, Glenn C. (1979) "Market Structure and Innovation." *Quarterly Journal of Economics* 93(3), 395–410.

Mándi, Attila (2003) "Protection and Challenge of Pharmaceutical Patents." *Journal of Generic Medicines* 1(1), 72–82.

Manuszak, Mark D. (2010) "Predicting the Impact of Upstream Mergers on Downstream Markets with an Application to the Retail Gasoline Industry." *International Journal of Industrial Organization* 28(1), 99–111.

Martin, Stephen, Hans-Theo Normann and Christopher M. Snyder (2001) "Vertical Foreclosure in Experimental Markets." *Rand Journal of Economics* 32(3), 466–496.

Mazzeo, Michael J. (2002a) "Competitive Outcomes in Product-differentiated Oligopoly." *Review of Economics and Statistics* 84(4), 716–728.

_ (2002b) "Product Choice and Oligopoly Market Structure." *Rand Journal of Economics* 33(2), 221–242.

McLaren, John (2000) "'Globalization' and Vertical Structure." *American Economic Review* 90(5), 1239–1254.

Meng, Chun-Lo and Peter Schmidt (1985) "On the Cost of Partial Observability in the Bivariate Probit Model." *International Economic Review* 26(1), 71–85.

Miller, Joseph Scott (2004) "Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents." *Berkeley Technology Law Journal* 19(2), 667–739.

Murphy, Kevin M. and Robert H. Topel (1985) "Estimation and Inference in Two-Step Econometric Models." *Journal of Business and Economic Statistics* 3(4), 370–379.

Normann, Hans Theo (2007) "Vertical Mergers, Foreclosure and Raising Rivals' Costs: Experimental Evidence." Mimeograph.

Ordover, Janusz A., Steven C. Salop and Garth Saloner (1990) "Equilibrium Vertical Foreclosure." *American Economic Review* 80, 127–142.

Orhun, A. Yeşim (2005) "Spatial Differentiation in the Supermarket Industry." Mimeograph.

Pakes, Ariel, Jack Porter, Kate Ho and Joy Ishii (2006) "Moment Inequalities and Their Application." Mimeograph.

Perry, Martin K. and Robert H. Groff (1988) "Vertical Integration and Industry Growth." Discussion Paper, Rutgers University and Bell Communications Research.

Reiffen, David and Michael Ward (2005) "Generic Drug Industry Dynamics." *Review of Economics and Statistics* 87(1), 37–49.

Reinganum, Jennifer F. (1983) "Uncertain Innovation and the Persistence of Monopoly." *American Economic Review* 73(4), 741–748.

Rey, Patrick and Jean Tirole (2007) "A Primer on Foreclosure." In *Handbook of Industrial Organization*, ed. Mark Armstrong and Robert Porter, vol. 3 (Amsterdam: Elsevier) chapter 33, pp. 2145–2220.

Richard, Jean-Francois and Wei Zhang (2007) "Efficient High-Dimensional Importance Sampling." *Journal of Econometrics* 141, 1385–1411.

Riordan, Michael H. (2008) "Competitive Effects of Vertical Integration." In *Handbook of Antitrust Economics,* ed. Paolo Buccirossi (Cambridge, MA: MIT Press) pp. 145–182.

Robert, Christian P. and George Casella (2004) *Monte Carlo Statistical Methods,* 2 ed. (New York, NY: Springer-Verlag).

Rosengren, Eric S. and James W. Meehan (1994) "Empirical Evidence on Vertical Foreclosure." *Economic Inquiry* 32, 303–317.

Roumeliotis, Gregory (2006) "Mylan in India's Biggest Pharmaceutical Takeover as it Enters the Matrix." *In-PharmaTechnologist.com.* August 30.

Salinger, Michael A. (1988) "Vertical Mergers and Market Foreclosure." *Quarterly Journal of Economics* 103(2), 345–356.

Schweitzer, Stuart O. (2007) *Pharmaceutical Economics and Policy,* 2 ed. (New York, NY: Oxford University Press).

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

Seim, Katja (2006) "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." *Rand Journal of Economics* 37(3), 619–640.

Shaw, Arthur B. (2008) "Drug Master Files." Presentation slides prepared for Drug Information Association Webinar.

Silber, Seth C. (2007) "FTC Review of Generic Drug Mergers – Four Recent Consents Provide Significant Guidance on Mode of Analysis and Divestiture Requirements." *Journal of Generic Medicines* 4(3), 162–168.

Snyder, Christopher M. (1995) "Empirical Studies of Vertical Foreclosure." In "Industry Economics Conference Papers and Proceedings," vol. 23 (Canberra: University of Melbourne and Bureau of Industry Economics, AGPS) pp. 98–125.

Spengler, Joseph J. (1950) "Vertical Integration and Antitrust Policy." *Journal of Political Economy* 58(4), 347–352.

Stafford, Ambrose (2006) "The Evolution of the Global Active Pharmaceutical Ingredients Supply Market and the Impact on the Generic Pharmaceuticals Business in the United States." *Journal of Generic Medicines* 3(4), 295–305.

Stigler, George J. (1951) "The Division of Labor Is Limited by the Extent of the Market." *Journal of Political Economy* 59(3), 185–193.

Surendar, T. (2009) "Just What the Doctor Ordered." *Forbes India.* December 4.

Suzuki, Ayako (2008) "Market Foreclosure and Vertical Merger: A Case Study of the Vertical Merger between Turner Broadcasting and Time Warner." *International Journal of Industrial Organization* 27(4), 532–543.

Teva Pharmaceutical Industries (2008) *Annual Report (Form 20-F)* (Financial statement).

_ (2009) *Annual Report (Form 20-F)* (Financial statement).

Train, Kenneth E. (2009) *Discrete Choice Methods with Simulation,* 2 ed. (Cambridge University Press).

Turocy, Theodore L. (2008) "Towards a Black-Box Solver for Finite Games: Finding All Nash Equilibria with Gambit and PHCpack." In *Software for Algebraic Geometry,* ed. Michael E. Stillman, Nobuki Takayama and Jan Verschelde, vol. 148 of *IMA Volumes in Mathematics and its Applications* (New York: Springer) pp. 133–147.

Villas-Boas, Sofia Berto (2007) "Using Retail Data for Upstream Merger Analysis." *Journal of Competition Law and Economics* 3(4), 689–715.

Vitorino, Maria Ana (2010) "Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry." Mimeograph.

Whinston, Michael D. (2003) "On the Transaction Cost Determinants of Vertical Integration." *Journal of Law, Economics, and Organization* 19(1), 1–23.

Williamson, Oliver E. (1971) "The Vertical Integration of Production: Market Failure Considerations." *American Economic Review Papers and Proceedings* 61(2), 112–123.

Woodruff, Christopher (2002) "Non-contractible Investments and Vertical Integration in the Mexican Footwear Industry." *International Journal of Industrial Organization* 20(8), 1197–1224.

Wooldridge, Jeffrey M. (2002) *Econometric Analysis of Cross Section and Panel Data* (Cambridge, MA: MIT Press).

# Appendix A

# Ancillary Material for Chapter 2

## A.1 Marginal Effects on Conditional Outcome Probability in Trivariate Probit Model

Here, I derive the marginal effect of changes in the covariates on the probability of vertical integration by a potential downstream entrant, conditional on the firm having entered the downstream segment and on the market not being subject to paragraph IV certification. The conditional probability is written as

$$Prob(VI = 1 \mid DE = 1,\, PF = 0,\, \bar{\mathbf{x}}) = \frac{Prob(VI = 1,\, DE = 1,\, PF = 0 \mid \bar{\mathbf{x}})}{Prob(DE = 1,\, PF = 0 \mid \bar{\mathbf{x}})},$$

where $\bar{\mathbf{x}} = \bar{\mathbf{x}}_1 \cup \bar{\mathbf{x}}_2 \cup \bar{\mathbf{x}}_3$ contains representative values of the covariates. The probabilities on the right-hand side are written out as

$$Prob(VI = 1,\, DE = 1,\, PF = 0 \mid \bar{\mathbf{x}}) = \int_{-\infty}^{-\beta_3' \bar{\mathbf{x}}_{3m}} \int_{-\beta_2' \bar{\mathbf{x}}_{2mi}}^{\infty} \int_{-\beta_1' \bar{\mathbf{x}}_{1mi}}^{\infty} f_3\left(\varepsilon_1, \varepsilon_2, \varepsilon_3; \Sigma\right) d\varepsilon_1 d\varepsilon_2 d\varepsilon_3,$$

$$Prob(DE = 1,\, PF = 0 \mid \bar{\mathbf{x}}) = \int_{-\infty}^{-\beta_3' \bar{\mathbf{x}}_{3m}} \int_{-\beta_2' \bar{\mathbf{x}}_{2mi}}^{\infty} \phi_2\left(\varepsilon_2, \varepsilon_3; 0\right) d\varepsilon_2 d\varepsilon_3,$$

where $\phi_2(\cdot;\rho)$ is the density of a standard bivariate normal distribution with correlation coefficient $\rho$. The marginal effect of a continuous covariate $x_k$, which may belong to one, two, or all three of the covariate vectors $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$, is derived as follows:

$$\frac{\partial Prob(VI=1 \mid DE=1,\, PF=0,\, \bar{\mathbf{x}})}{\partial x_k} = \frac{1}{[Prob(DE=1,\, PF=0 \mid \bar{\mathbf{x}})]^2}$$

$$\times \left\{ \frac{\partial Prob(VI=1,\, DE=1,\, PF=0 \mid \bar{\mathbf{x}})}{\partial x_k} Prob(DE=1,\, PF=0 \mid \bar{\mathbf{x}}) \right.$$

$$\left. - Prob(VI=1,\, DE=1,\, PF=0 \mid \bar{\mathbf{x}}) \frac{\partial Prob(DE=1,\, PF=0 \mid \bar{\mathbf{x}})}{\partial x_k} \right\}$$

$$= \frac{1}{\Phi(\beta_2'\bar{\mathbf{x}}_2)\Phi(-\beta_3'\bar{\mathbf{x}}_3)}$$

$$\times \left[ \beta_{1k}\phi(\beta_1'\bar{\mathbf{x}}_1) \int_{-\infty}^{-\beta_3'\bar{\mathbf{x}}_3} \int_{-\beta_2'\bar{\mathbf{x}}_2}^{\infty} f_2\left(\varepsilon_2 + \rho_{12}\beta_1'\bar{\mathbf{x}}_1,\, \varepsilon_3 + \rho_{13}\beta_1'\bar{\mathbf{x}}_1;\, \Sigma_{23|1}\right) d\varepsilon_2 d\varepsilon_3 \right.$$

$$+ \beta_{2k}\phi(\beta_2'\bar{\mathbf{x}}_2) \int_{-\infty}^{-\beta_3'\bar{\mathbf{x}}_3} \int_{-\beta_1'\bar{\mathbf{x}}_1}^{\infty} f_2\left(\varepsilon_1 + \rho_{12}\beta_2'\bar{\mathbf{x}}_2,\, \varepsilon_3;\, \Sigma_{13|2}\right) d\varepsilon_1 d\varepsilon_3$$

$$\left. - \beta_{3k}\phi(\beta_3'\bar{\mathbf{x}}_3) \int_{-\beta_2'\bar{\mathbf{x}}_2}^{\infty} \int_{-\beta_1'\bar{\mathbf{x}}_1}^{\infty} f_2\left(\varepsilon_1 + \rho_{13}\beta_3'\bar{\mathbf{x}}_3,\, \varepsilon_2;\, \Sigma_{12|3}\right) d\varepsilon_1 d\varepsilon_2 \right]$$

$$- \frac{\beta_{2k}\phi(\beta_2'\bar{\mathbf{x}}_2)\Phi(-\beta_3'\bar{\mathbf{x}}_3) - \beta_{3k}\phi(\beta_3'\bar{\mathbf{x}}_3)\Phi(\beta_2'\bar{\mathbf{x}}_2)}{\left[\Phi(\beta_2'\bar{\mathbf{x}}_2)\Phi(-\beta_3'\bar{\mathbf{x}}_3)\right]^2}$$

$$\times \int_{-\infty}^{-\beta_3'\bar{\mathbf{x}}_3} \int_{-\beta_2'\bar{\mathbf{x}}_2}^{\infty} \int_{-\beta_1'\bar{\mathbf{x}}_1}^{\infty} f_3(\varepsilon_1, \varepsilon_2, \varepsilon_3;\, \Sigma) d\varepsilon_1 d\varepsilon_2 d\varepsilon_3, \tag{A.1}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function and $f_2(\cdot;\Sigma)$ is the density of a bivariate normal distribution with zero mean vector and covariance matrix $\Sigma$. The conditional covariance matrices are written out as follows:

$$\Sigma_{23|1} = \begin{bmatrix} 1-\rho_{12}^2 & -\rho_{12}\rho_{13} \\ -\rho_{12}\rho_{13} & 1-\rho_{13}^2 \end{bmatrix}, \Sigma_{13|2} = \begin{bmatrix} 1-\rho_{12}^2 & \rho_{13} \\ \rho_{13} & 1 \end{bmatrix}, \Sigma_{12|3} = \begin{bmatrix} 1-\rho_{13}^2 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}.$$

For the dichotomous covariates in $x$, the marginal effect on the conditional probability is calculated as

$$Prob(VI=1 \mid DE=1,\, PF=0,\, \bar{\mathbf{x}}_{-k},\, x_k=1) - Prob(VI=1 \mid DE=1,\, PF=0,\, \bar{\mathbf{x}}_{-k},\, x_k=0).$$

## A.2 Dataset Construction Details

### Identifying Generic Products in the FDA's Database

The US Food and Drug Administration's Orange Book contains information on all pharmaceutical finished formulations that have ever been approved, including those that have been discontin-

ued. [1] There are several methods to Identify generic approvals in the Orange Book data. One way is to refer to another database of the FDA, called Drugs@FDA, which identifies generic approvals with the term "ANDA".[2] However, the FDA's own classification appears to be imperfect. For instance, several drug approvals from before 1984 are classified as ANDAs, even though abbreviated new drug applications did not exist until after the passage of the Hatch-Waxman Amendments in 1984. Therefore, I use the FDA's classification in conjunction with another classification rule based on the trade name, or brand name, of a drug. Under this rule, an approved drug is classified as a generic if its trade name is the same as the generic name of the API contained in the drug. After applying both rules, I visually inspect all approvals in the database to correct obvious misclassifications.

## Identifying Firms and Treating Mergers

The FDA's data on ANDAs and DMFs often contain multiple (sometimes erroneous) names for the same firm. Moreover, different firms belonging to the same corporate group are not identified as such. To resolve this problem, I refer to the Newport Sourcing $^{TM}$ database, which classifies finished formulation manufacturers and API manufacturers into uniquely defined corporate groups. A firm in my dataset is equivalent to a corporate group as defined by Newport Sourcing.

Since Newport Sourcing identifies the older ANDAs and DMFs in terms of their current corporate group affiliations, one must take into account the many mergers and acquisitions – both horizontal and vertical – that have taken place in the generics industry during and around the observation period. For instance, Teva and IVAX were rivals in both the API and finished formulation industries until IVAX was acquired by Teva in January 2006. In the raw data from Newport Sourcing, however, the two firms are treated as being part of the same corporate group, even in markets that opened up prior to the acquisition. To fix this problem, I designate a separate corporate group for the observations for IVAX prior to the acquisition. Other ownership changes are similarly accounted for on the basis of news information on the timings of mergers and acquisitions that involve in-sample firms.

Merger and acquisition histories are also taken into account when determining a firm's potential entrant status on the basis of its past experience, or when constructing variables that measure a firm's entry experience. In doing so, I assume that an acquired firm's past entry experience is carried over to the acquiring firm, and that the new entity's entry experience is calculated as the sum of the two firms' experience levels.

[1] The Orange Book files are available from the FDA's website: `http://www.fda.gov/CDER/orange/obreadme.htm`.

[2] Drugs@FDA is accessible online at `http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm`.

Table A.1: List of Drugs in the Dataset

| Active Pharmaceutical Ingredient | Therapeutic Class | Dose Form | Market Opening |
|---|---|---|---|
| acebutolol hydrochloride | Cardiovascular | capsule | 1995 |
| acyclovir | Anti-Infective | capsule | 1997 |
| acyclovir | Anti-Infective | tablet | 1997 |
| alprazolam | Central Nervous System | tablet | 1993 |
| alprostadil | Endocrine-Metabolic | injectable | 1998 |
| amiodarone hydrochloride | Cardiovascular | tablet | 1998 |
| anagrelide hydrochloride | Blood Modifier | capsule | 2005 |
| azathioprine | Musculoskeletal | tablet | 1996 |
| azithromycin | Anti-Infective | tablet | 2005 |
| benazepril hydrochloride | Cardiovascular | tablet | 2004 |
| benzonatate | Respiratory | capsule | 1993 |
| betaxolol hydrochloride | Cardiovascular | tablet | 1999 |
| bromocriptine mesylate | Central Nervous System | tablet | 1998 |
| bumetanide | Cardiovascular | tablet | 1995 |
| bupropion hydrochloride | Central Nervous System | tablet | 1999 |
| buspirone hydrochloride | Central Nervous System | tablet | 2001 |
| cabergoline | Endocrine-Metabolic | tablet | 2005 |
| captopril | Cardiovascular | tablet | 1995 |
| carboplatin | Oncology | injectable | 2004 |
| cefotaxime sodium | Anti-Infective | injectable | 2002 |
| cefoxitin sodium | Anti-Infective | injectable | 2000 |
| cefpodoxime proxetil | Anti-Infective | tablet | 2004 |
| cefprozil | Anti-Infective | tablet | 2005 |
| cefuroxime axetil | Anti-Infective | tablet | 2002 |
| ciclopirox olamine | Dermatological | topical | 2004 |
| cilostazol | Blood Modifier | tablet | 2004 |
| cimetidine | Gastrointestinal | tablet | 1994 |
| ciprofloxacin hydrochloride | Anti-Infective | tablet | 2004 |
| cisplatin | Oncology | injectable | 1999 |
| citalopram hydrobromide | Central Nervous System | tablet | 2004 |
| clarithromycin | Anti-Infective | tablet | 2005 |
| clonazepam | Central Nervous System | tablet | 1997 |
| clozapine | Central Nervous System | tablet | 1997 |
| diclofenac potassium | Central Nervous System | tablet | 1998 |
| diclofenac sodium | Central Nervous System | ER tablet | 1995 |
| didanosine | Anti-Infective | ER capsule | 2004 |

*(Table continued on next page.)*

*(Continued from previous page.)*

| Active Pharmaceutical Ingredient | Therapeutic Class | Dose Form | Market Opening |
|---|---|---|---|
| dihydroergotamine mesylate | Central Nervous System | injectable | 2003 |
| doxazosin mesylate | Cardiovascular | tablet | 2000 |
| econazole nitrate | Dermatological | topical | 2002 |
| enalapril maleate | Cardiovascular | tablet | 2000 |
| estazolam | Central Nervous System | tablet | 1997 |
| ethambutol hydrochloride | Anti-Infective | tablet | 2000 |
| etodolac | Central Nervous System | tablet | 1997 |
| etoposide | Oncology | injectable | 1994 |
| famotidine | Gastrointestinal | tablet | 2001 |
| felodipine | Cardiovascular | ER tablet | 2004 |
| fenofibrate | Cardiovascular | capsule | 2002 |
| fexofenadine hydrochloride | Respiratory | tablet | 2005 |
| flecainide acetate | Cardiovascular | tablet | 2002 |
| fluconazole | Anti-Infective | injectable | 2004 |
| fluconazole | Anti-Infective | tablet | 2004 |
| fludarabine phosphate | Oncology | injectable | 2003 |
| fludrocortisone acetate | Endocrine-Metabolic | tablet | 2002 |
| fluoxetine hydrochloride | Central Nervous System | capsule | 2001 |
| flurbiprofen | Central Nervous System | tablet | 1994 |
| flutamide | Oncology | capsule | 2001 |
| fluvoxamine maleate | Central Nervous System | tablet | 2001 |
| fosinopril sodium | Cardiovascular | tablet | 2003 |
| gabapentin | Central Nervous System | capsule | 2004 |
| gabapentin | Central Nervous System | tablet | 2004 |
| ganciclovir | Anti-Infective | capsule | 2003 |
| gemfibrozil | Cardiovascular | tablet | 1993 |
| glimepiride | Endocrine-Metabolic | tablet | 2005 |
| glipizide | Endocrine-Metabolic | tablet | 1994 |
| glyburide | Endocrine-Metabolic | tablet | 1995 |
| guanfacine hydrochloride | Cardiovascular | tablet | 1995 |
| hydroxychloroquine sulfate | Anti-Infective | tablet | 1995 |
| hydroxyurea | Oncology | capsule | 1995 |
| indapamide | Cardiovascular | tablet | 1995 |
| itraconazole | Anti-Infective | capsule | 2005 |
| ketoconazole | Anti-Infective | tablet | 1999 |
| ketorolac tromethamine | Central Nervous System | tablet | 1997 |

*(Table continued on next page.)*

*(Continued from previous page.)*

| Active Pharmaceutical Ingredient | Therapeutic Class | Dose Form | Market Opening |
|---|---|---|---|
| labetalol hydrochloride | Cardiovascular | tablet | 1998 |
| leflunomide | Musculoskeletal | tablet | 2005 |
| leuprolide acetate | Endocrine-Metabolic | injectable | 1998 |
| lisinopril | Cardiovascular | tablet | 2002 |
| lovastatin | Cardiovascular | tablet | 2001 |
| mefloquine hydrochloride | Anti-Infective | tablet | 2002 |
| metformin hydrochloride | Endocrine-Metabolic | tablet | 2002 |
| methazolamide | Ophthalmologic | tablet | 1993 |
| methimazole | Endocrine-Metabolic | tablet | 2000 |
| metolazone | Cardiovascular | tablet | 2003 |
| metoprolol tartrate | Cardiovascular | tablet | 1993 |
| mexiletine hydrochloride | Cardiovascular | capsule | 1995 |
| midazolam hydrochloride | Central Nervous System | injectable | 2000 |
| midodrine hydrochloride | Cardiovascular | tablet | 2003 |
| mirtazapine | Central Nervous System | tablet | 2003 |
| misoprostol | Endocrine-Metabolic | tablet | 2002 |
| moexipril hydrochloride | Cardiovascular | tablet | 2003 |
| mupirocin | Dermatological | topical | 2003 |
| nabumetone | Central Nervous System | tablet | 2001 |
| nadolol | Cardiovascular | tablet | 1993 |
| naltrexone hydrochloride | Central Nervous System | tablet | 1998 |
| naproxen | Central Nervous System | tablet | 1993 |
| naproxen sodium | Central Nervous System | tablet | 1993 |
| nefazodone hydrochloride | Central Nervous System | tablet | 2003 |
| nicardipine hydrochloride | Cardiovascular | capsule | 1996 |
| nizatidine | Gastrointestinal | capsule | 2002 |
| norethindrone acetate | Endocrine-Metabolic | tablet | 2001 |
| ofloxacin | Anti-Infective | tablet | 2003 |
| omeprazole | Gastrointestinal | ER capsule | 2002 |
| oxaprozin | Central Nervous System | tablet | 2001 |
| paclitaxel | Oncology | injectable | 2002 |
| pamidronate disodium | Endocrine-Metabolic | injectable | 2001 |
| paroxetine hydrochloride | Central Nervous System | tablet | 2003 |
| pentoxifylline | Blood Modifier | ER tablet | 1997 |
| pergolide mesylate | Central Nervous System | tablet | 2002 |
| propafenone hydrochloride | Cardiovascular | tablet | 2000 |

*(Table continued on next page.)*

*(Continued from previous page.)*

| Active Pharmaceutical Ingredient | Therapeutic Class | Dose Form | Market Opening |
|---|---|---|---|
| quinapril hydrochloride | Cardiovascular | tablet | 2004 |
| ranitidine hydrochloride | Gastrointestinal | tablet | 1997 |
| ribavirin | Anti-Infective | capsule | 2004 |
| rimantadine hydrochloride | Anti-Infective | tablet | 2001 |
| selegiline hydrochloride | Central Nervous System | tablet | 1996 |
| sotalol hydrochloride | Cardiovascular | tablet | 2000 |
| sucralfate | Gastrointestinal | tablet | 1996 |
| tamoxifen citrate | Oncology | tablet | 2003 |
| terazosin hydrochloride | Cardiovascular | capsule | 1999 |
| terbutaline sulfate | Respiratory | tablet | 2001 |
| terconazole | Genitourinary | topical | 2004 |
| ticlopidine hydrochloride | Blood Modifier | tablet | 1999 |
| tizanidine hydrochloride | Musculoskeletal | tablet | 2002 |
| torsemide | Cardiovascular | tablet | 2002 |
| tramadol hydrochloride | Central Nervous System | tablet | 2002 |
| triazolam | Central Nervous System | tablet | 1994 |
| ursodiol | Gastrointestinal | capsule | 2000 |
| vinorelbine tartrate | Oncology | injectable | 2003 |
| zidovudine | Anti-Infective | tablet | 2005 |
| zonisamide | Central Nervous System | capsule | 2005 |

# Appendix B

# Ancillary Material for Chapter 3

## B.1 Importance Sampling with Change-of-Variables

The importance sampling with change-of-variables method proposed by Ackerberg (2009) proceeds in two steps. In the first step, the econometrician redefines the variable of integration so that all of the unknown parameters of the model are incorporated into it. To fix ideas, consider a simplified representation of the likelihood function for my model:

$$L(\varpi) = \prod_m \int v(\varpi, e) f_e(e) de, \tag{B.1}$$

where $\varpi$ is the vector of unknown parameters and $e$, the variable of integration, is a random error term with known density $f_e(\cdot)$. $v(\cdot)$ is an indicator function that is equal to one if the outcome predicted by the model matches the outcome observed in the data.[1] The product is taken over sample markets.

Now, define $u(\varpi, e)$ as the new variable of integration such that $v(\varpi, e) = \tilde{v}[u(\varpi, e)]$. After the change-of-variables, the likelihood function becomes

$$L(\varpi) = \prod_m \int \tilde{v}(u) \frac{1}{\partial u / \partial e} f_u(u; \varpi) du. \tag{B.2}$$

I assume that $u(\cdot)$ has the form $u(\varpi, e) = \tilde{u}(\varpi) + e$ so that the fractional term drops out. Note that the value of the indicator function $\tilde{v}(\cdot)$ does not depend directly on the parameter vector $\varpi$; only the value of the density $f_u(\cdot)$ changes with $\varpi$.

[1] The function $v(\cdot)$ corresponds to the function $\lambda(\cdot)$ in the main text. For the sake of simplicity, I ignore the equilibrium multiplicity problem here.

In the second step, the econometrician implements importance sampling with respect to the newly defined variable of integration, $u$. Importance sampling is a technique that was developed to numerically approximate an integral when the variable of integration has a distribution that cannot be drawn from easily.[2] For instance, the probability that a random variable $x$, whose density is $f^{\text{T}}(\cdot)$, is less than some value $c$ can be written as $Pr(x < c) = \int_{-\infty}^{c} \frac{f^{\text{T}}(x)}{f^{\text{IS}}(x)} f^{\text{IS}}(x)dx$. The distribution represented by $f^{\text{T}}(\cdot)$ is called the target distribution. $f^{\text{IS}}(x)$ is the density for an alternative distribution called the importance sampling distribution. When it is difficult to take simulated draws from the target distribution, but easy to draw from the importance sampling distribution, $Pr(x < c)$ can be approximated by drawing $\{x^r\}_{r=1}^{R}$ from the importance sampling distribution and calculating the frequency simulator $\frac{1}{R}\sum_{r=1}^{R} \mathbf{1}(x^r < c)\frac{f^{\text{T}}(x^r)}{f^{\text{IS}}(x^r)}$.

Ackerberg's (2009) insight is that importance sampling lets the econometrician draw the variable of integration from a distribution that does not depend on the model parameters. In the preceding example, the likelihood function is modified to

$$L(\varpi) = \prod_m \int \tilde{v}(u) \frac{f_u(u;\varpi)}{f_u^{\text{IS}}(u)} f_u^{\text{IS}}(u)du. \tag{B.3}$$

(B.3) has a computational advantage over (B.1) and (B.2). To obtain estimates based on (B.1), the function $v(\cdot)$ has to be calculated anew for every iteration of the parameter search. Simulation-based estimation via (B.2) also requires evaluation of $\tilde{v}(\cdot)$ at every iteration, because new draws of $u$ are taken for each candidate value of the parameters. On the other hand, estimating the model by the simulated analog of (B.3), $\prod_m \frac{1}{R}\sum_{r=1}^{R} \tilde{v}(u^r)[f_u(u^r;\varpi)/f_u^{\text{IS}}(u^r)]$, requires evaluation of $\tilde{v}(\cdot)$ only at the first iteration; the same draws for $u$ are used throughout the parameter search so that the value of $\tilde{v}(u)$ stays fixed. The only extra requirement is to calculate the ratio of densities, called the "importance sampling weight", at every iteration.

Intuitively, the estimator works by changing the importance sampling weight given to each draw of the random terms as the parameter values change. Each draw is associated with its own predicted outcome, with some draws being associated with predictions that are closer to the observed outcome than others. The parameters that maximize the simulated likelihood function are those that assign higher importance sampling weights to draws whose predicted outcomes are closer to the observed outcome.

While the computational efficiency of importance sampling with change-of-variables makes it attractive, two conditions must be fulfilled to make it work. First, the support of the target distribution $f_u(u;\varpi)$ must not vary with the parameters $\varpi$ (Ackerberg, 2009). When this condition is violated, it is possible that at some parameter values, all draws have an importance sampling weight of zero, which drives the likelihood function down to zero.

The second condition is that the dimensionality of the variable of integration cannot be too high. When $u$ has high dimensionality, the importance sampling weight $f_u(u^r;\varpi)/f_u^{\text{IS}}(u^r)$ tends to

[2] A description of importance sampling can be found in Robert and Casella (2004).

have a very high variance across draws. In the extreme case, one observes some draws having a weight that overflows to machine infinity while for other draws the weight underflows to machine zero. Even if the weights are normalize to take finite values, the high variance among them implies that too much importance is assigned to a small number of draws (Robert and Casella, 2004). When this happens, the value of the likelihood function jumps around greatly in response to small changes in the parameter values, as weight is shifted abruptly from one set of draws to another. The resulting estimates are likely to be unreliable. Some authors have devised methods to implement importance sampling reliably even when the variable of integration has high dimensionality (e.g., Richard and Zhang, 2007). However, these methods require the importance sampling distribution to be updated numerous times during estimation. In the present context, this implies taking new draws for $u$ and running the equilibrium finding algorithm many times. This defeats our purpose of employing importance sampling, which is to economize on computational time.