PERFORMANCE EFFECTS OF IMITATIVE ENTRY

SENDIL K. ETHIRAJ* and DAVID H. ZHU
Stephen M. Ross School of Business, University of Michigan, Ann Arbor, Michigan, U.S.A.

This article examines how waiting to imitate a product affects the performance of the imitator compared to the innovator. Specifically, we address two research questions. Under what conditions does imitation erode the advantage of the innovator? What strategies of imitators help overcome the innovator's advantage? Our main argument is that the increasing availability of information on the innovator's product increases the imitator's returns to waiting. With this increasing availability of information, imitators' products transition from those that are horizontally differentiated (products are similar in quality but differ in their attributes) to those that are vertically differentiated (products differ in quality). Thus, we hypothesize that shifts in the nature of competition over time from horizontal differentiation to vertical differentiation account for why the innovator's advantage is not preserved. Imitation timing simply reflects the uncertainty inherent in imitation efforts. One such uncertainty is the extent of product differentiation that the imitator can achieve. We develop several hypotheses that elaborate this basic intuition. We obtained detailed data on innovator-imitator competition in the branded drug industry to test the hypotheses. All our hypotheses are supported. The main contribution of the article is in showing that the nature of product differentiation in product categories is endogenous to the imitative entry decisions of firms. Copyright © 2008 John Wiley & Sons, Ltd.

INTRODUCTION

How do firms generate and sustain superior performance is arguably a central question in the strategy research enterprise (Nelson, 1991; Rumelt, Schendel, and Teece, 1994). A flip side of this question is what deters imitation efforts that preserve the competitive advantage of leading firms or under what conditions are imitation efforts successful. In explaining imitation deterrence, the extant literature emphasizes two mechanisms: causal ambiguity or uncertainty around the choices of superior performing firms (Barney,

1986; Rumelt, 1984) and complexity and path dependence that impedes piecemeal imitation (Dierickx and Cool, 1989; Rivkin, 2000). A related literature examines the conditions under which early entry advantages are durable or why later entrants (or imitators) never catch up. Though the two literatures share similar concerns with understanding sources of durable performance differences, they have progressed along largely distinct and independent trajectories. In this article, we seek to exploit the common concerns of the two literatures and in the process contribute both to a better understanding of sources of early entry advantages and the conditions under which imitation is effective.

The literature on imitation may be usefully decomposed into two related research questions: why do firms imitate, and what are the performance

effects of imitation. On the question of why firms imitate, there is a large literature spanning strategy, economics, and organization theory and an attendant empirical literature (see Lieberman and Asaba, 2006 for a comprehensive review). In contrast, the question of the performance effects of imitation is much less studied. The explanations for why imitation may be (un)successful are mostly conceptual in nature (see Reed and DeFillippi, 1990 for a review) or based on computer simulations (Ethiraj, Levinthal, and Roy, 2008; Rivkin, 2000). Little or no research has empirically examined what firms choose to imitate and how those choices create performance consequences.

What explains entry timing advantages? Two classes of theories agree on the prediction that early entry is good but differ in their explanation (see Golder and Tellis, 1993 for an exception). Explanations on the *demand side* argue that consumer familiarity with the early entrant's product, uncertainty avoidance in trying later entrants' products, choice of optimal locations in product space by early entrants, or consumer switching costs provide durable advantage to early entrants (Greve, 2000; Schmalensee, 1982). This relationship is particularly strong in markets with horizontally differentiated products, that is, when consumers choose products based on attribute preferences rather than quality differences (Bohlman, Golder, and Mitra, 2002). The *supply-side* theories argue that early entry matters since it provides advantages of size (Klepper, 1996). Large firms have a larger sales base over which their process innovations can be utilized that in turn allows them to lower costs in proportion to their sales. Later entrants are smaller and thus continue to lag the early entrants in cost competitiveness. This literature has a strong empirical tradition (see Lieberman and Montgomery, 1988; 1998 for a review). However, the empirical findings have been mixed with several studies finding no advantages to early entry (see VanderWerf and Mahon, 1997 for a meta-analysis).

The point of departure for this article is, on the one hand, the lack of empirical research on the imitation choices of firms and their performance consequences and, on the other hand, inadequate explanations for the inconsistent empirical findings on the performance effects of entry timing. The entry timing literature is mostly focused on identifying industry characteristics that create early entry advantages while largely ignoring the firm-specific

choices that determine entry timing decisions. In short, the imitation literature is long on theory and short on evidence and the converse is true for the entry timing literature. We saw an opportunity to unite both literatures to address these gaps. In addressing them we ask two research questions. Under what conditions does imitation erode the advantage of the innovator? What strategies of imitators help overcome the innovator's advantage?

We enjoined our interest in these research questions with an empirical puzzle in the branded drug industry. Cohen (2006) documented that the innovator did not always enjoy a market share advantage in new drug product categories. This is surprising especially since theory predicts durable innovator advantages in the pharmaceutical industry because of strong patent protection and the importance of complementary assets in the form of sales forces (Cohen, Nelson, and Walsh, 2000; Teece, 1986). In explaining this anomaly, we focus on the imitation time lag between the innovator and the imitator. We argue that when products are horizontally differentiated (i.e., products differ in their characteristics but are similar in quality) the innovator's advantages are relatively durable. In contrast, when products are vertically differentiated (i.e., products differ in quality) imitators enjoy significant advantages. Thus, we hypothesized that shifts in the nature of competition over time from horizontal to vertical differentiation account for the loss of the innovator's advantage in the branded drug industry. Greater imitation time lag is correlated with greater information about the innovator's drug that the imitator can use to exploit vertical differentiation opportunities. In testing this idea, we exploit another unique feature of the branded drug industry wherein drug information leaks gradually as the drug progresses through various stages of clinical trials (DiMasi, 1995). We show that the success of imitators is linked to the quality of information about the innovator's drug that is available for potential imitators. Thus, our primary explanation for variance in innovator's advantage in the pharmaceutical industry revolves around a transformation of competition over time from one of horizontal differentiation to one of vertical differentiation.

We would like to emphasize three main contributions of this study. First, the literature on entry timing distinguishes between timing advantages within horizontally or vertically differentiated products. It assumes that whether products are

horizontally or vertically differentiated is exogenous to the entry decisions of firms. In this article, we endogenize the nature of differentiation in products to imitators' entry decisions. We argue and show that it is the imitators' entry decisions that create horizontal or vertical differentiation in product categories. Thus, an insight that our analysis affords is that the same product category can transition from horizontal to vertical differentiation depending on the imitation decisions of firms. Second, we show that entry timing is perhaps simply an artifact of the imitation decisions of firms that in turn generate entry timing advantages or disadvantages. Timing is important to the extent that imitating firms have the requisite information about the innovator's product to generate horizontally or vertically differentiated products. This affords a theory of timing advantages rooted in managerial decisions and, thus, contributes to increasing the theoretical rigor of the entry timing literature. Finally, with respect to the imitation literature, we show how uncertainty and its progressive resolution over time create imitation advantages. More important, this represents a first, albeit modest, attempt at empirically examining imitation choices and their performance consequences.

First, we briefly review the literature on entry timing. We then draw on the imitation literature to develop the principal theory and hypotheses and follow that with our description of the research context and the research methods respectively. We then present the results and discussion respectively and provide our conclusion

PRIOR LITERATURE

The literature on entry timing focuses on whether timing has an impact on firm performance and the theoretical basis for such differences. As suggested earlier, the literature may be usefully partitioned into theories focusing on the *supply side* and *demand side* respectively. We briefly review both literatures.

Supply side: experience and entry timing

Klepper (1996) develops a simple model to explain the importance of early entry into a new product category. In his model, a firm introduces a product innovation that induces a set of consumers to buy its product. The product innovation advantage

is temporary and imitators costlessly imitate the innovation with a lag. In parallel, the innovator is also engaged in process innovation that reduces average production cost. The returns to process innovation are increasing in the sales volume or market share of the firm. The imitator entering with a one-period lag attracts a new set of consumers. By implication, later entrants entering the market with an identical offering will garner smaller and smaller shares of the market. A smaller market share also will leave the later entrants at a disadvantage vis-à-vis the innovator since the cost advantage from process innovation is increasing in market share. Thus, early entrants enjoy a durable advantage against new entrants. The empirical evidence is largely supportive of this model (Klepper, 2002).

It is useful to understand the effects of early entry if we relax the two assumptions in Klepper (1996). The assumption of product innovation being costlessly imitable is critical. It implies that the incremental benefits of imitators' products will not cause existing customers to switch because they are costlessly imitated by the innovators, albeit with a lag. Thus, relaxing this assumption can have two effects. One, if imitation is costly, then the imitators are more likely to wait for better information about the innovator's product so that they can improve upon it (Dasgupta, 1988). Two, in the reverse cycle, if the imitators' products are costly for the innovators to imitate, there is an incentive for existing customers of the innovators to switch to the imitators' product. This would undercut the innovator's advantage.

With respect to the assumption of process innovation, it is unlikely to provide an early entry advantage if production cost is only a small proportion of total cost (e.g., if research and development [R&D] comprises a large portion of total costs) Further, if product innovation is costly, then firms may prefer to wait and watch rather than race to the market since competition is based on product quality that cannot be costlessly imitated. Thus, relaxing the twin assumptions in Klepper's (1996) model suggests that the benefits of early entry are contingent. For instance, in the branded drug industry imitation is costly since imitators also have to engage in independent clinical trials. Similarly, process innovation has a negligible effect since production cost accounts for only about 10–15 percent of the total cost of drugs (Berndt *et al.*, 1995).

The current study is by no means the first one to suggest that entry timing is a double-edged sword. Lieberman and Montgomery (1988) outline four factors that can create early mover disadvantages: free rider, shifts in technology, shifts in customer needs, and incumbent inertia (see also Golder and Tellis, 1993). Our contribution to this literature is twofold. First, much of the extant studies have focused on cost-side asymmetries between early and late entrants to draw entry timing implications. Consequently, most of the tests of entry timing effects are directed at the industry level where the industry supplies the variance in costs (see VanderWerf and Mahon, 1997 for a review and critique). In contrast, we show that even when such cost asymmetries within one industry are absent, entry timing effects can still exist. Our second contribution to this literature is in highlighting the managerial decisions that underpin such entry timing effects.

Demand side: early mover advantages

The early literature on entry timing effects argued for a negative relationship between order of entry into new markets and market share in horizontally differentiated experience goods (see Schmalensee, 1982). In horizontally differentiated markets, products offer identical price-performance tradeoffs but differ in the bundle of attributes (Eaton and Lipsey, 1998). These bundles of attributes appeal to different groups of consumers and create imperfect substitution across products. Thus, consumer choice is driven by individual preferences. For instance, products such as video games, cereals, paper towels, and razor blades are examples of experience goods (i.e., product information is obtained only after use) that are horizontally differentiated. The first product in a category induces trial by consumers. Once adoption occurs, consumers have little incentive to switch to the products of later entrants since there is little incremental benefit. Furthermore, in experience goods, because of the uncertainty associated with trying new products, the absence of clear incremental benefit induces consumers to stay with the first product. Several empirical studies fit this theoretical prediction (Huff and Robinson, 1994; Lee *et al.*, 2000; Robinson and Fornell, 1985; Urban *et al.*, 1986), though several other studies have found no relationship and sometimes a positive relationship (see Golder and Tellis, 1993; VanderWerf and Mahon, 1997 for a summary).

The obvious question is whether the relationship between order of entry and market share is robust to the relaxation of the two boundary conditions of horizontal differentiation and experience goods. In the case of vertically differentiated products, there are clear differences in the price-performance tradeoff offered, that is, higher price is associated with higher performance or quality (Eaton and Lipsey, 1998). If quality is observable, ceteris paribus, consumers always prefer the highest-quality product. Thus, consumer choice is independent of individual preferences. Examples of vertically differentiated products include wristwatches (Rolex to Swatch), automobiles (Cadillac to Saturn), and cameras (Hasselblad to Kodak). In vertically differentiated markets, if later entrants offer superior price-performance tradeoffs then consumers have an incentive to switch. This incentive to switch, however, is dampened if the product is an experience good because the higher quality is unknown. Conversely, the incentive to switch is amplified if the product is a search good (i.e., information on product characteristics can be obtained prior to purchase and/or use). In other words, if later entrants offer superior quality products (vertical differentiation) and such quality differences are discernible prior to purchase then the advantage of early entry is fragile. For instance, in a recent paper, Bohlman *et al.* (2002) find that pioneers do worse than later entrants in product categories where quality (vertical differentiation) is more important than variety (horizontal differentiation).

In sum, much of the extant literature has tended to view products either as horizontally differentiated or vertically differentiated and rarely examined transitions from one to the other. Such transitions can have important implications for early entry advantages for firms within a single industry. Moreover, understanding how and why transitions in the nature of differentiation occur will help better understand the managerial decision processes that generate or dampen early entry advantages. This article seeks to address these issues.

THEORY AND HYPOTHESES

Entry timing and imitation benefits

The challenge of imitation and its consequences become interesting only in the presence of uncer-

tainty (Lieberman and Asaba, 2006). Consider, for instance, a firm choosing between an internally generated product development project and one that imitates a competitor's product. Only when the costs and benefits of each project is uncertain will the decision become nonobvious and also present the possibility of observing variation in choices. In general, imitation under uncertainty presents a simple tradeoff (Dasgupta, 1988). If a firm chooses to imitate in the early stages when uncertainty is high, the imitator is subject to the same uncertainties that the innovator faces, which increases the riskiness of the project. The benefit to the imitator, however, is that eventual leadership in the market is highly contested since the innovator does not enjoy a significant monopoly presence in the market that will enable it to erect entry barriers or raise the cost of competition for later entrants such as through patents, exclusive contracts, or spatial saturation (Lieberman and Montgomery, 1988). In contrast, if the imitator chooses to wait until the uncertainty clears, it reduces the riskiness of the investment but also gives the innovator enough time to erect entry barriers and raise the cost of competition for the imitator. It is useful to probe more carefully how this tradeoff affects the nature and quality of products that emerge from imitation.

Consider the contemporary automotive industry. We have two different technologies with different levels of uncertainty competing with the established internal combustion engine—the gas-electric hybrid and fuel cells. The gas-electric hybrid technology pioneered by Toyota presents much less technical uncertainty to a potential imitator but all the post-entry downsides outlined above. Information about the cost and features of the technology is relatively well known as are the attributes desired by customers. In contrast, the fuel cell technology pursued by General Motors presents much greater uncertainty. Both the costs and technical challenges of the technology are unknown, as are the market preferences and the complementary assets required for commercialization (e.g., hydrogen filling stations). What are the implications of choosing the gas-electric or the fuel cell for a potential imitator (e.g., Ford)?

Looking at Toyota's gas-electric hybrid, an imitator faces two choices for imitation. It can choose to horizontally differentiate the product by changing the design of the car (e.g., offering a station wagon) but maintain parity in quality (e.g., fuel

efficiency, acceleration) or choose to vertically differentiate via significant improvements on the principal quality dimensions on which the gas-electric hybrid is evaluated. There are obvious cost-benefit tradeoffs for either choice. The horizontal differentiation option clearly demands less R&D dollars and lower risk in comparison with the vertical differentiation option. However, the tradeoff is in the market potential. The horizontal differentiation option runs up against an established incumbent, and market prospects are unlikely to be significantly superior to that of Toyota. The converse is true for the vertical differentiation option. Improving the principal dimensions of quality is likely to demand much greater R&D investments and entail greater uncertainty in cost, time, and likelihood of success. However, the benefit to this larger investment is higher in that it presents consumers with a clearly superior product in comparison with the product of the established incumbent. Thus, with greater availability of information about the innovator's product, the option to pursue vertical differentiation becomes a real possibility. To the extent that the imitator chooses to pursue vertical differentiation, other things held constant, it has a higher likelihood of outperforming the innovator.

The imitator's options in the case of fuel cells are more constrained. There is much greater uncertainty about the technology and the cost of various choices, as is the market's preferences about the resulting product. The imitator can, at best, glean sketchy information about the path being pursued by the innovator and pursue a similar path. Consequently, the imitator is subject to more or less the same kinds of uncertainties that the innovator faces. The imitator, therefore, is unlikely to be able to use information from the innovator's development effort to develop a superior product. Given the differences in the choices available to the imitator in the gas-electric hybrid versus the fuel cell option, we surmise that the imitator is more likely to come up with a horizontally differentiated product in the fuel cell case.

Stepping back from the examples, it is now possible to develop a simple theory relating imitation timing and the returns to imitation. Imitation is most useful in the face of uncertainty (Lieberman and Asaba, 2006). Imitation reduces costs for the imitator by helping update the beliefs of the imitator using information from the innovator or improving estimates of expected costs and benefits of pursuing a course of action (Cyert and March,

1963). At high levels of uncertainty, the quality of information about the innovator's product is poor and any possibility of using that information to develop a superior product is relatively limited. In contrast, when uncertainty is lower, the quality of information is greater, and this can be used to develop a superior imitative product. Thus, uncertainty about the imitator's product characteristics affects the imitator's opportunities for differentiation. In other words, greater uncertainty about the innovator's product is more likely to result in a horizontally differentiated imitative product, and lower uncertainty is more likely to lead to a vertically differentiated imitative product. If imitation timing is a proxy for the underlying uncertainty about the imitator's product, we might conjecture that the later the imitation timing, the greater the likelihood of a vertically differentiated imitator's product.

The nature of product differentiation affects the returns to imitation (Pepall, 1997). The tradeoff is again between the cost of differentiation and market competition. Horizontal differentiation consumes less R&D dollars since the imitative product is largely similar to the innovator's product. The lower cost of incremental change to achieve horizontal differentiation is traded off against the cost of market competition. Such costs might involve short-term outlays such as advertising or providing free samples to induce trials. It also might involve long-term costs such as price discounting to build market share. In either case, the cost of competition for the imitator should be weighed against the lower cost of horizontal differentiation. In contrast, vertical differentiation consumes more R&D dollars. The upside on the market, however, is that the superior quality makes it easier to induce consumer trial and also opens up the possibility of a price premium vis-à-vis the innovator. Thus, the nature of product differentiation affects the ultimate success of the imitator in the market. Other things held constant, vertically differentiated imitative products have a higher likelihood of outperforming the innovator's product in the market.

Having linked uncertainty and imitation timing to product differentiation opportunities and then product differentiation to returns from imitation, it is now possible to connect all three. Over time, there is greater information about the innovator's product available to potential imitators (Abernathy and Utterback, 1978; Tushman and Anderson, 1986). The greater availability of

information reduces uncertainty and increases the opportunities for vertical differentiation among imitators. Again, if imitation timing is a proxy for uncertainty about the innovator's product, over time, with increased availability of information, the opportunity for imitation changes from horizontal differentiation to vertical differentiation. This change in turn affects the returns to imitation. Thus, products of later imitative entrants are more likely to outperform the innovator's product. This sums up our theory relating imitation timing, nature of product differentiation, and the returns to imitative entry. Thus, we hypothesize that, *ceteris paribus*:

> *Hypothesis 1: The greater the imitation time lag between the innovator's product and the imitator's product, the greater the likelihood that the imitator will beat the innovator in sales.*

Information leakage, opportunities for differentiation, and returns to imitation

In Hypothesis 1 we argued that with the passage of time greater information about the innovator's product becomes available. Such information can form the basis for potential imitators to transform the product category into a vertically differentiated one. If this is true, we should observe the direct impact of differentiation on the likelihood of beating the innovator. Thus, we hypothesize that, *ceteris paribus*,

> *Hypothesis 2: The greater the vertical differentiation of the imitator's product relative to the innovator's product, the greater the likelihood that the imitator will beat the innovator in sales.*

Differentiation, the information leakage trajectory, and imitation benefits

In Hypothesis 1 we argued that a change in the quality of information over time alters the uncertainty associated with imitation. The changing uncertainty in turn presents different opportunities for differentiation and affects the outcome of imitation. This led to the first hypothesis that the greater the imitation time lag between the introduction of the innovator's drug and the imitator's drug, the more likely the imitator will beat the innovator in the market. If more information allows imitators to engage in better product differentiation,

then we should directly observe the returns to differentiation. This led to Hypothesis 2 where we argued that vertically differentiated products are more likely to beat the innovator. Finally, if product differentiation accounts for the positive slope between imitation time lag and the likelihood of beating the innovator, we expect that the positive slope should be dampened when product differentiation is accounted for. In other words, we expect product differentiation to mediate the relationship between imitation time lag and the likelihood of beating the innovator. Thus, we hypothesize that,

> *Hypothesis 3: Product differentiation will mediate the relationship between entry timing and the likelihood of beating the innovator.*

RESEARCH CONTEXT: THE BRANDED DRUG INDUSTRY

The branded drug industry provides a unique context to study firm-level differences in the returns to early entry. Product development in the branded drug industry usually comprises two stages: drug discovery and drug development. See Appendix for a description.

We chose the branded drug industry to test our hypotheses for four reasons. First, in order to study imitation efforts across firms, we needed to identify product categories within which products are largely substitutes (albeit, imperfect) for one another. Following Lu and Comanor (1998), we defined product categories in drugs by targets or mechanisms of action. A drug target is a protein, cell, or human organ that is affected by or is a cause of disease. Drugs that act on a single target via a mechanism of action constitute a reasonably well-defined product category because drugs within the group are close substitutes of one another (Lu and Comanor, 1998). This enabled us to identify clear product categories within which we could examine innovator-imitator dynamics. We defined the first drug in a mechanism of action as *innovator* and all subsequent entrants as *imitators*. More generally, the terms innovator and imitator span a continuum. Anchoring one end of the continuum, a pure imitator is one that completely copies an innovator and produces an identical product. All the information and knowledge that goes into making the imitative product comes

from the innovator. Generic drugs are an example of such pure-type imitation. A pure innovator is one whose product is based on the innovator's own knowledge and information since there are no substitute products available. In between these two ends lie a range of imitative behaviors that draw greater or lesser information from the innovator. Thus, the notion of imitation employed in this article refers to the imitators or later entrants using information from the innovator to guide their imitation decisions. This is consistent with equating informational spillovers with imitation (see Dasgupta, 1988: 74–75).

Second, in order to track information and knowledge spillovers between firms, we needed products or product categories where there is significant uncertainty about the technical characteristics of the product and the market preferences for those characteristics. When imitation is costly and there is uncertainty about product characteristics and market preferences, the imitation decision is far from straightforward. In other words, neither the 'race to imitate' nor the 'wait and watch' strategies are optimal strategies for all firms. This will generate variance in the behavior of imitators on their imitation timing. Without variance in imitation timing we cannot test any of our hypotheses.

Third, we needed to observe leakage or spillover of information about the innovator's product over time. The regulatory review process of drug development (see Appendix) and the long timelines involved enabled relatively accurate inferences about the quality of information about the innovator's drug leaking at different time points. The run-up to R&D and eventual Food and Drug Administration (FDA) approval of the drug is highly uncertain. New targets or mechanisms of action represent new science, and it is incumbent upon the innovating firm to convince the scientific community that the new target is based on research that meets the rigors of science (Polidoro, 2006). This often involves publishing the details of the scientific breakthrough in prestigious journals such as *JAMA (Journal of the American Medical Association), Nature*, or *New England Journal of Medicine*. This provides the basis for observing information leakage. Without a clear information transmission mechanism it is difficult to connect imitation timing outcomes to information leakage about the innovator's product. In addition, we can map the quality of information leakage to different

phases of the clinical trial process. At the conclusion of phase I clinical trials, the only information available is whether or not the drug is safe for use with humans. There is little or no information on the characteristics of the drug or its effectiveness in disease treatment. So any imitation based on this information should result only in horizontal differentiation. In contrast, the conclusion of phase II clinical trials reveals information about the efficacy of the drug and its characteristics. Imitation based on this information affords vertical differentiation. Thus, being able to assess the quality of information leakage over time made the branded drug industry an attractive setting for this study.

Finally, we needed to measure the characteristics of products to distinguish between horizontal and vertical differentiation. Prior studies in this area have simply used different product categories to capture horizontal or vertical differentiation using a dummy variable. Having chosen product categories within which products are close substitutes, it was important for us to be able to measure product characteristics that made them horizontally or vertically differentiated. Product differentiation in the branded drug industry is along two dimensions: efficacy and safety (Yeoh and Roth, 1999). Whereas efficacy is seen from the number and nature of treatment indications for which the drug is approved, safety is seen from the number and nature of side effects that the drug engenders (Berndt *et al.*, 1995). Using these two dimensions, it is possible to characterize horizontal and vertical differentiation. Consider two cholesterol drugs A and B that are horizontally differentiated. Drug A reduces cholesterol in about four weeks and has side effects that include headache and nausea. Drug B reduces cholesterol in about four weeks and causes headache and drowsiness. If the two drugs are identically priced, consumers will choose drug A or B depending on their preferences for nausea or drowsiness. In contrast, the two drugs are vertically differentiated if the number of indications and side effects, respectively, are significantly different. For instance, if drug A lowers only LDL cholesterol and drug B both lowers LDL cholesterol and increases HDL cholesterol, then the two drugs are vertically differentiated. Similarly, if drug A causes headache and nausea and drug B cause headache, nausea, drowsiness, and stomach pains, then quality differences between the two drugs become more apparent. Holding price constant, consumers are always expected to choose drug A over drug B.

These four reasons made the branded drug industry an attractive setting to test the hypotheses presented earlier. The following section outlines the empirical methodology that we employed.

METHODS

Sample and data

The data for the sample was compiled from multiple sources. We first obtained the full list of therapeutic classes and mechanisms of action for existing branded drugs from *Mosby's Rx*, and *Drug Facts and Comparisons*, two major physicians' reference directories of drug information. We used each mechanism of action to identify a distinct product category.[1]

We then obtained drug sales ranks from *Med Ad News*, a trade journal that tracks the pharmaceutical and biotech industries. For the period 1994–2004, we obtained an annual list of the top 200 drugs by sales revenue in the global market. This source did not compile drug sales information prior to this period. The top-selling drug in any year was Lipitor with sales revenues of $11,594M in 2004. The lowest-selling drug in the top-200 list in any year was Alesse with sales revenues of $20M in 1997. This wide range of dollar sales volumes indicates that branded drugs with even modest sales were captured in our dataset.

We then matched this list of drugs to a list of all drugs approved by the FDA and listed in the Orange Book.[2] The Orange Book yielded data on FDA approval dates, names of firms that received drug approval, drug therapeutic characteristics, and patents that protect the drug from generic competition.[3]

---

[1] We did not include biological drugs or generic drugs in the sample. The information leakage in biological drugs over the phases of clinical trials is not comparable with small molecule drugs, which complicates comparisons. We also excluded generic drugs since they are identical to their branded drug counterparts (i.e., not differentiated). The list of targets that was included in the study is available from the authors.

[2] The publication, *Approved Drug Products with Therapeutic Equivalence Evaluations* (commonly known as the Orange Book), identifies drug products approved by the FDA on the basis of safety and effectiveness under the Federal Food, Drug, and Cosmetic Act.

[3] Of the 309 drugs for which we had sales rank data, 54 drugs had two corresponding 'trade names' in the Orange Book. For

We collected data on the indications and side effects for all drugs in the sample from three independent sources: *Mosby's Rx*, Micromedex, and Efacts. All three databases are physician reference sources that list important drug characteristics such as approved indications, known side effects, dosage, and package labeling. We extracted the indications and side effects respectively for each drug. We purchased data on drug level detailing expenditures from Verispan.[4] The detailing data (for the United States) spanned the period 1991–2004 for the top-500 drugs by dollar sales revenues. Firm-specific data such as sales, R&D expenditures, employees, and assets were compiled from the Compustat database.

Identifying the innovators and imitators involved mapping drugs to dates. This could be done in two ways: (1) identify the date of FDA approval of each drug; (2) identify the date of commencement of drug development for each drug. Ideally, the study demands the identification of dates that correspond to the start of drug development. Unfortunately, this date was difficult to identify, so we used the FDA approval date to identify innovators and imitators. We attempted to collect patent listing information for all drugs in the sample. Since patent filing constitutes the first identifiable date for commencement of a drug development program, it provides a good approximation for when drug development effort began.[5]

The patent listing information is provided in the Orange Book released by the FDA each year. The listing of drug patents in the Orange Book is mandated by the Hatch-Waxman Act of 1984 that regulates generic drug competition. We collected all the Orange Books published since 1984 and compiled the patent listing information for all the drugs in our sample.[6]

Our final sample comprised 171 drugs in 14 therapeutic classes across 32 mechanisms of action. This included 912 observations spanning the sales years 1994–2004 and drugs approved during the period 1984–2004.

Variables

Dependent variable

Beat innovator. For each year we ranked all drugs within a product class (i.e., mechanism of action) by sales revenues. Each imitator drug for each year was coded zero if its sales rank was lower than that of the innovator and coded one if its sales rank was higher. In all cases, when the imitator beat the innovator, it also became the market share leader in the product class.

Explanatory variables

Imitation lag. Imitation time lag was measured as the time difference in years between the FDA approval date of the innovator drug and the approval date of the imitator drug. Consistent with Hypothesis 1, we expect the sign on this coefficient to be positive.

Product differentiation. Product differentiation was measured using two variables—indications and side effects respectively offered by the imitator's product in comparison with the innovator.

Difference in indications. For each imitator drug, let *NewInd* be the indications offered by the imitator drug but not the innovator drug, *RedInd* be

---

example, AVIANE corresponds to AVIANE-21 and AVIANE-28 in the Orange Book. Twenty other drugs correspond to more than two 'trade names' in the Orange Book. We merged all such variants into a single drug in order to obtain a match with the sales rank data. We also found that some drugs were issued to subsidiaries of a company. We used Hoover's Online and *Who Owns Whom* to identify the parent companies. We also referred to Compustat, Hoover's Online, and company annual reports to identify any mergers or acquisitions among our sample firms. We treat firms as separate entities before a merger or acquisition.

[4] Verispan, a healthcare information company, is the leading provider of de-identified patient-centric, longitudinal data.

[5] We did not adopt patent dates as the primary identification procedure because we found that several drugs in our final sample did not have patents associated with them. This made it difficult to precisely identify the date of commencement of drug development. Hence we retained drug approval dates as a proxy for the innovator and imitator respectively and did the patent-based analysis as a robustness check.

[6] We identified the patent application dates for all the drugs in our sample to map correspondence between first to FDA approval and first to commence development. We found that many drugs did not have corresponding patents. For 149 of the 171 drugs that had patents associated with them, the innovator defined by FDA approval date also matched the innovator by patent application date, i.e., 87 percent of the sample was unaffected. Of the 22 drugs (11 pairs) whose status as innovators or imitators were affected, we found that three pairs of drugs had less than one year difference in patent application dates, another three pairs had a difference of less than two years, and the remaining five pairs had an average difference of seven years. Our hypotheses tests may be confounded if the last group of five pairs of drugs is driving the results. So for robustness we reran all the analyses by altering the classification of the innovator and the imitator in our final sample whenever applicable. The results were unchanged. This gave us confidence that our approach to classifying innovators and imitators does not account for the results.

the indications offered by the innovator drug but not the imitator drug, and *InnTotInd* be the total number of indications offered by the innovator drug. Difference in indications was measured as *(NewInd-RedInd)/InnTotInd*. We expect the sign on this coefficient to be positive.

Difference in side effects. For each imitator drug, let *NewSE* be the side effects engendered by the imitator drug but not the innovator drug, *RedSE* be the side effects in the innovator drug but not the imitator drug, and *InnTotSE* be the total number of side effects in the innovator drug. Difference in side effects was measured as *(NewSE-RedSE)/InnTotSE*. We expect the sign on this coefficient to be negative.

Control variables

We included a variety of controls that are described in Table 1. The controls may be subsumed into imitator characteristics (size, detailing, R&D capabilities, non-U.S. sales dummy), innovator characteristics (innovator patent protection, innovator beat itself, pre-1984 approval of innovator drug), competition (entry order, coefficient of variation in indications and side effects, respectively), market size (therapeutic class dummies), and time-specific effects (sales year dummies).

Model specification and estimation

We estimated a probit choice model of the likelihood that an imitator will beat the innovator in a drug class in sales in any given year. We included three sets of predictors—imitation time lag, product differentiation, and other controls to account for alternative explanations. Thus, the probit equation was,

$$P(z_{ijt} = 1) = \alpha_t + \tau_j + \beta m_{ij} + \delta p d_{ij}$$
$$+ \gamma c_{ij} + u_i \qquad (1)$$

where, z_{ijt}, denotes imitator drug i beating the innovator in drug class j in sales year t, m_{ij} is the imitation time lag between approval of the innovator's drug in drug class j and the approval of the imitator drug i in the same drug class j, pd_{ij} is a vector of product differentiation covariates, c_{ij} is a vector of other controls, α_t the sales year dummies, τ_j the therapeutic class

dummies, and $u_i \sim N(0, 1)$ the error term. Since firms could be imitators on multiple drugs across different drug classes, we report robust standard errors.

RESULTS

Table 2 presents the descriptive statistics and correlation matrix of variables employed in the regression estimation. About one-fourth of the imitators across the 32 new drug classes beat the innovator in the class. The average timing lag between the entry of the innovator and the imitator was 6.8 years. The correlations were all in the expected direction.[7] Figure 1 presents the mean ratio of indications and side effects respectively for imitator drugs that beat the innovator and those that did not. The graph confirms our basic hypothesis that differentiation in imitator drugs accounts for differences in whether imitators beat the innovator.

Table 3 presents a series of nested probit regression estimates of the likelihood that an imitator beats the innovator. Model 1 presents the results of a specification that includes only the control variables. All signs on the controls were in the expected direction. Larger firms were more likely to beat the innovator. Consistent with prior research, later entrants were less likely to beat the innovator (Robinson and Fornell, 1985; Urban *et al.*, 1986). This is because multiple entries fragment the total market for the drug and reduce the
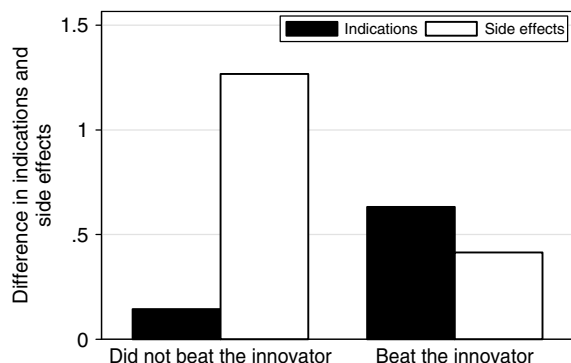

Figure 1.   Differentiation of imitator drugs

---

[7] The correlation between entry order and imitation lag is 0.71. A correlation of this magnitude inflates the standard errors in our coefficients and thus biases our hypotheses tests to be conservative (Chatterjee, Hadi, and Price, 2000).

Table 1. Description of control variables

Variables	Measurement details	Notes
Beat innovator U.S.	0, if the imitator's U.S. drug sales were smaller than that of the innovator's U.S. drug sales in a given year; 1 if the imitator's drug sales were larger	Alternative DV to assess robustness
Difference in drug sales	Global sales of imitator's drug minus global sales of innovator's drug (in $million)	Alternative DV to assess robustness
Indication ratio	Number of indications of a drug divided by the total number of indications offered by all existing drugs in a given market-year	Used to construct variable 'variance in indications'
Side effect ratio	Number of side effects of a drug divided by the total number of side effects reported by all existing drugs in a given market-year	Used to construct variable 'variance in side effects'
Variance in indications	$\frac{SD(indication\_ratio)}{Mean(indication\_ratio)} \times \frac{N}{2(N-1)}$	Captures competition in drug quality (on indications) among all drugs in a given mechanism of action. Based on the Gini coefficient (see Theil, 1967)
Variance in side effects	$\frac{SD(sideeffect\_ratio)}{Mean(sideeffect\_ratio)} \times \frac{N}{2(N-1)}$	Captures competition in drug quality (on side effects) among all drugs in a given mechanism of action. Based on the Gini coefficient (see Theil, 1967)
Non-U.S. sales dummy	Dummy variable. Equals 1 if the drug has over 10% sales outside the U.S.	Account for market share differences due to cross-country differences in revenues
Imitator R&D intensity	Imitator firm's R&D expenditure divided by total sales	Included in regression of differentiation on imitation lag
Patents	Moving three year sum of patents filed by the firm in the corresponding therapeutic class	Included in regression of differentiation on imitation lag
Publications	Moving three year sum of scientific publications of the firm	Included in regression of differentiation on imitation lag
Detailing (000s)	Annual detailing expenditures on the imitator drug ($thousand)	Measuring marketing differentiation
Log(sales)	Log of contemporaneous firm sales for the imitator	Measuring firm size differences
Entry order	Time based rank ordering of FDA approval time	Measuring entry order and competition effects
Innovator patent protection	Remaining patent protection years for the innovator	Measuring patent protection of the leader drug
Innovator beat itself	Dummy variable. Equals 1 if the same firm produced an imitator drug that beat its own innovator drug in sales	Capturing cases where innovator launched a new drug that beat its own drug
Pre-1984 innovator	Dummy variable equals 1 if the innovator drug was approved before 1984	Captures effect of drug approval in the pre-Hatch-Waxman Act era
Sales year dummies	Sales year dummies	Time-specific effects
Therapeutic class dummies	Therapeutic class dummies	Control variable, capturing market size differences

likelihood that the later entrants will beat the innovator. We also find that the greater the remaining patent life of the innovator's drug, the lower the likelihood that the imitator will beat the innovator. This is because the innovator's incentive to defend market share via detailing and advertising is increasing in the remaining useful patent life on the drug. Finally, we found a positive likelihood that an innovator will beat itself though it was not statistically significant.

Model 2 includes imitation time lag. The coefficient on imitation time lag is positive and statistically significant. All other variables held at their means, one additional year increase in the imitation

Table 2.   Descriptive statistics and correlation matrix of variables employed in estimation (N = 912)

	Variable name	Mean	SD	1	2	3	4	5	6	7	8	9	10
1	Beat innovator	0.25	0.44	1									
2	Imitation Lag	6.84	4.69	−0.138	1								
3	Difference in indications	0.27	1.08	0.194	0.087	1							
4	Difference in side effects	1.08	7.27	−0.054	0.067	−0.032	1						
5	Detaling (000s)	9.81	24.72	0.238	−0.020	0.091	−0.029	1					
6	Log(Sales)	8.93	1.83	0.231	−0.155	0.018	0.089	0.142	1				
7	Entry order	5.67	4.09	−0.160	0.717	0.065	−0.097	−0.153	−0.143	1			
8	Variance in indications	0.32	0.20	0.008	0.208	0.188	−0.022	0.006	−0.245	0.288	1		
9	Variance in side effects	0.34	0.24	−0.097	0.118	−0.036	0.154	−0.107	−0.204	−0.050	−0.037	1	
10	Innovator patent protection	0.68	8.49	0.029	−0.248	−0.033	−0.155	0.206	0.028	−0.353	−0.262	0.100	1
11	Innovator beat itself	0.11	0.32										
12	Pre-1984 innovator	0.01	0.10										
13	Global sales dummy	0.15	0.36										

Note: Correlations greater than 0.06 are significant at the .05 level or less.

time lag beyond the mean increases the probability of beating the innovator by about 1.4 percent. This provides baseline support for Hypothesis 1. Models 3 and 4, respectively, include the product differentiation variables. Model 5 presents the full model. The sign on the coefficients of the controls were all in the expected direction. We find that the difference in the number of indications is positive and statistically significant. All other variables held at their means, one percentage point increase in the number of unique indications offered by the imitator drug beyond the mean increases the probability of beating the innovator by 7.1 percent. The coefficient on difference in side effects is negative and statistically significant. All other variables held at their means, one percentage point increase in the number of unique side effects in the imitator drug beyond the mean decreases the probability of beating the innovator by 0.6 percent. Figures 2 and 3 plot the marginal effects of the two differentiation variables on the probability of beating the innovator. Each figure plots three lines. One sets all other variables in the regression at the mean in the dataset, the second sets the other variables at their minimum, and the last sets the other variables at their maximum. These graphs show how the impact of differentiation varies over the range of observed data. As can be seen, the impact of indications is robust in the full range of observed data. In contrast, the impact of side effects is somewhat weaker since its effect depends on the range of the other variables in the regression equation. Taken together this provides strong support for Hypothesis 2.

Hypothesis 1 argued that increasing information about the innovator's drug that is available over time allows imitators to vertically differentiate their products. Since this implies that as long
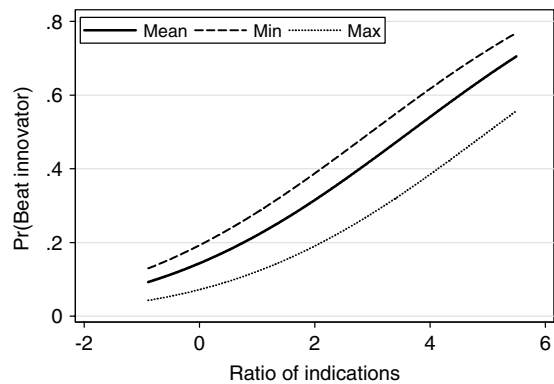

Figure 2.   Marginal impact of indications

Table 3.　Probit estimates for probability of beating the innovator

	(1) Beat innovator	(2) Beat innovator	(3) Beat innovator	(4) Beat innovator	(5) Beat innovator	(6) Beat innovator	(7) Beat innovator	(8) Beat innovator	(9) Beat innovator
Imitation lag (+)		0.055* (0.024)	0.032 (0.024)	0.052* (0.024)	0.027 (0.024)	0.068 (0.056)	0.122+ (0.065)	0.008 (0.061)	0.121+ (0.069)
Difference in indications (+)			0.283** (0.061)		0.292** (0.061)			0.876** (0.133)	0.192* (0.084)
Difference in side effects (−)				−0.020* (0.009)	−0.025* (0.010)			−0.047** (0.017)	−0.036* (0.018)
Log(Sales)	0.227** (0.052)	0.230** (0.053)	0.237** (0.055)	0.236** (0.055)	0.245** (0.057)	0.132* (0.064)	0.437** (0.112)	0.067 (0.059)	0.511** (0.111)
Variance in indications	−0.243 (0.376)	−0.116 (0.379)	−0.286 (0.384)	−0.164 (0.376)	−0.353 (0.382)	−0.991+ (0.520)	2.891** (0.919)	−1.769** (0.524)	1.159 (0.766)
Variance in side effects	0.434 (0.295)	0.513+ (0.287)	0.518+ (0.299)	0.483+ (0.282)	0.497+ (0.289)	0.642* (0.313)	−1.752+ (0.928)	0.785* (0.338)	0.313 (0.660)
Non-US sales dummy	0.551** (0.161)	0.489** (0.165)	0.515** (0.168)	0.497** (0.166)	0.527** (0.169)	0.089 (0.216)	1.506** (0.322)	0.193 (0.239)	0.633 (0.522)
Entry order	−0.060** (0.019)	−0.111** (0.028)	−0.103** (0.029)	−0.112** (0.028)	−0.103** (0.028)	0.020 (0.066)	−0.166** (0.053)	0.104 (0.068)	−0.190** (0.070)
Detailing (000s)	0.011** (0.003)	0.011** (0.003)	0.012** (0.003)	0.011** (0.003)	0.012** (0.003)	0.013** (0.003)	0.010* (0.005)	0.020** (0.003)	0.015** (0.005)
Innovator patent protection	−0.042** (0.009)	−0.040** (0.009)	−0.049** (0.010)	−0.044** (0.009)	−0.053** (0.010)	−0.051** (0.012)	−0.023 (0.024)	−0.084** (0.014)	−0.039 (0.032)
Innovator beat itself	0.024 (0.155)	0.075 (0.158)	0.104 (0.157)	0.041 (0.159)	0.064 (0.158)	0.104 (0.194)	0.198 (0.313)	−0.169 (0.214)	0.157 (0.321)
Pre-1984 innovator drug	0.768+ (0.437)	0.806+ (0.438)	−0.047 (0.475)	0.762+ (0.438)	−0.129 (0.477)	1.224* (0.482)		−1.308* (0.603)	
Sales year dummies	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Therapeutic class dummies	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Constant	−2.541** (0.619)	−2.667** (0.638)	−2.709** (0.657)	−2.680** (0.643)	−2.728** (0.664)	−2.131** (0.781)	−6.249** (1.245)	−1.577* (0.730)	−6.951** (1.238)
Observations	912	912	912	912	912	475	437	475	437
Pseudo R-squared	0.280	0.286	0.303	0.288	0.307	0.226	0.524	0.297	0.492
Pseudo log likelihood	−371.51	−368.691	−359.453	−367.391	−357.741	−228.058	−102.093	−207.072	−107.468

Robust standard errors in parentheses. Pre-1984 innovator drug dropped out in Models 7 and 9 due to collinearity with the year dummies
Two-tailed tests. + significant at 10%; * significant at 5%; ** significant at 1%
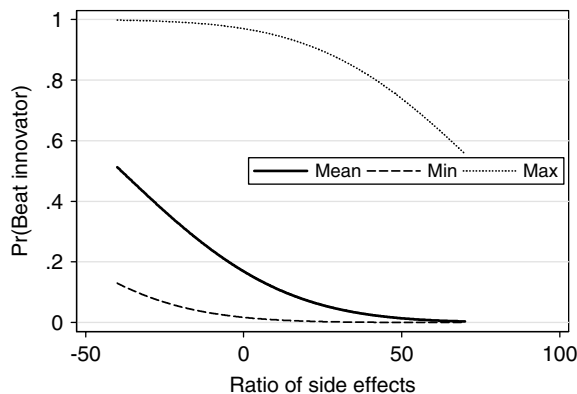
Figure 3.   Marginal impact of side effects

as products are horizontally differentiated the likelihood of beating the innovator is no better than random, we expect the regression line for imitation time lag to be flat (i.e., parallel to the x-axis) until there is enough information about the innovator drug to afford vertical differentiation. From this point on we expect the slope to turn positive. This relationship cannot be teased out from the imitation lag variable. In order to identify whether the pattern was indeed as we expected, we split the sample into two: imitation time lag was seven years or less and greater than seven years respectively. This marks the conclusion of phase II clinical trials when the quality of information about the innovator's drug allows imitators to pursue vertical differentiation (DiMasi, 2001a). Models 6–9 present the results of the split sample regressions. Model 6 presents estimates for the sample when imitation lag is seven years or less. As expected, imitation lag is nonsignificant. In contrast, when imitation lag is greater than seven years (Model 7), it is positive and statistically significant. This supports our argument that the quality of information leakage after phase II clinical trials allows vertical differentiation. Models 8 and 9 include the product differentiation variables to the two subsamples and they continue to be significant. This lends further support to the theory underpinning Hypothesis 2.

Finally, turning to the mediation hypothesis, the results were supportive. In Model 5, adding the product differentiation variables causes the imitation time lag coefficient to turn nonsignificant. This suggests complete mediation of imitation lag. We also regressed indications and side

effects on imitation lag.[8] Models I–J in Table 4 present the results of firm fixed-effect regressions and Models K–L the random-effect regressions. We included two measures of capabilities (count of firm patents in the therapeutic class and count of firm-scientific publications) to account for the fact that the availability of information is not sufficient for a firm to be able to exploit it. The requisite absorptive capacity is a necessary condition (Cohen and Levinthal, 1990). Though the coefficients on both variables are in the expected direction, they are statistically nonsignificant. We conjecture that the nonsignificance is due to measurement error in the two variables, that is, they do not precisely measure the capabilities specific to the drug being imitated. Consistent with prior research, in the absence of good measures rooted in the context, we are unlikely to discern the importance of capabilities (Ethiraj, Kale, Krishnan, and Singh, 2005). As expected, imitation lag is positive and significant in the case of indications and negative in the case of side effects in both fixed-effect and random-effect models. The causal test for mediation outlined in Baron and Kenny (1986) was completely met: (1) imitation time lag is statistically significant in the absence of the differentiation variables (Model 2); (2) regression of the differentiation variables on imitation time lag yielded statistically significant coefficients for both indications and side effects ($p < 0.000$) suggesting that these two variables are the primary drivers of mediation; (3) the differentiation variables are statistically significant; and (4) imitation time lag turns nonsignificant when the differentiation variables were added to the equation. This suggests that the differentiation variables completely mediate the relationship between imitation time lag and the likelihood of beating the innovator (MacKinnon *et al.*, 2002). Thus, we found strong support for Hypothesis 3.

---

[8] Regressing indications and side effects on imitation lag implies that indications and side effects are endogenous to the system. To alleviate such concerns of endogeneity, we estimated a three-stage least squares model with three equations. The first two equations were Models I and J in Table 4. These two equations were identified with R&D intensity, firm patents in the therapeutic class, firm publications, and the approval year dummies. The third equation was Model B in Table 4. The results were robust in this specification as well. This alleviates concerns about the endogeneity between imitation time lag and the product differentiation measures. These results are available from the authors.

Table 4. Alternative specifications for robustness

	(A) Beat innovator FE	(B) Beat innovator FE	(C) Beat innovator RE	(D) Beat innovator RE	(E) Beat innovator GLS-PH	(F) Beat innovator GLS-PH	(G) Beat innovator GLS-PA	(H) Beat innovator GLS-PA	(I) Indications FE	(J) Side effects FE	(K) Indications RE	(L) Side effects RE
Imitation lag (+)	33.604* (15.614)	3.060 (16.771)	26.617+ (14.006)	−2.520 (15.102)	9.359* (4.521)	−3.590 (5.217)	45.999** (16.644)	5.980 (18.491)	0.029** (0.006)	−0.285** (0.093)	0.029** (0.006)	−0.293** (0.078)
Difference in indications (+)		506.106** (65.721)		273.641** (43.974)		159.146** (19.786)		302.780** (34.969)				
Difference in side effects (−)		−13.339* (5.242)		−15.092** (5.066)		−8.007** (1.414)		−9.938* (4.146)				
Log(Sales)	−30.636 (104.366)	−56.058 (99.960)	39.257 (26.633)	39.594 (28.630)	23.748** (5.716)	39.417** (6.642)	21.299 (20.429)	35.974+ (18.758)	−0.045 (0.039)	0.644 (0.566)	−0.048 (0.036)	0.810** (0.175)
Variance in indications	−400.566+ (235.995)	−800.304** (273.783)	−289.809 (210.978)	−1396.873** (268.710)	−51.711 (77.569)	−524.073** (109.710)	10.299 (226.430)	−432.953+ (251.020)	−0.691** (0.130)		−0.654** (0.129)	
Variance in side effects	172.012 (173.763)	234.592 (184.598)	113.846 (166.785)	208.630 (179.675)	19.459 (44.522)	107.480 (75.676)	141.276 (206.713)	602.759** (217.611)		−0.361 (1.402)		−0.104 (1.316)
Entry order	−26.640+ (14.657)	−16.660 (15.205)	−27.017+ (13.819)	−16.155 (14.389)	−10.235** (3.762)	−10.727* (4.933)	−42.763** (14.458)	−15.535 (18.305)				
Detailing (000s)	9.578** (1.632)	8.810** (1.614)	8.947** (1.559)	8.175** (1.551)	3.962** (1.072)	4.861** (1.149)	3.394** (0.867)	2.859** (0.859)				
Innovator patent protection	−43.523** (5.777)	−52.981** (5.977)	−40.007** (5.270)	−42.617** (5.781)	−10.416** (2.037)	−22.742** (2.918)	−20.480** (6.394)	−30.739** (6.493)				
Innovator beat itself	260.624+ (135.111)	213.825 (131.778)	196.685 (122.319)	172.928 (121.751)	68.055* (28.238)	18.695 (29.885)	166.951 (118.214)	144.042 (134.176)				
Pre-1984 innovator drug	993.277** (371.406)	−518.892 (407.273)	745.579+ (360.078)	132.644 (374.827)	328.107** (89.084)	−159.699 (108.232)	502.253 (366.823)	−539.384 (416.690)	2.935** (0.199)	−0.671 (2.895)	2.937** (0.199)	−1.744 (2.526)
Imitator R&D intensity									−0.111 (0.127)	1.031 (1.846)	−0.101 (0.125)	1.883 (1.237)
Patents									0.0002 (0.0008)	−0.002 (0.011)	0.000 (0.001)	−0.001 (0.006)
Publications									0.0000 (0.0001)	−0.001 (0.002)	0.000 (0.000)	−0.001+ (0.001)
Sales year dummies	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Therapeutic class dummies	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Approval year dummies									Incl.	Incl.	Incl.	Incl.
Constant	123.293 (984.032)	566.729 (950.382)	−511.546+ (289.601)	401.968 (316.974)	−359.870** (70.676)	−53.482 (100.079)	−251.783 (219.137)	−390.964 (239.723)	0.526 (0.358)	4.491 (5.167)	0.604+ (0.353)	−6.388** (1.969)
Observations	912	912	912	912	912	912	912	912	912	912	912	912
R-squared	0.216	0.286							0.740	0.300		

Standard errors in parentheses. R-squares for the random effects models and the GLS models are not computed.
Two-tailed tests. + significant at 10%; * significant at 5%; ** significant at 1%

Robustness tests

We performed a variety of robustness tests to establish confidence in the results. Table 4 presents the alternative specifications. Columns A and B present the results of a firm fixed-effects regression using a continuous dependent variable (imitator drug sales in dollars minus innovator drug sales in dollars).[9] The results are qualitatively similar to the probit results presented earlier. Columns C and D present the same results using firm random effects. Since the dependent variable is annual drug sales, there is the possibility of serial correlation. To account for this, we estimated generalized least squares models with a correction for drug panel heteroskedasticity. These models are presented in Columns E and F. We estimated the same models allowing for correlations within drug panels. These results are presented in Columns G and H. Once again the results are consistent across all these variants, giving us confidence that the results are robust.

Finally, we reran all the models using patent filing dates to construct the imitation time lag variable to ensure that the unobserved factors that affect market launch dates were not driving the results. All the results were qualitatively similar to that in Table 3. Also, to ensure that the results are not simply a function of the estimation method, we reran all models using logit estimation. All the results were again largely similar to that in Table 3.

DISCUSSION

The primary objective of this study was to employ a theory of information-based imitation to explain the inconsistent empirical findings on the relationship between entry timing and the performance of imitators. In addressing this question, we saw an opportunity to resolve an empirical puzzle in the branded drug industry where an innovator's advantage was not always durable. In contrast with the existing literature, we put forward a positive relationship between imitation time lag

and the likelihood of beating the innovator. The theoretical basis lies in the amount of information on the innovator's drug that is available to imitators. In the early stages, with relatively sparse information on the innovator's development program, any imitation efforts will yield only horizontal differentiation. Over time, however, as greater information about the innovator drug leaks out the potential for vertical differentiation increases, giving the imitator significant advantages over the innovator. In other words, our principal answer to the question of why the innovator's advantage in the branded drug industry is not always durable is that competition within a new product category in the branded drug industry shifts over time from horizontal differentiation to vertical differentiation. Whereas horizontal differentiation does little to threaten the innovator's advantage, vertical differentiation benefits the imitator. We developed several hypotheses to explore this intuition. All the hypotheses were supported.

Our results offer three important takeaways for understanding industry structure, the dynamics of innovator-imitator competition, and the literature on entry timing respectively. First, much of the entry timing literature assumed that products are either horizontally or vertically differentiated and examined the effects of entry timing on the imitator's performance. This approach assumes that the nature of product differentiation is exogenous to the entry decisions of firms. We argue and show that imitative entry decisions are conditioned on the amount and quality of information available about the innovator's product. This information shapes choices about product differentiation for the imitators and thus affects the outcome of whether product categories are horizontally or vertically differentiated. We outline a theory of imitative entry that endogenizes the nature of product differentiation and shows how the same product can transition from horizontal to vertical differentiation over time.

A related issue of industry structure is whether firms specialize as innovators or imitators? Is there any evidence for specialization that would suggest such choices may be endogenous to firm capabilities? Within the branded drug industry, the answer is an unequivocal no to the question of specialization. Innovators and imitators often switch roles across drug product categories. That we find no evidence for specialization is not surprising. For

---

[9] We did not use difference in sales revenues as the primary specification for three reasons: (1) we had to make assumptions about sales revenues when data was unavailable, (2) using sales revenues entails demand estimation via a system of simultaneous equations, which is challenging in the pharmaceutical industry (see Berndt *et al.*, 1995), and (3) given the *ex ante* nature of the imitation decision, our theory better predicts the likelihood of beating the innovator in sales rather than by what magnitude.

imitators to be able to use the information leaked from the innovator they need the requisite absorptive capacity (Cohen and Levinthal, 1990) that is built up over several years of cumulative investments. For instance, upon observing an innovator filing a patent on a compound, only those imitators that have built up large banks of molecules can generate a variant of the innovator's molecule. Without the ability to generate such variants, the innovator's patent can be a strong entry barrier. We sought to measure differences in capabilities using the firm's cumulative record of patents and publications. But we did not find statistical support for the measures, likely because of measurement error.

Second, with respect to the dynamics of innovator-imitator competition, much of the theoretical and empirical literature on entry timing has examined industries and/or products that are characterized by either horizontal or vertical differentiation respectively. We are not aware of any studies that examine industry transitions from horizontal to vertical differentiation or vice versa and how it affects early entry advantages. This study showed that such transitions have important implications for the advantage of the innovator. Turning this result into a prescriptive lesson, it appears that one way for imitators to overcome the advantage of the innovator in horizontally differentiated product categories is to attempt to transform them into vertically differentiated ones. In other words, imitators might benefit from shifting the consumers' focus on variety or attribute differences (i.e., horizontal differentiation) to a focus on quality or performance differences (i.e., vertical differentiation). It appears that vertically differentiating a product is an important lever for imitators to alter the advantage of the innovator and turn it into a pioneering cost.

Third, there is the question of what our study offers to the well-established empirical literature on entry timing. We believe that our study both embraces and extends it. In embracing it, we do find support, albeit weak, for the negative relationship between order of entry and the likelihood of beating the innovator (a proxy for market share). This is because with increasing entry into a product category, there is fragmentation of total revenues that decreases the chances that later movers will dislodge the innovator. We extend the literature by showing that timing of imitation does make a crucial difference to the success of imitation efforts.

That said, timing is itself an explanation that is disembodied from managerial actions and strategies. We show that the effects of entry timing are simply an artifact of important managerial actions that create or undermine timing advantages. More generally, such timing differences embody uncertainties inherent in imitation efforts that in turn affect the cost versus benefit of imitation. In this article, we show that one such uncertainty is the nature of differentiation that is achievable by the imitator. Thus, our important contribution to the literature on entry timing is in uncovering some important firm strategies that underpin the relationship between entry timing and competitive (dis)advantage.

Finally, a practical implication of our study is that imitators will be better off delaying imitation until they obtain useful information about the innovator's drug that will enable vertical differentiation. While this will ensure that they will surpass the innovator in terms of market share, it is unclear whether this strategy is also profit maximizing. In other words, is it ever too late to imitate? Answering this question turns on the issue of the life cycle of the product category. If the product category has a finite life and it is likely to be substituted by another product category, then waiting for better information on the innovator's drug will be counter-productive from the standpoint of profit maximization. Even though the imitator might come up with a vertically differentiated product, it may not be profitable if the entire product category is substituted by a new one. In contrast, if product categories coexist in the market (i.e., they are complements) then waiting to get information to develop a higher-quality drug will also be profit maximizing for the imitator since the drug category does not have a finite life cycle. The life cycle of a drug within a product category is simply a function of remaining patent life before generics enter the market. Moreover, if there are generic substitutes for the early entrants, the importance of a higher-quality imitative drug is amplified.

CONCLUSION

Having concluded the study of entry timing and imitation in the branded drug industry, it is useful to address the question of generalizability or the boundary conditions under which we expect our

results to hold. We believe that there are at least two boundary conditions that merit close scrutiny. First, the branded drug industry is characterized by long life cycles of development that span 10 years or more. During this long run up there is gradual leakage of information about the innovator's effort that can aid the imitator's decision calculus. The thought experiment to consider is whether our results will hold in products with shorter development life cycles. We surmise that life cycles should have little impact on our results. Even in shorter life cycle industries if imitators can transform an industry from horizontal to vertical differentiation, we expect innovators' advantages to erode as well. However, life cycles become important and interesting when we consider the time for the innovator to respond to the imitator's differentiation attempts. Whereas in the pharmaceutical industry, the innovator's ability to respond quickly is limited by regulatory constraints, it is not likely to be the case in products without such regulatory scrutiny. If we allow for the innovator to respond rapidly to the differentiation attempts of the imitator, we expect a weakening of our results especially if the capabilities of the innovator and the imitator are largely similar.

A second boundary condition that is important is the cost of imitation. In the branded drug industry, a significant proportion of development cost is already sunk by potential imitators, that is, all the big pharmaceutical firms own large banks of molecules that they can draw upon to quickly commence rival drug programs. For instance, this explains why two imitator drugs—Cialis and Levitra—were able to enter the market so quickly after the launch of Viagra. When a significant cost of imitation is already sunk, the incentives for imitation are significantly higher, especially if the incremental cost of imitation is a small proportion of the total cost (Mansfield, Schwartz, and Wagner, 1981). Similarly, to the extent that marketing costs are already sunk in the form of large, underutilized sales forces, the incentive to imitate is amplified. Thus, we believe that to the extent imitation costs are not sunk, our results are likely to be weakened since firms may wait considerably longer for the uncertainty to resolve before they commence imitation. This would extend the early entry advantage of the innovator.

Finally, our study, like all other empirical studies, is not without its limitations. We think three limitations are significant. First, the idea of vertical differentiation being possible after the leakage of information about phase II clinical trials of the innovator could not be directly tested. The results fit well with the industry average of seven years reported in prior research (DiMasi, 2001a). Ideally we would have liked to collect the actual phase II clinical trial completion dates for the innovator. Unfortunately, we were unable to find reliable information on this, especially for the older drugs. Second, we measured horizontal and vertical differentiation as a count of the indications and side effects. It is possible that there is a vertical dimension within each indication. For instance, one drug might act more quickly than another. Such differences are not captured in our measures and thus classified as horizontal differentiation. This means that our measure of vertical differentiation is conservative, which perhaps strengthens the confidence in our results. Finally, our regressions estimate the likelihood of beating the innovator in sales. This parallels the problem of demand estimation, which requires simultaneous consideration of prices and quantities. For a variety of reasons outlined earlier, we could not control for the effect of prices. In addition, it is possible that prices (of the imitator drug and the response of the innovator) may be systematically related to horizontal or vertical differentiation of the imitator drug. While this is a clear limitation, it is also common to most other studies of entry timing effects. We hope that future research can address these three limitations.

In conclusion, notwithstanding the peculiarities of the branded drug industry, we believe that our theory of imitative entry is fairly general. If the boundary conditions outlined above are met, we expect the conclusions to carry over into other industries and products as well. Nevertheless, a systematic empirical test of our theory in other products and industries will be a useful extension.

ACKNOWLEDGEMENTS

Duke University, the University of North Carolina-Chapel Hill, and Harvard Business School for helpful comments and suggestions. Deborah Gray provided excellent editorial assistance. Errors and omissions remain our own.

REFERENCES

Abernathy WJ, Utterback J. 1978. Patterns of industrial innovation. *Technology Review* **80**(7): 40–47.

Banerjee PK, Myers SD, Baker A, Andersen M, Bichet S. 2001. *High performance drug discovery*. Accenture Executive Briefing: Philadelphia, PA.

Barney JB. 1986. Strategic factor markets: expectations, luck, and business strategy. *Management Science* **32**(10): 1231–1241.

Baron RM, Kenny DA. 1986. The moderator mediator variable distinction in social psychological-research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* **51**(6): 1173–1182.

Berndt ER, Bui L, Reiley DR, Urban GL. 1995. Information, marketing, and pricing in the U.S. antiulcer drug market. *American Economic Review* **85**(2): 100–105.

Bohlman JD, Golder PN, Mitra D. 2002. Deconstructing the pioneer's advantage: examining vintage effects and consumer valuations of quality. *Management Science* **48**(9): 1175–1195.

CDER. 2006. *From test tube to patient: Protecting America's health through human drugs*. Center for Drug Evaluation and Research, Food and Drug Administration: Washington, D.C.

Chatterjee S, Hadi AS, Price B. 2000. *Regression Analysis by Example* (3rd edn). John Wiley & Sons: New York.

Cohen FJ. 2006. Entry order as a consideration for innovation strategies. *Nature Reviews Drug Discovery* **5**: 285–294.

Cohen WM, Levinthal DA. 1990. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* **35**(1): 128–152.

Cohen WM, Nelson RR, Walsh JP. 2000. Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not). Working Paper, National Bureau of Economic Research. NBER: Cambridge, MA.

Cyert RM, March JG. 1963. *A Behavioral Theory of the Firm*. Prentice-Hall: Englewood Cliffs, NJ.

Dasgupta P. 1988. Patents, priority and imitation or, the economics of races and waiting games. *Economic Journal* **98**(389): 66–80.

Dierickx I, Cool K. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science* **35**(12): 1504–1512.

DiMasi JA. 1995. Success rates for new drugs entering clinical testing in the United States. *Clinical Pharmacology and Therapeutics* **58**: 1–14.

DiMasi JA. 2001a. New drug development in the United States from 1963 to 1999. *Clinical Pharmacology and Therapeutics* **69**: 286–296.

DiMasi JA. 2001b. Risks in new drug development: approval success rates for investigational drugs. *Clinical Pharmacology and Therapeutics* **69**: 297–307.

Eaton BC, Lipsey RG. 1998. Product differentiation. In *Handbook of Industrial Organization* Vol. 1 (5th edn), Schmalensee R, Willig R (eds). Elsevier Science: North Holland, The Netherlands; 723–768.

Ethiraj SK, Kale P, Krishnan MS, Singh JV. 2005. Where do capabilities come from and how do they matter? a study in the software services industry. *Strategic Management Journal* **26**(1): 25–45.

Ethiraj SK, Levinthal DA, Roy RR. 2008. The dual role of modularity: innovation and imitation. *Management Science* forthcoming.

Golder PN, Tellis GJ. 1993. Pioneer advantage: marketing logic or marketing legend. *Journal of Marketing Research* **30**(2): 158–170.

Greve HR. 2000. Market niche entry decisions: competition, learning, and strategy in tokyo banking, 1894–1936. *Academy of Management Journal* **43**(5): 816–836.

Huff LC, Robinson WT. 1994. The impact of leadtime and years of competitive rivalry on pioneer market share advantages. *Management Science* **40**(10): 1370–1177.

Klepper S. 1996. Entry, exit, growth, and innovation over the product life cycle. *American Economic Review* **86**(3): 562–583.

Klepper S. 2002. Firm survival and the evolution of oligopoly. *RAND Journal of Economics* **33**(1): 37–61.

Lee H, Smith KG, Grimm CM, Schomburg A. 2000. Timing, order and durability of new product advantages with imitation. *Strategic Management Journal* **21**(1): 23–30.

Lieberman MB, Asaba S. 2006. Why do firms imitate each other? *Academy of Management Review* **31**(2): 366–385.

Lieberman MB, Montgomery DB. 1988. First-mover advantages. *Strategic Management Journal,* Summer Special Issue **9**: 41–58.

Lieberman MB, Montgomery DB. 1998. First-mover (dis)advantages: retrospective and link with the resource-based view. *Strategic Management Journal* **19**(12): 1111–1125.

Lu JZ, Comanor WS. 1998. Strategic pricing of new pharmaceuticals. *Review of Economics & Statistics* **80**(1): 108–118.

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. 2002. A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods* **7**(1): 83–104.

Mansfield E, Schwartz M, Wagner S. 1981. Imitation costs and patents: an empirical study. *Economic Journal* **91**(364): 907–918.

Mathieu MP. 2001. *Parexel's Pharmaceutical R&D Statistical Sourcebook*. Parexel International Corporation: Waltham, MA.

Nelson RR. 1991. Why do firms differ, and how does it matter? *Strategic Management Journal,* Winter Special Issue **12**: 61–74.

Pepall L. 1997. Imitative competition and product innovation in a duopoly model. *Economica* **64**(254): 265–279.

Polidoro F. 2006. Institutions and innovations: a longitudinal study of institutional endorsement and firm innovation in the pharmaceutical industry, 1980–2004. PhD diss. Strategy Dept., Ross School of Business. University of Michigan: Ann Arbor.

Reed R, DeFillippi RJ. 1990. Causal ambiguity, barriers to imitation, and sustainable competitive advantage. *Academy of Management Review* **15**(1): 88–102.

Rivkin J. 2000. Imitation of complex strategies. *Management Science* **46**(6): 824–844.

Robinson WT, Fornell C. 1985. Sources of market pioneer advantages in consumer goods industries. *Journal of Marketing Research* **22**(3): 305–317.

Rumelt RP. 1984. Toward a strategic theory of the firm. In *Competitive Strategic Management*, Lamb RB (Ed). Prentice-Hall: Englewood Cliffs: NJ; 556–570.

Rumelt RP, Schendel DE, Teece DJ. 1994. Fundamental issues in strategy. In *Fundamental Issues in Strategy: A Research Agenda*, Rumelt RP, Schendel DE, Teece DJ (eds). Harvard Business School Press: Boston, MA; 9–47.

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

Teece DJ. 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy* **15**(6): 285–305.

Theil H. 1967. *Economics and Information Theory*. Rand-McNally: Chicago.

Tushman ML, Anderson P. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* **31**(3): 439–465.

Urban GL, Carter T, Gaskin S, Mucha Z. 1986. Market share rewards to pioneering brands: an empirical analysis and strategic implications. *Management Science* **32**(6): 645–659.

VanderWerf PA, Mahon JF. 1997. Meta-analysis of the impact of research methods on findings of first-mover advantage. *Management Science* **43**(11): 1510–1519.

Yeoh PL, Roth K. 1999. An empirical analysis of sustained advantage in the US pharmaceutical industry: impact of firm resources and capabilities. *Strategic Management Journal* **20**(7): 637–653.

APPENDIX: DRUG DEVELOPMENT IN THE BRANDED DRUG INDUSTRY[10]

The process of *drug discovery* involves the identification and validation of candidates or targets for disease, and the design, synthesis, screening, and validation of molecules or drug leads that will act on the targets. A drug target is a protein, cell, or human organ that is affected by or is a cause of disease. A single disease may be caused by multiple targets, implying that there are often multiple approaches to disease treatment.[11] The discovery of a new target (termed new mechanism of action) and a corresponding drug that acts on the target signals the emergence of a new product category for a particular disease treatment. The successful identification of a drug target and the design of a molecule that acts on the target concludes the drug discovery process. The drug discovery process is estimated to span an average of about four years (Banerjee *et al.*, 2001) and consumes about 27 percent of total R&D budget in the branded drug industry (Mathieu, 2001) or about $10 billion in 2001. The drug discovery process culminates with the synthesis of a drug molecule and the filing of a patent with the United States Patent and Trademark Office.

The process of *drug development* begins with the testing of the new molecule or drug on animals. Once the safety and efficacy of the drug is established in animal trials, the firm applies to the FDA (called an investigational new drug application or INDA) for permission to start human trials. Human clinical trials evolve through three sequential phases. Phase I trials involve demonstrating safety in human use, yielding preliminary data on drug side effects. The success rate of phase I trials is about 22 percent (DiMasi, 2001b). Phase II trials test for drug efficacy in disease treatment in addition to assessing side effects. In phase II trials,

firms define explicit markers for disease treatment called *indications* and look for improvement in them. For instance, Lipitor, a widely prescribed drug for cholesterol treatment, lists three indications: reduction in total cholesterol, LDL cholesterol, and triglyceride levels respectively. These indications have to be measurable. The demonstration of drug efficacy involves showing improvement over a placebo. The conclusion of phase II clinical trials yields fairly reliable information on both efficacy (i.e., indications that the drug treats) and safety (i.e., side effects associated with the drug). About 32 percent of all drugs entering phase II trials eventually move on to phase III (DiMasi, 2001b). Finally, phase III clinical trials involve demonstration of safety and efficacy in long-term use. The success rate for phase III trials is about 78 percent (DiMasi, 2001b). This suggests that a significant proportion of the uncertainty surrounding new drug development is resolved with the successful conclusion of phase II clinical trials. The human trials process usually spans about 7–12 years and consumes about 38 percent of total R&D budgets.

Upon successful completion of the human trials, firms submit the clinical trial data to the FDA and seek approval for marketing the drug. After FDA review and approval, which typically takes about two years, the drug is launched in the market. Between the filing of the patent after synthesis of the molecule and final FDA approval (typically about 8–13 years), there is gradual leakage of drug efficacy and safety information for the innovator's drug. The leakage of information occurs via three different sources: (1) the filing of patents reveals the composition of the drug; (2) the recruitment of patients for each stage of clinical trials signals success in the prior stage; and (3) the publication of clinical trial research results in scientific journals, which is considered essential to building up the evidence for FDA approval especially in the case of new drug targets. This gradual leakage of efficacy and safety information on the innovator's drug over a decade sets the stage for the dynamics of innovator-imitator competition in the branded drug industry. Thus, imitators can choose to initiate product development efforts at any point after the filing of the innovator's patent.

---

[10] This section draws extensively from an FDA publication (CDER, 2006).

[11] For instance, there are four different targets for AIDS treatment. Entry inhibitors prevent the Human Immunodeficiency Virus (HIV) from entering cells. Protease inhibitors inhibit protease activity since protease is involved in the replication of HIV within T-cells. Both Nucleoside reverse transcriptase inhibitors (NRTIs) and Non-nucleoside reverse transcriptase inhibitors (NNRTIs) prevent healthy T-cells in the human body from becoming infected. Whereas NRTIs work by delivering faulty nucleotide building blocks that inhibit the HIV genetic material from replicating, NNRTIs work by attaching themselves to the reverse transcriptase protein and inhibiting HIV replication.