

CENTRO DE INVESTIGACIÓN ECONÓMICA

Discussion Paper Series

**Developer's Expertise and the Dynamics of
Financial Innovation: Theory and Evidence**

Helios Herrera
Instituto Tecnológico Autónomo de México

and

Enrique Schroth
HEC Lausanne

August 2005
Discussion Paper 05-04

Developer's Expertise and the Dynamics of Financial Innovation: Theory and Evidence

Helios Herrera (ITAM) and Enrique Schroth (HEC Lausanne)

June 2005

Abstract

We study product innovation and imitation in the market of corporate underwriting with a dynamic model where client switching costs and the bankers' expertise in deal structuring characterize the life cycle of a security. While the clientele loyalty allows positive rent extraction, the superior expertise can account for the documented market leadership of the innovator. As expertise on product structuring is acquired by imitators, the innovator's market share advantage decreases. Also, the speed of entry by imitators increases for later generation products. Our predictions are consistent with well documented evidence on the market share leadership of innovators. We also present new evidence from equity-linked and derivative corporate products that supports the dynamic predictions of our learning model.

JEL Classification: G24, L12, L89.

Keywords: Innovation and imitation, first-mover advantages, product differentiation, learning.

1 Introduction

Investment banks have been at the forefront of financial innovation in the last two decades, increasing the number of security designs that issuing firms can use to raise new money. The volume of cash that banks underwrite using these products has also taken an increasing proportion of the overall underwriting market. But innovation is often followed by imitation and even large banks of big reputations avoid expenditures in research and development and compete vis à vis the innovator with an imitation of the original product. Yet, the empirical evidence suggests quite strongly that the developers of new securities are able to preserve a competitive advantage over their imitators. Why this is the case is still an open question in the finance field. It is important and timely to study the source and the evolution of the developer's advantage if we want to understand the incentives of banks to innovate, how these incentives affect the speed of innovation and, in turn, how the protection of innovation through patent laws may affect these incentives.¹

Peter Tufano's seminal empirical study of innovation in corporate products showed that investment banks that develop new corporate products enjoy a market share leadership over equally reputed rivals (Tufano, 1989).² Thus, despite the fact that imitators are offering similar products, a significant share of the issuing firms are more likely to choose the innovator as their underwriter. One reason they may have such a preference is that the imitator can underwrite deals with only an imperfect version of the product. In fact, for many product innovations, the reverse engineering often does not result in a perfect substitute (Toy, 2001).³ Similarly, in their study of the investment

¹It is widely recognized that patents have been ineffective ways of protection from competition by imitators in finance. While it was always possible to obtain a patent on an innovative corporate product, it was virtually impossible to enforce the patent before 1999, as most financial innovations are considered "business methods or formulas". Under patent laws, business methods were unpatentable until the US Supreme Court upheld a patent on a "business method" in 1999. It is believed that the *State Street Case* has set the precedent required to make patents effective to protect R&D in financial products development.

²Thorough surveys of innovations in corporate finance instruments are provided by Peter Tufano (Tufano, 1995) and John Finnerty (Finnerty, 1992). A more comprehensive survey of financial innovation in general is provided by Franklin Allen and Douglas Gale (Allen and Gale, 1994).

³This view that imitations are imperfect substitutes is summarized in the testimony by William Toy, a Managing

banking industry, Dwight Eccles and Robert Crane argue that the skill to structure the issue of a corporate security takes time to acquire (Eccles and Crane, 1988). These views suggest that the innovator is effectively the expert underwriter, and the competitors are imperfect imitators.

This paper argues that the innovator has an advantage over the imitator that consists of a superior expertise in structuring a complex security for any given client. The innovator acquires the expertise about the new security during its the development stage, whereas an imitator cannot reverse-engineer it perfectly or acquire the innovator's expertise instantly after the new security is marketed. To back our claim we build a model of the underwriting market around this main feature and then test the validity of its empirical predictions with the existing empirical evidence of financial innovation, and with new evidence that we provide here.

To our knowledge very few researchers try to answer why firms may have a preference for the innovator rather than the imitator.⁴Our model considers the fact that the imitators do not have the same expertise as the innovator initially but may acquire it as more deals are being completed. Thus, the model characterizes the timing of the entry of imitators and the pricing behavior before and after imitation. It has distinctive dynamic implications that are comparable with existing and new evidence on financial innovation. as more issues of the innovative security are underwritten and imitators perfect their own expertise, the innovator's advantage decreases and eventually disappears as the innovation approaches a commodity product status. This dynamic pattern suggests that the innovator's profits are eroded gradually along the life cycle of new financial products (see Van Horne, 1985).

The theoretical part of the paper is a duopoly model, i.e., an innovator vs. an imitator, where the comparative statics show that the smaller the innovator's quality advantage the faster the imitators

Director at CDC Capital:

"There is at least a perception that the first mover is more familiar with the product he issues than the imitator,"
(William Toy, personal interview, New York City, February 2001).

⁴In a recent survey, Peter Tufano argues that many mechanisms that reward innovation still remain to be studied (Tufano, 2003).

enter the market and the faster the market shares of innovators and imitators converge. Intuitively, the more highly structured is the product the harder it is to reverse-engineer and the longer the innovator can maintain its market share advantage. To verify empirically these predictions we look into innovations that can be classified into product groups and generations of products within a group according to the relationship of a product to its predecessors (i.e., the prior art). We find such type of innovations in the equity-linked and derivatives class of the Securities Data Company New Issues Database. This class of products has become increasingly important, not only as a fertile ground for innovation in corporate products, but also as a large source of funds: between 1985 and 2002, firms have raised over US\$ 200 billion, which represents almost 16% of the cash that was raised using common stock. For the case of equity-linked securities, some products are radically innovative while others are only enhancements of previous products. Since a later generation product builds on the prior security designs, i.e., is less innovative than a first generation product, it could be reverse-engineered more effectively. Thus, the initial expertise advantage of the first-mover is expected to be stronger in first generation products than on later generations. Indeed, for the later generation products our model predicts faster imitation and faster market share convergence than for earlier generation products. We find that the theoretical predictions on the speed of imitation match the empirical evidence on equity linked securities that we present.

This paper relates to the work of Sugato Bhattacharyya and Vikram Nanda on the role of switching costs in financial innovation (Bhattacharyya and Nanda, 2000). They show that a potential innovator can charge underwriting fees over marginal costs to loyal clients who face costs to switching to other banks. Thus, banks endowed with a broader or more loyal client base can profit more in new markets. While client loyalty certainly affects the innovator's profits they may not eliminate the free-rider problem. If client loyalty were the only source of rent extraction then any potential innovator would rather be an imitator than develop, save the development costs and still underwrite the security for its loyal clientele at a premium. In other words, the advantage belongs to the second mover rather than to the first. Our model exhibits the developer's expertise advantage feature together with the switching costs feature. Innovators and imitators compete with

products that are differentiated horizontally and vertically. The horizontal dimension represents the switching costs and the vertical dimension represents the innovator’s expertise in structuring deals. The vertical dimension is crucial to account for the stylized facts in the literature and the new evidence provided here. The different expertise advantages across subsequent generations of a family of innovations accounts for the faster expected timing of entry of imitation and the faster speed of convergence of market shares for later generations that our data show. It is important to stress that other explanations that do not rely on the developer’s expertise can hardly explain the dynamic pattern of the decreasing market share advantage of the innovator and the faster speed of entry in equity-linked securities.

Our model can also address the interactions between the size of switching costs or the size of initial clienteles and the incentives to innovate. As in the model by Bhattacharyya and Nanda, our model predicts that banks with smaller initial clienteles may never innovate. In fact, we do see in the data that competition in derivative corporate products involves mostly the “bulge bracket” Wall Street banks. However, innovators do not appear to have large market shares only *because* of their large initial clienteles: we also observe that within the group of large banks, they are sometimes leaders and other times followers. Enrique Schroth estimates the demand function for a given underwriter and finds that the leadership is systematic to the innovator, even after controlling for the size of the clientele of the bank (Schroth, 2003). Thus, even large banks also have small market shares, and this is generally when they are imitators.

Our analysis has implications too about the speed at which innovations are introduced. Innovator’s may have an idea for a marketable security, but may not offer it as soon as they have it. In our model, this happens because the arrival of issuing firms is random, so even if the bank has completed the design privately, there may not be any close clients looking for external finance. Underwriting the security with a distant client is not very profitable because the client is switching away from his bank. Moreover, it triggers the competitor’s learning process by the imitators too soon. Thus, our model predicts that innovators will wait for good clients to come to market, or, market their innovations aggressively to their clients, or, alternatively, innovate based on their

clients' capital structure targets.

The profits of innovation in our model increase when imitators cannot learn too much information about optimal product engineering from each deal. Banks innovate more often in markets where inference about the optimal engineering by the imitator is clouded by a changing economic environment, i.e., a higher volatility. Innovation is more frequent in volatile markets not because firms demand new products to hedge risks but because innovating banks will expect larger first-mover advantages there.

John Persons and Vincent Warther propose a theory of the adoption of financial innovations that explains some documented cases of boom and bust cycles (Persons and Warther, 1997). In their model, issuing firms are differentiated in terms of their cost of adopting an innovation of uncertain value. At every period, all firms that have not yet been financed will choose to use the innovation as a financing instrument if its expected value exceeds its cost. The expected value to the non-adopters is updated after a noisy signal of the true value is revealed from those that had adopted it. Since the precision of this signal would depend on the total number of adopters, they can generate waves of adoption in equilibrium. Our model is different in several respects. In first place, our model differs in the timing: here one client is drawn independently at each time, while in their model, all firms are potential issuers every period. We chose to model the market this way in order to exploit fully the available data on new issues of corporate derivatives: the data is recorded deal by deal, and banks bid for each underwriting deal at a time. Therefore, the model can make predictions of the future expected market shares after each deal is completed. But most crucially, their model abstracts from competition between rival banks while ours focuses precisely in the demand for the underwriting services of different competing banks and the relative advantage of the innovator over its imitator across time. Finally, in our model the size of the aggregate demand for the new product is fixed every period (i.e., one client at a time). Allowing for changes in the demand every period may not affect significantly our predictions in terms of the *relative* advantage of one bank over the other, or the speed of entry by imitators.

We proceed with Section 2, where we describe the elements of the model, and explain how

imitation may be imperfect. Section 3 characterizes the equilibrium when the innovator and the imitator have identical underwriting market shares before the innovation is introduced, and Section 4 formalizes the acquisition of underwriting expertise by the imitator. Section 5 pins down the equilibrium profits of innovation and discusses the incentives to innovate. Section 6 generalizes the results to the case where the competitors have initially asymmetric client bases and reputation is accumulated throughout the product's life. Section 7 tests the predictions with existing and new evidence found in the underwriting market of equity-linked and corporate derivative products. Section 8 summarizes briefly.

2 The Setup of the Model

2.1 The Underwriting Market

In this section we model the market of corporate underwriting. There are two types of underwriters: the innovator and the imitator. Each type offers its own variety of a corporate finance product, i.e., a structured security, that firms can issue to raise funds. The innovator is the bank that first developed a new security design and competes with the imitator to underwrite every issue of the innovative security by a given firm. Let the underwriters be indexed by i , the innovator is $i = 0$ and the imitator is $i = 1$. The case with more imitators is developed in the appendix and produces the same qualitative results. An underwriter is hired by an issuer to structure the deal and sell the securities to investors. The underwriter charges its client a fee, i.e., the underwriting spread, for such a service.

The game starts at period 0 when the innovator ($i = 0$) gets an idea about a new corporate security design. The potential innovator can choose to develop and market the new security by paying a fixed R&D cost, F_0 . As soon as it chooses to innovate, it starts underwriting issues of the new security. After the innovator completes its first underwriting deal, information about the security design is revealed. With this information, the other bank can develop a similar product and become an imitator. We assume that the imitator can free-ride completely the R&D, that

is, $F_1 = 0$.⁵The innovator is a monopolist only for the first deal. After that deal, the presence of imitation limits his market power as both banks compete in underwriting spreads.

The underwriting service provided by banks is differentiated, both vertically and horizontally. The vertical dimension measures the quality of the product: all other things constant, any issuer derives a higher value if she hires an investment bank that provides a higher quality underwriting.

The horizontal dimension describes the preferences of issuers for a particular bank. Issuers are “located” on a unit interval, and their mass is distributed over it following a given distribution. For now we assume that the distribution of issuers is uniform over $[0, 1]$ and relax this assumption in Section 6. At every period nature draws the next firm who will seek for an underwriting deal from the said distribution. We assume that the two competing investment banks in this economy offer differentiated varieties and are located at the two extremes of the interval (without loss of generality, the innovator is located at 0; see Figure 1 for an illustration).

Let the quality of underwriter i 's product be q_i and assume that the preference for quality, location and the price paid, p_i , enter linearly into the firm's valuation of an underwriting deal. Then, the values of a client located at x of hiring either type of banker as its underwriter are given by:

$$\begin{aligned} u_0(x) &= q_0 - p_0 - sx, \\ u_1(x) &= q_1 - p_1 - s(1 - x), \end{aligned}$$

where s is the cost per unit of distance of choosing a variety located away from the preferred one.⁶

Note that with this setup each bank will have its own clientele of financing firms. The value to a firm of adopting the product of bank 0 or 1 depends on relative prices, on the quality of the product, but also on the proximity of the bank's variety to its preferred one. The horizontal

⁵We could relax this extreme assumption to $F_0 > F_1 > 0$. This would only strengthen the innovator's advantage and not change at all the comparative statics.

⁶Note that the price, p , is not the price at which the issue of the new security is sold by the client to investors, but the price that the client pays to the bank (innovator or imitator) for the engineering of the new security. It is thus the underwriting fee.

dimension represents then the degree of loyalty that issuing firms have to the available underwriters since a firm always belongs to a given bank's clientele and it faces a cost of switching bankers. Thus, hereafter we refer to s as the size of the switching cost or the loyalty of the client interchangeably.⁷

Every time a firm is drawn, she chooses its underwriter, i , to maximize the value of its contract, u_i . We assume that the issuing firm has a reservation value normalized to zero and cannot delay the financing decision. After a firm is drawn, both banks compete in prices to sign an underwriting deal with her. Given the firm type, qualities, and switching costs, each banker's per-deal profits are:

$$\pi_i = (p_i - c) D_i(x, p_0, p_1, x, q_0, q_1, s, t) \quad (1)$$

for $i = 0, 1$. The term c represents the marginal cost of underwriting (e.g., SEC filing, advertising, legal fees) and t the order of the draw, i.e., the security has a history of $t - 1$ deals. Only one bank gets the current deal so the demand functions are given by:

$$\begin{aligned} D_0 &= \begin{cases} 1 & \text{if } u_0(x) > u_1(x), \\ 0 & \text{otherwise,} \end{cases} \\ D_1 &= 1 - D_0. \end{aligned}$$

At period zero, the expected profits to the innovator are:

$$\Pi_0^e = -F_0 + \pi_0^e(0) + E \sum_{t=1}^{\infty} (1 - \delta)^t \pi_0(t)$$

where $\pi_0^e(0)$ denotes the innovator's expected profits in the first deal, which he gets for sure being still the only issuer. Note that most innovations in corporate security designs are finite-lived. The infinite-horizon assumption is a natural way to model the problem if we introduce the probability that the game continues for one more period as a discount factor, which we call $(1 - \delta)$. To save in

⁷We consider only the type of switching costs where firms are more likely to choose a bank with which they have had relationships in the past in other lines of business, i.e., where banks have a loyal clientele from the outset. Client loyalty might also be developed during the life cycle of an innovation. However, we do not model this other source of client stickyness because it does not allow a better interpretation of the data: no firm has issued the same security more than once.

notation we have excluded the “pure time” discount factor. However, this can be easily incorporated to the model if we interpret $(1 - \delta)$ as the product of the probability of continuation and the pure time discount.

2.2 Financial Innovation

An innovation is a new corporate security that a firm can issue to raise money. Due to disclosure regulations, the design of the new security is revealed to imitators. However, this design typically has several parameters that have to be set for each deal. For example, among other things, a PERCs (Preferred Equity Redemption Cumulative Stock) issue has to specify the conversion rate of preferred to common shares as a function of the returns of common stock, by choosing a cap to the appreciation, \bar{r} . The contract also must specify the dividends paid and the sale price (see Figure 2).⁸

A bank that wants to imitate PERCs can see what is the general structure of the product but still does not know how to set optimally for his client the specific parameters of the security such as caps, conversion rates and price. For this reason a client who decides to issue PERCs would expect a higher quality of underwriting from the original developer of the security, all other things being equal.

Similarly, generic equity-linked debt products must specify the stock or stock index whose price is tied to the adjustable face value. Thus, to underwrite an issue of any given security, the underwriter has to structure each deal by customizing the parameters specified by the design.⁹ Deal customization has been well documented. It is depicted in testimonies by bankers collected by Eccles and Crane (Eccles and Crane, 1988). Recently, Schroth analyzes the structuring of equity-linked deals and finds a significant variation across the parameters within same designs (Schroth, 2003).

We assume that the skill needed to customize deals is acquired with expertise. If the innovator

⁸PERCs are shares of preferred stock that are mandatorily convertible to common stock after 3 years.

⁹A notable example of an equity-linked bond is Salomon Brother’s invention, the ELK, a bond whose face value is pegged to the appreciation of a chosen traded stock.

has superior expertise than the imitator he structures the deals better and, *ceteris paribus*, he provides a higher quality underwriting. We let the investment banks's expertise be q , the quality parameter of the product. While the imitator can learn the design structure immediately and for free, he may only be able to imitate the innovator's new product imperfectly or with an inferior customizing skill than the innovator. In such case, $q_0 > q_1$, and let $\Delta q \equiv q_0 - q_1 \geq 0$ be the quality differential.

The characteristic of corporate derivatives that we highlight in this paper is that the product design does not disclose immediately all the private information of the innovator. We believe that other prominent financial vehicles share these features. For example, Goldman pioneered and remained the lead underwriter of puttable securities indexed to the Nikkei Index. The idea of issuing Nikkei Put warrants was disclosed rapidly to competitors but Goldman also hedged the issuer's exposure to the Nikkei discretionally, profiting from private knowledge acquired during the development of the hybrid security.

Some anecdotal evidence also exhibits similar features to the corporate products in our study. Thackray (1985), for example, documents how Drexel, Burnham, Lambert did not disclose its "junk-bond" prospectuses to Wall Street insiders because of fears that competitor's imitations may challenge their lead in the market for underwriting high-yield debt. J.P. Morgan's lead in underwriting asset-backed securities using its so called BISTRO variety of a collateralized loan obligation arguably hinges on the discretion with which it manages the pool of assets used as collateral (Roper, 1999). Thus, while data is not available we would expect to observe similar patterns of innovation and imitation in highly structured debt products, such as credit derivatives, collateralized loans and asset-backed obligations.

3 Equilibrium

We assume that q_1 , which is smaller than q_0 , is large enough to guarantee that the ex-post profits to the innovator or the imitator from a deal with any potential client are positive if either were a monopolist. Thus, we assume that

$$q_1 \geq c + s. \quad (2)$$

In the first deal, the innovator is a monopolist and makes a certain profit π_M . After the first deal the innovator loses part of its market power.

After the first deal, underwriters compete for the following client that wants to issue the new security. Banks compete by undercutting prices until one of them reaches its marginal cost. Define \hat{x} as the client that, when offered a deal priced at marginal cost by *both* banks, is indifferent between either. That is, \hat{x} equates $u_0(\hat{x})$ and $u_1(\hat{x})$ for $p_0 = p_1 = c$. Solving, we obtain

$$\begin{aligned} q_0 - c - s\hat{x} &= q_1 - c - s(1 - \hat{x}) \\ \hat{x} &= \frac{1}{2} + \frac{\Delta q}{2s}. \end{aligned}$$

The value of \hat{x} denotes the location (loyalty) of the indifferent client. Whenever the innovator's quality advantage is high relative to the clientele effect, i.e., $\Delta q > s$, then the "indifferent" client lies outside the unit interval and the innovator gets the next deal for sure. Still, the presence of the imitative competitor puts a bound on the markup that the innovator can obtain.

Let $\bar{x} = \min(1, \hat{x})$. Figure 3 illustrates the probability that either competitor obtains the next deal as a function of the quality advantage, Δq . For any client $x \in [0, \bar{x})$ the innovator can undercut the imitator's price below its break-even level, making the the client indifferent and thus getting the deal while making a profit). The innovator's price is solves:

$$\begin{aligned} q_0 - p_0 - sx &= q_1 - c - s(1 - x), \\ \Rightarrow p_0 &= c + (1 - 2x)s + \Delta q. \end{aligned}$$

Similarly, an imitator gets any client $x \in [\bar{x}, 1]$ and prices the deal at

$$p_1 = c + (2x - 1)s - \Delta q.$$

Given that in every period one firm is drawn uniformly, the expected one period profits for both

banks are:

$$\pi_0^e = \int_0^{\bar{x}} (p_0 - c)dx = \bar{x}((1 - \bar{x})s + \Delta q) = \begin{cases} \Delta q & \text{for } \Delta q > s, \\ s\left(\frac{1}{2} + \frac{\Delta q}{2s}\right)^2 & \text{for } \Delta q < s, \end{cases}$$

$$\pi_1^e = \int_{\bar{x}}^1 (p_1 - c)dx = (1 - \bar{x})(\bar{x}s + \Delta q) = \begin{cases} 0 & \text{for } \Delta q > s, \\ \left(\frac{1}{2} - \frac{\Delta q}{2s}\right)^2 & \text{for } \Delta q < s. \end{cases}$$

The innovator's expected profits are higher, due to the higher quality. In fact, $\pi_0^e - \pi_1^e = \Delta q$ and $\Delta q > 0$ as long as the imitator cannot reverse engineer the innovation perfectly. As the quality advantage vanishes for uniformly distributed clients the probabilities of obtaining the deal converge to one half for both competitors and the expected profits are equal.

The quality differential is a crucial element of innovation in this model. The model exhibits the typical free-rider problem in product innovation because the security design is disclosed publicly and $F_1 < F_0$. However, deals have to be customized within the design of the product, and this leaves room for quality differences. In the next section we formalize how expertise is acquired as deals are completed and the innovation develops into a commodity.

4 The Acquisition of Product Expertise

We now focus on the learning process that describes the dynamics of Δq . We use the dynamics of the expertise acquired by both competing banks to analyze the underwriting game equilibrium and make comparative statics predictions.

An imitator can improve his deal structuring from the moment he observes the new security. He acquires product expertise as more deals on that security are completed in the market. Let the expertise specific to a given security be summarized by the knowledge of a variable, a . To understand better the meaning of a , consider the following factors that affect the quality of the underwriting service: (i) the underwriter must learn how to choose the right parameters that are best for different issuers; (ii) investment banks need to identify changes in the tastes of investors or changes in market conditions and structure each issue accordingly to maximize the proceeds;

and (iii) underwriters also provide advice to issuers on how to hedge the liabilities or to invest the proceeds associated to the issue of the securities they engineer. In fact, in some cases the underwriters may buy some of the issued shares, in which case they need to understand the product's effect on the risk and returns of a portfolio.¹⁰ The innovator may provide a superior quality underwriting by improving his knowledge on all these dimensions during the development process. Additionally, his superior expertise may come from: (iv) the innovator contacting potential issuers before the first deal to have a leg-up on the first few deals, (v) the innovator having its lawyers sign-off on the documentation while an imitator would have to get used to the legal or tax risks before the deal is presented to clients, or (vi) the innovator having collected prior information in potential investors' demand in the secondary markets.¹¹

Thus, we can think of a as a mapping parameter of these changing conditions (clients, markets, investors, own investments) to the optimal deal structure. A higher quality is tied to a superior product or market expertise, which is itself a better knowledge of a . Formally, product expertise is the precision of the information that the underwriter has about the unknown value of a . The prior density of a is normally distributed with variance A^{-1} . In the case of the innovator, R&D provides him information about a through a signal:

$$z_0 = a + \varepsilon_0, \tag{3}$$

where the noise component ε_0 is a normally distributed variable with $E(\varepsilon_0) = 0$ and $Var(\varepsilon_0) = \tau^{-1}$.

Bayesian updating gives the posterior precision or the knowledge of the innovator about the engineering choice a :

$$q_0 = A + \tau$$

we identify this precision with the quality q_0 of the product engineered by the the innovator.

¹⁰The case of the Nikkei Put Warrants introduced by Goldman, Sachs & Co. in 1990 illustrates these factors very well.

¹¹Robert Hauswald and Robert Marquez propose a model where banks improve the technology for screening loan applicants to give them a first mover in the loan market (Hauswald and Marquez, 2003).

Even though we have illustrated product expertise as multidimensional, we prefer to treat a as a scalar. We believe that making a a vector does not add any important insight, while treating it as a scalar keeps our exposition parsimonious.

4.1 Learning by the Imitator

Before the new security is issued the first time, the precision of the imitator's information about the security engineering parameters a is also A^{-1} . After observing the first deal completed by the innovator or any later deal underwritten by himself or by the innovator, the imitator is able to update his information about a . In other words, the imitator observes a noisy signal z_1 , which reveals information about a . Even if the design of the innovative security is disclosed publicly after the first deal, the leakage of information about a is only partial and the imitator's signal has an additional normally distributed noise η_1 relative to the innovator's information:

$$z_1 = (a + \varepsilon_0) + \eta_1,$$

where $E(\eta_1) = 0$, and $Var(\eta_1) = \Sigma^{-1}$. Note that Σ is the precision of the imitator's signal beyond the incompressible component $(a + \varepsilon_0)$, i.e., how much is revealed after each deal is completed and a signal is extracted. Let $\tau_1(t)$ be the imitator's precision or his product quality after observing t deals. The quality differential between the products is equal to the difference in precision:

$$\Delta q(t) = \tau_0 - \tau_1(t)$$

Lemma 1 *The difference in quality after t deals is:*

$$\Delta q(t) = \tau \frac{1}{1 + \frac{\Sigma}{\tau} t}. \quad (4)$$

We prove this Lemma in the appendix. The quality difference decreases and converges to zero as the imitator observes or underwrites more deals. Since: $\Delta q(0) = \tau$, we can re-express the innovator's precision by:

$$\Delta q(t) = \Delta q(0) \left(1 + \frac{\Sigma}{\Delta q(0)} t \right)^{-1}. \quad (5)$$

The imitator’s entry coincides by definition with the realization of his first deal. The dynamics of the quality advantage, allows us to characterize the timing of the entry. The probability of entry becomes positive as soon as $\Delta q(t)$ becomes smaller than s . As long as $\Delta q(t) > 0$, the innovator has a higher probability of getting the next deal. This advantage of the innovator is decreasing in time. It follows from (5) that his expected advantage disappears faster if the initial advantage $\Delta q(0)$ is smaller and switching costs s are bigger.

The dynamic pattern of Δq is crucial to distinguish the predictions of this model from models of horizontal differentiation only. If imitation is perfect, the loyalty of the client base may still provide the required incentive to innovate, as in the model of Bhattacharyya and Nanda (Bhattacharyya and Nanda, 2000). However, switching costs alone predict that imitation is immediate and that the expected market shares are stationary.

An important measure of the innovator’s advantage is the number of deals after which his superior initial expertise is reduced by half. This measure of the “half life” of the advantage is related to the initial advantage and to the amount of information revealed per deal in a simple way:

$$t_{\frac{1}{2}} = \frac{\Delta q(0)}{\Sigma} \tag{6}$$

4.2 Subsequent Generations of Products

In some equity linked securities we observe that the design of some products relies on earlier ones. In particular, new generations are improvements of their older versions. We incorporate this feature to the model as an improvement in the upper bound of the quality τ of the previous products.

Suppose that a bank invents a product that is an enhancement of an earlier product characterized by quality τ . We assume that this enhancement improves the quality by τ' that the new product has maximum quality of

$$\tau + \tau'$$

After a new product that relies on earlier ones is issued, the second generation process of learning-by-doing can start. For this generation, the competing banks (innovators and imitators) start to acquire knowledge from the precision of the earlier generation, τ (the prior art). For

example, Dividend Enhanced Convertible Stock, or DECS (Figure 4), are a second generation innovation derived from the PERCs.¹²

The security design of subsequent generations is not as innovative as the design of the elders. In other words, the maximum potential value that a new security adds to its issuer is decreasing in the generation number of the security. This implies that

$$\tau' < \tau$$

or since (for any t) $\Delta q(\tau, t)$ is decreasing in τ :

$$\Delta q(\tau', t) = \tau' \frac{1}{1 + \frac{\Sigma}{\tau'} t} < \tau \frac{1}{1 + \frac{\Sigma}{\tau} t} = \Delta q(\tau, t)$$

a given generation's product quality gap is larger than the quality gaps of the products of later generations.

4.2.1 Speed of Entry

The number of deals done by the innovator after which the imitator closes his first imitative deal is a random variable that depends on how innovative or hard to imitate is the original product. More precisely, consider the probability distribution that the imitator closes his first deal anytime after $N - 1$ deals closed by the innovator. This is a cumulative probability function equal to:

$$\Pr(N) = 1 - \prod_{t=1}^{N-1} (\bar{x}_t)$$

where \bar{x}_t is the probability that the innovator closes the t -th deal, i.e.:

$$\begin{aligned} \bar{x}_t &= \min \left(1, \frac{1}{2} + \frac{\Delta q(t)}{2s} \right) \\ \Delta q(t) &= \Delta q(0) \left(1 + \frac{\Sigma}{\Delta q(0)} t \right)^{-1} \end{aligned}$$

Since for every t , $\Delta q(t)$ is increasing in the initial advantage $\Delta q(0)$, than for every N , $\Pr(N)$ decreases in $\Delta q(0)$. This implies that:

¹²Subsequent generations of convertible preferred stock are ACES and PEPS. ACES convert one to one mandatorily after 4 years, but flooring and capping the appreciation of common. PEPS convert mandatorily one to one after 4 years only if the common stock appreciates more than a threshold return.

Proposition 1 *The probability distribution of the time of entry of the imitator at or after the N -th deal is first order stochastically dominated by the distribution of the time of the entry when the initial expertise advantage is larger.*

This implies, for instance, that the expected time of entry of an imitator is lower the lower the initial disadvantage. This stochastic dominance can be verified in the data by comparing the sample distribution of the times of entry of competitors across subsequent generations of innovations within the same family. Indeed, as we argued later generation products should have lower initial expertise advantages relative to earlier ones.

4.2.2 Equilibrium Market Shares

The expected market share of the innovator after M deals plus the monopolistic deal ($t = 0$) is:

$$MS_0(M) = \left(1 + (N - 1) + \sum_{t=N}^M \left(\frac{1}{2} + \frac{\Delta q(t)}{2s} \right) \right) / (M + 1)$$

The expected market share of the imitator after a $M + 1$ deals is:

$$MS_1(M) = \left(\sum_{t=N}^M \left(\frac{1}{2} - \frac{\Delta q(t)}{2s} \right) \right) / (M + 1),$$

The expected market share of the innovator is always larger than the expected market share of the imitator and the difference decreases with the “age” of the security, i.e., with M :

$$MS_0(M) - MS_1(M) = \left(N + \frac{1}{s} \sum_{t=N}^M \Delta q(t) \right) / (M + 1)$$

Since $\Delta q(t) \leq s$ for $t \geq N$ then, at any given period M , if the innovator’s expertise is higher or if the speed of learning of the innovator is smaller or if the switching cost are smaller, then the market share of the innovator becomes relatively larger than the imitator’s. This happens for two reasons. First, the possible entry of the imitator happens later (after more deals are underwritten by the innovator, i.e., a larger N). Second, even after the “entry” of competition, the probability that the imitator obtains the deal in any given period is smaller (larger \bar{x}). Clearly, $MS_0(M) - MS_1(M)$ converges to zero.

Proposition 2 *If next generation products are associated with decreasing incremental innovations then market share convergence occurs faster for later generations.*

5 The Incentives to Innovate

Let N be the first deal that the imitator can get with positive probability. Thus, N solves

$$\begin{aligned} \Delta q(N-1) &> s > \Delta q(N), \\ \Rightarrow N &= 1 + \text{Int} \left[\frac{\Delta q(0)}{\Sigma} \left(\frac{\Delta q(0)}{s} - 1 \right) \right]. \end{aligned}$$

This threshold N is higher the higher the expertise advantage of the innovator and the smaller the switching cost (client loyalty) and slower the information spillover (learning of the imitator).

The total expected profits from innovating must account for four terms: the development cost, the expected profits from the first deal, the expected profits from the periods where the expertise advantage still allows to drive out competition with certainty, and the expected profits in the presence of competition:

$$\Pi_0^e = -F_0 + \pi_M^e + \sum_{t=1}^{N-1} (1-\delta)^t \Delta q(t) + \sum_{t=N}^{\infty} (1-\delta)^t s \left(\frac{1}{2} + \frac{\Delta q(t)}{2s} \right)^2$$

where the evolution of the advantage and the monopoly profits are:

$$\begin{aligned} \Delta q(t) &= \Delta q(0) \left(1 + \frac{\Sigma}{\Delta q(0)} t \right)^{-1} \\ \pi_M^e &= q_0 - \left(c + \frac{s}{2} \right) > \Delta q(0) \end{aligned}$$

The innovator's total profits and his incentives to innovate increase with his initial expertise advantage $\Delta q(0)$ and decrease with Σ , the amount of information that the imitator learns after every deal underwritten for that security by any bank.

The total expected profits from imitating account the expected profits from the period when the probability of obtaining the deal becomes positive:

$$\Pi_1^e = \sum_{t=N}^{\infty} (1 - \delta)^{t_s} \left(\frac{1}{2} - \frac{\Delta q(t)}{2s} \right)^2$$

The imitator's total profits decrease with his initial quality disadvantage $\Delta q(0)$ and increase with Σ .

Since the innovator's profits decrease in Σ , they have incentives to innovate in markets where the precision of the updating process by imitators is smaller. In other words, they will innovate where imitators can extract less information from observing each deal. Such will be the case of highly volatile markets, where the changes in the economic environment will prompt changes in the engineering of each deal, given the product design. This clouds the inference that imitators make about the optimal mapping of the deals' parameters. Innovation should be more frequent in volatile markets not because in such a context firms demand new risk hedging products but because banks have bigger first mover advantages.

In this model the innovator does not have a choice of when to introduce the new product. Clearly, if an innovator develops a new security he may wait to market it when the demand for the product is high. In the context of this model, the innovator may have the design ready but may wait until the client who is in the market is one that can be charged the highest underwriting fee. Since clients are drawn independently, the innovator's equilibrium profits increase if the first deal is with a more loyal client, i.e., a client closer to location 0. Thus, an innovator with a new security design has an incentive to wait until his most loyal client is in need of finance.

Waiting for the most loyal client can be too costly if there is a risk that other competitors may come up with the same innovation. Thus, an underwriter with a new design has incentives to market the innovation to its most loyal client base. Alternatively, banks may tailor the design of their innovations to suit best the needs of their most loyal clientele. For example, the design of their products may be destined to meet the targets of their client's capital structure, or their needs to save taxes.

5.1 Speed of Innovation

We argued above that later generation products typically rely more on prior art and hence can be imitated faster and more accurately. One implication of this was that the half-life of the innovator's advantage is shorter for later generation products. Also in general next generation products are improvements on previous ones. Thus, the actual life span of a security design depends on the speed at which the next generation arrives.

To understand how the life span of a security depends on its generation number, consider this simple setup. Assume that in each period any given bank has some exogenous probability of discovering a later generation security, namely the improvement over the current one. Let the probabilities be δ_0 and δ_1 , which are different in general, for innovator and imitator. We can think that $\delta_0 > \delta_1$ because the innovator has an expertise advantage in the engineering of the current product that gives him a lead in the research and development for a later generation product. The probability that some bank innovates in any given period is

$$1 - (1 - \delta_0)(1 - \delta_1) \equiv \delta.$$

With later generation products, the initial advantage in product engineering of the innovator over the imitator decreases. In this same spirit and for the same reason we can assume also that the one period chance that the imitator is able to develop a new improved product, δ_1 , gets closer to the innovator's chance, δ_0 ¹³. This implies that the probability that an improved security is created, δ , is increasing in the generation number. Since the earlier product is replaced (“cannibalized”) by its improvement, we have that later generation products should last on average less or be replaced faster. Indeed, if δ_0 and δ_1 are constant within a generation, then the expected number of deals (or, alternatively, time periods) before a given product is replaced is $\frac{1}{\delta}$.¹⁴ We conclude that later generation products last less, i.e., that new products arrive faster or after less deals of the previous

¹³The distribution of the time of innovation is geometric. Thus, if an event can occur independently every period with a constant probability p , then the expected number of periods before this event occurs is p^{-1} .

¹⁴We can also allow for δ_0 and δ_1 to increase with every deal. This would speed up the introduction of next generations even more.

ones.

Note that, all other things constant, imitators can enter the market faster for later generations. However, the shortening of the life cycle of later generation securities decreases the chances that imitators complete their first deal. Thus, while it is true that, conditional on being imitated, later generations are imitated faster, it is not clear whether we should see more or less frequent imitation in later generations. If the expected life cycle of a security shortens faster (slower) than the expected imitation time, then we would expect less (more) imitation later in a sequence. This is an empirical issue that we explore later in the paper.

6 Client Base Heterogeneity and Reputation Effects of Innovation

We assumed that the potential clients were uniformly distributed on the unit interval to explore a situation where no bank had an advantage over the other prior to the creation of the new security. After the innovation comes to life the innovator has an advantage over the imitator that eventually fades. In fact, with uniformly distributed clients the situation in the long run returns to the equal sharing of the market, just like before the innovation occurred.

In this section we depart from that equal advantage benchmark and explore the dynamics of the first mover advantage when the two competitors do not have a client base of the same size to begin with. To model this in a simple way, we assume that clients are distributed on the unit line according to a density function of the following kind:

$$f_{\alpha}(x) = \alpha x^{\alpha-1} \quad 0 \leq x \leq 1$$

where α is a positive real parameter. This type of distribution is a subclass of the beta family and it allows us to capture the following features. For $\alpha < 1$ the innovator (located at 0) has a client base advantage. For $\alpha > 1$ the imitator has the client base advantage and for $\alpha = 1$ we are back in the uniform benchmark case of equal client bases. Note that despite the non-uniform distribution, the location of the indifferent client is still $\bar{x} = \min\left(1, \frac{1}{2} + \frac{\Delta q}{2s}\right)$. What changes is the measure of clients are located to his left and to his right.

The expected one period profits of the innovator and the profit difference in all cases are:

$$\pi_0^e = \int_0^{\bar{x}} [(1-2x)s + \Delta q] \alpha x^{\alpha-1} dx = \begin{cases} \Delta q + s \frac{1-\alpha}{1+\alpha} & \text{for } \Delta q > s, \\ \frac{2s}{1+\alpha} \left(\frac{1}{2} + \frac{\Delta q}{2s} \right)^{1+\alpha} & \text{for } \Delta q < s, \end{cases}$$

$$\pi_1^e = \begin{cases} 0 & \text{for } \Delta q > s, \\ \frac{s}{1+\alpha} \left(2 \left(\frac{1}{2} + \frac{\Delta q}{2s} \right)^{1+\alpha} - (1-\alpha) \right) - \Delta q & \text{for } \Delta q < s. \end{cases}$$

where $\pi_0^e - \pi_1^e = \Delta q + s \frac{1-\alpha}{1+\alpha}$.

Proposition 3 *A larger initial clientele of the innovator relative to the imitator, results in higher innovator's profits from the new security and lower profits from imitation.*

This proposition is proved in the appendix. From it we learn that the initial client base can have an important effect on the incentives to innovate. Everything else being equal, it may not be profitable for a bank with a smaller *initial* client base to develop a new product that will later be imitated, whereas it may be profitable for a bank with a larger initial client base. As a result, banks with larger client bases should innovate more often.

The above argument brings us to the relation between innovation and reputation. It is often argued that in the financial sector there are returns to being a leader rather than a follower. Many firms prefer to be clients of a bank that innovates more frequently than other banks. This effect can be captured in this model if we assume that every innovation makes α decrease. If the potential developer of a new cutting edge product can expand its client base, i.e., gain additional clients to do other regular with as a result of enhanced reputation, then it has an additional incentive to develop the product. Not only that, this innovation-reputation effect on the client base can feed back on itself and spur even more innovation. If a bank by creating a new product can later increase its client base for *future* innovations, it will have higher expected profits from its next innovations, because he will be serving a larger initial potential set of clients.¹⁵

¹⁵Morgan Stanley's dominance in convertible preferred stock in the early and mid nineties is a notable example consistent with this prediction.

7 Empirical Evidence Related to the Model

7.1 Summary of Predictions

Here we address how the predictions of the model are consistent with the evidence found in the issues of corporate derivatives. The implications of the model that can be tested empirically are:

Prediction 1 The market share for the innovator's variety of the product is larger than for an imitator's and the difference is decreasing with time.

Prediction 2 If an innovation is an improvement (i.e. is a later generation) of a previous one, the market share advantage of the innovator is smaller and decreases faster than the earlier generation.

Prediction 3 Later generation securities are imitated faster than earlier generations.

To test these predictions we use data from the Securities Data Company's on-line databases of financial transactions. We use all the private and public offerings of equity-linked and derivative corporate securities in the New Issues database and record characteristics such as the name of the issuer, the principal issued, the name of the underwriter and the dates. There are 665 of such issues from 1985 until December of 2002 and involve 51 different securities (innovations) by 30 different lead underwriters. Not all banks compete in all products markets, so there are 98 different bank-security couples. As we argue above, the complexity of the design of corporate derivatives, rather than standard debt or equity, makes it more appropriate to evaluate the predictions of a model with different expertise between underwriters. We also refer to the results of the empirical study by Schroth, that uses the same database (Schroth, 2003), and by (Tufano, 1989).

The model also predicts that later generation products are shorter lived. We verify empirically whether this is true or not.

7.2 Product Groups and Sequences of Innovations

Schroth classifies Equity-Linked securities into product groups (or “families”) and generations within these groups (Schroth, 2003). We consider each one of the 51 different corporate derivatives in the SDC database as an innovation since for each one there is a unique feature that distinguishes it from everything that already existed. Each security has its generation number, which is their order of appearance within its product group. The innovator of a security is the lead underwriter of the first offer ever. Any other bank underwriting deals using the same security is called an imitator.¹⁶

Panel A of Table 1 compares the 11 different product groups for corporate derivatives. Some innovations spur the development of further improvements while others do not. Families with the largest number of improvements (later generations) have been those of convertible preferred equities, and the tax-saving perpetual or convertible securities. Innovations in more standard debt products (RISRS) or zero-coupon convertible debt (LYONS) brought about relatively large and long lasting underwriting markets but do not seem to have provided a fertile ground for subsequent development. The second and third columns of this table suggest that product groups with longer sequences of innovations seem to be associated with more competitors and more innovators. These are expected features of a fertile product group, in which during the sequence more information about the products would have diffused from innovators to potential competitors.

7.3 Evidence on Market Share Dominance

>From Tufano’s study, it is clear that the innovator’s average share of underwriting of all the deals done with his innovation is larger than any of the imitators’ average share (Tufano, 1989). Thus, our first prediction is verified for all corporate securities innovations between 1971 and 1989.

¹⁶Innovative corporate products are classified by Schroth using a compilation of articles in *Investment Dealers’ Digest*, *American Banker*, *Dow Jones Newswires* and others found using the ABI Search Engine (see Schroth, 2003). For every product, there is at least one description in these databases and a reference to an older product which was similar to it. Tom Pratt writes a descriptive article in the *Investment Dealer’s Digest* of almost every corporate security invented.

Using the equity-linked and corporate securities data, Schroth (2003) estimates the demand for the innovators' and imitators' varieties at any point in the securities life cycle. In first place, he confirms that, on average the market demand for the innovator's variety is bigger than for the imitator's in an arbitrary time period. This is a more direct test to our model since it is actually a consistent estimation of the demand function rather than a consistency check through the observed the market shares.

Schroth's study also measures the innovator's advantage over time. Our second and third predictions are also verified by this study: the difference between the innovator's and the imitators demand is decreasing in time, and the time required for convergence is smaller for later generations. Panel B of Table 1 summarizes the results of these two studies.

7.4 The Speed of Imitation and Product Life

Prediction 3 states that imitation is expected to be faster in later generations. Figure 5 plots the empirical cumulative function of the speed at which a security is imitated. The speed of imitation is measured by the number of the deal, in chronological order, at which a given security was imitated. The dotted line is the CDF corresponding to those imitated securities that were first generation products, i.e., the first product in a sequence of related innovations. The solid line is the CDF of the speed of imitation of products that appear in the sequence after the first generation. In this figure we can see that the empirical CDF of the speed of imitation for late generation securities first-order stochastically dominates the one for first generation securities, confirming the increased speed at which the former are imitated.

To get a more precise assessment of the speed of entry across different generations, we fit a hazard rate model where the survival time is the time in days before a given security is imitated. We construct a panel that consists of all the deals from the second to the first imitation of each imitated security and associate the time elapsed between each deal and the first one to time invariant covariates included in the following specification:

$$\lambda_i = \exp\{-(\beta_0 + \beta_1 * \text{generation number}_i + \beta_2 * \text{expected size of the market}_i + \varepsilon)\}; \quad (7)$$

where λ_i is the probability that security i is imitated immediately after time t (measured in days) given that it has not been imitated by time t . We use several proxies for the expected size of the market for a security: the size of the first deal ever, the moving average size of all deals before the observed deal, and the size of the previous deal before the observed one. The larger the expected market, the larger the incentives that imitators would have to introduce their varieties faster. We estimate the parameters β_0, β_1 , and β_2 by maximum likelihood, and their estimated standard errors are consistent in the presence of heteroskedasticity and correlation within securities in the same product group. The baseline hazard rate is determined by ε and to compute our estimates we must choose its probability distribution. We fit the model for three cases, i.e., Weibull, Exponential and the Log-normal.

The best fit of our model happens under the Log-normal assumption. We report the estimates for this case in Table 2. The first column in Panel A shows the benchmark estimates for the parameters in (7), omitting the expected market size control. A higher generation number is associated on average with a larger hazard rate, and thus, with a faster expected time of imitation. The estimate is significantly different from zero at the 99% level. The joint hypotheses that all parameters are zero is also rejected. All the other columns show the results when we use different proxies for the expected size of the market before imitation happened. The estimates and the inference performed with them are the same when we use the size of the first deal, the average size of all deals before prior or the size of the previous deal as proxies. In all these cases the sign of β_1 remains negative and significantly different from zero with at least 95% confidence. We use also as proxies the total size of the market and the total number of deals. Again, the results regarding β_1 do not change but the proxies are insignificant themselves, most likely because they are ex-post measures that do not measure well the expected market size before imitation took place.

The Log-normal assumption implies that the baseline hazard rate, namely the instantaneous chance that a security is imitated conditional on not having been imitated before, is initially zero and increasing. Its good fit is consistent with the following intuition: *ceteris paribus*, it is initially very hard for a competitor to imitate a new security, yet as time passes it becomes easier for

competitors to come in the market with an imitation. By contrast, the fit of the model under the Exponential and Weibull assumptions is poor: the joint hypotheses that all parameters are zero under such assumptions cannot be rejected. In these cases the implied baseline hazard rate is time-invariant (exponential) or decreasing in the whole time domain (Weibull). Indeed, the Weibull (or exponential) hazard rates imply that imitating a security becomes harder (or does not get easier) with time, which is counter-intuitive. The estimation using other distributions belonging to the generalized F class also give positive results. Since these hazard rates have shapes increasing as in the Log-normal case all estimates and goodness of fit measures are basically the same as in the Log-normal case. The latter results are not shown here for parsimony, but are included in a supplement to this paper. Figure 6 illustrates the survival probabilities, i.e., the probability that a security has not been yet imitated, implied by the estimates of the lognormal model.

Panel B shows the estimated median times to the arrival of the first imitative deal, conditional on the generation number of the innovation. We compute these estimates at four different measures of the sample distribution of the average of all deals before the entry of imitators: all the quartiles and the mean. We use the parameters in the fourth column of Panel A. The estimated time of entry of imitation is given by $\frac{1}{\lambda}$. At the median of the previous deals average size, the median imitation time is almost a year. The median imitation time decreases to just over six-months by the fifth generation. At the third quartile, the times are much smaller. For example, a tenth generation security’s predicted median time of imitation is less than two months. This is also depicted in Figure 7.

Table 3 summarizes further characteristics of early and late generation products. The first row confirms a result found by Schroth: the innovator’s market share is larger for a first generation innovator than for later generation innovators (Schroth, 2003).¹⁷In consistency with the previous results, the second row of Table 3 shows that later generation products, if imitated, are imitated

¹⁷The measure of market share used here and by Schroth (Schroth, 2003) is the number of deals that a given bank has underwritten within a product or within a product group divided by the respective total number of underwriting deals. Note that the measure is not the share of the underwritten *principal*. Implicit is the assumption that the amount to finance required by an issuer is given at the time it has to choose its underwriter.

on average much faster than early generation products.

In Section 5.1, we argue that the predictions of the half-life of the innovator's advantage across generations could not be verified from the data because a product typically disappears because a next generation replaces it. Thus, the observed number of deals is rather a measure of the speed of next generation innovations. The instantaneous probability of discovery of the next innovation by either bank should decrease along the sequence of innovations if next generations were marginally decreasing improvements of quality of the previous ones, and if banks acquired more expertise about the product class. Table 3 shows how the life cycle of first generations and later generations differ. Measured by the total number of deals, it is clear that products that improve on the first generation are, on average, shorter lived.

We have seen that imitation occurs faster in later generations if the product is imitated. We may still ask why are some products imitated while others are not. In fact, only 18 of the 51 innovations in this sample were imitated. Table 4 addresses this concern by showing the distributions of imitated and non-imitated products conditional on whether these are a first or a later generation product. First generation products are significantly more likely to be imitated than later generation products. In fact, we can reject the null hypothesis of no association between the imitation and the generation number with a confidence level of 95%. Most likely, this is because later generation products become shorter lived and thus it is less likely to see that an imitator completes its first deal, even if it takes less deals for him to do so.

7.5 Who are the Innovators?

Investment banks that have large clienteles may have a captive market for their new corporate products and this may provide stronger incentives to innovate. In fact, if switching costs were the only source of monopolist rents then we would expect the same banks to innovate very frequently along the sequence. Further, if innovation increases the reputation of a bank as an underwriter, the effect of initial clientele on the incentives to innovate are magnified, predicting a persistence in the selection of the innovator: banks that develop the first generation would be more likely to continue

developing improvements, while other banks are always “relegated” to the role of imitators.

Table 5 shows that, on the contrary, a significant share of the later generation innovations are done by banks that did not develop the group’s first generation. This table takes the 61 combinations of groups and banks in the data and shows how the number of innovations is distributed, conditional on the banks being the group innovator or not. Of the 50 banks that were not the group innovator, 22 innovate at least once after the first generation has been introduced. More precisely, of the 39 innovations that appear after the first generation, 33 are not introduced by the group innovator.

7.6 Further comments

Besides switching costs and expertise advantages, we find in the literature one more explanation of why patents are not necessary for financial innovation. Vikram Nanda and Yeongkul Yun argue that banks coordinate their R&D effort and act as a research joint venture to overcome the free-riding incentives that ultimately eliminates the incentives to innovate (Nanda and Yun, 1995). We believe, however, that this hypothesis does not apply to our data set and the types of securities described in this paper. In first place, our data set and theirs have only one security in common. Second, of the 665 underwriting deals using equity-linked and derivative corporate securities only 13 are underwritten jointly by two lead underwriters. In fact, only once has the underwriting leadership ever been shared in the *first* issue of a security.

8 Conclusion

In this paper we have argued that the development process of new corporate products endows innovators with superior expertise in the structuring of deals for potential issuing firms. This feature is consistent with some stylized facts in the financial innovation literature. Namely, that the innovator has, *ceteris paribus*, a market share advantage in the market developed by his innovation, and that this advantage disappears with time. Beyond the existing evidence (Tufano, 1989, Schroth, 2003) we presented additional evidence on innovations developed more recently. The evidence on

innovations in equity-linked and corporate derivative products allowed us to identify families of innovations and different generations within them. We noticed that the innovator's advantage was smaller and shorter lived for later generations products. Our model is consistent, not only with the existing static evidence, but also with the dynamic patterns that the equity-linked securities innovations exhibit.

The expertise advantage of the innovator that emerges from this evidence makes the innovation more profitable. The innovator is more likely to recoup the development cost and have a positive profit from the innovation despite the absence of patent protection. The resolution of the *State Street Case*, in which the US Supreme Court decided to uphold a patent for a financial business method in 1999 has caused an arms-race-like run on patents by securities firms (Lerner, 2000 & 2004), which has received substantial press coverage recently. *State Street* may have provided unnecessary additional incentives to spend in R&D besides raising significantly the legal costs of patenting every new product often for mere defensive reasons. Whether *State Street* may have introduced welfare improving incentives as well is too early to tell. The effects of *State Street* on the amount of innovation and its profitability for investment banks remain to be seen and then studied.

References

- Allen, F., and D. Gale, (1994). *Financial Innovation and Risk Sharing*, MIT Press, Cambridge, MA.
- Bhattacharyya, S. and V. Nanda , (2000), “Client Discretion, Switching Costs, and Financial Innovation,” *Review of Financial Studies*, 13(4), 1101-1127.
- Boldrin, M. and D. Levine, (2002), “Perfectly Competitive Innovation,” Working Paper, University of Minnesota.
- Eccles, R. and D. Crane, (1988), *Doing Deals: investment banks at work*, Harvard Business School Press, Boston.
- Finnerty, J., (1992), “An Overview of Corporate Securities Innovation,” *Journal of Applied Corporate Finance*, 4, 23-39.
- Hauswald, R. and R. Marquez, (2003), “Loan Portfolio Quality and the Diffusion of Technological Innovation,” Working Paper, American University and University of Maryland.
- Lerner, J., (2000), “Where Does State Street Lead? A First Look at Finance Patents, 1971-2000,” *Journal of Finance*, 57, 901-930.
- Lerner, J. and A B. Jaffe (to be published in November 2004), *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What To Do About It*, Princeton University Press.
- Mason, R., R. Merton, A. Perold and P. Tufano, (1995), *Cases in Financial Engineering: Applied Studies of Financial Innovation*. Prentice Hall, New Jersey.
- Miller, M., (1986), “Financial Innovation: The Last Twenty Years and the Next,” *Journal of Financial and Quantitative Analysis*, 21, 459-471.
- Nanda, V., and Y. Yun, (1995), “Sharing the Limelight: On Why Investment Banks Co-manage Initial Offerings of Innovative Securities with Competitors,” Working Paper, University of Michigan.
- Persons, J. and V. Warther, (1997), “Boom and Bust Patterns in the Adoption of Financial Innovations,” *Review of Financial Studies*, 10(4), 939-967.

Roper, A. S. (1999): "CBO/CLO Ratings, Risk On Radar, S&P Says," *Private Placement Letter*, (December 6th Issue), New York.

Schroth, E., (2003), "Innovation, Differentiation and the Choice of an Underwriter: Evidence from Equity-Linked Securities," Working Paper, International Center for Financial Asset Management and Engineering (FAME).

Thackray, J. (1985): "Corporate Finance in the Golden Age of Innovation," *Forbes*, 136 (April).

Toy, W., (2001), Personal Interview. New York City, February.

Tufano, P., (1989), "Financial Innovation and First-Mover Advantages," *Journal of Financial Economics*, 25, 213-240.

Tufano, P., (1995), "Securities Innovations: A Historical and Functional Perspective," *Journal of Applied Corporate Finance*, 7(4), 90-104.

Tufano, P., (2003), "Financial Innovation," in G. Constantinides, M. Harris and R. Stulz (eds.), *Handbook of The Economics of Finance*, North-Holland, The Netherlands.

Van Horne, J., (1985), "Of Financial Innovations and Excesses," *The Journal of Finance*, 40(3), 621-631.

Appendix 1: Proofs

Proof of Lemma 1. The covariance matrix for the unknown variable, a , and its signal z_1 is

$$Var \begin{pmatrix} a \\ z_1 \end{pmatrix} = \begin{pmatrix} A^{-1} & A^{-1} \\ A^{-1} & A^{-1} + \tau^{-1} + \Sigma^{-1} \end{pmatrix}.$$

If all random variables are normally distributed then the posterior variance of a after receiving t signals is:

$$\frac{1}{\tau_1(t)} = \left(A^{-1} - A^{-1} \frac{1}{A^{-1} + \tau^{-1} + (t\Sigma)^{-1}} A^{-1} \right)$$

That is, the imitator updates the normal distribution of his estimator of a using the signals from each observed deal and Bayes Rule, and the posterior precision of the imitator after t deals, i.e., t signals, is:

$$\tau_1(t) = A + \tau \frac{\Sigma}{\Sigma + \frac{\tau}{t}}$$

Since the developer receives the signal $z_0 = (a + \varepsilon_0)$ and hence has precision:

$$\tau_0 = A + \tau$$

then the difference in quality between innovator and imitator is:

$$\Delta q(t) = \tau_0 - \tau_1(t) = \tau \frac{1}{1 + \frac{\Sigma}{\tau} t}.$$

■

Proof of Proposition 3. In all cases the innovators one period profits are decreasing in α . A small α that is a higher initial client base generates higher revenues from the innovation. Similarly for the imitator, because his profits increase in α :

$$\frac{\partial \pi_1^e}{\partial \alpha} = \frac{2sB^{1+\alpha}(1 + \ln B) + 2\alpha}{(1 + \alpha)^2} > 0$$

where: $B = \left(\frac{1}{2} + \frac{\Delta q}{2s} \right) > \frac{1}{2}$. ■

Appendix 2: The case of more than one imitator

To extend the model to more imitators, we consider here the case of two imitators and the innovator that are located at the extremes of an equilateral triangle (Figure A2.1). The extension to more than two imitators can be done in a similar way, e.g. we would consider a tetrahedron for three imitators and higher-dimensional abstract extensions of it for more than three imitators.

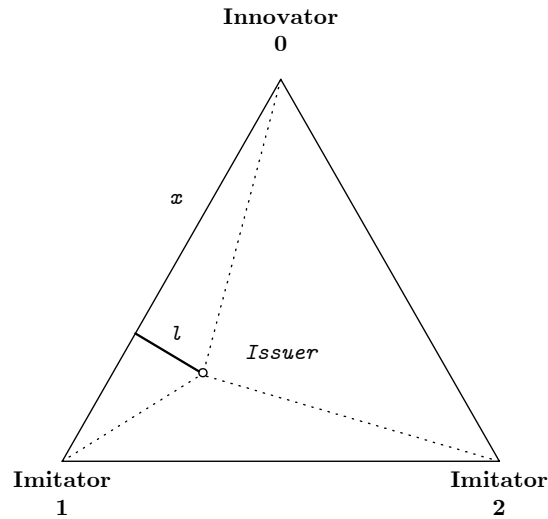


Figure A2.1.: This figure illustrates the type of imperfect competition in the model extended to two imitators. Issuers lie inside an equilateral triangle according to their degree of loyalty to either bank. The three underwriters are located at the extremes, and the position of an issuer relative to the three determines the loyalty ranking and intensities to either. The closer is the issuer to a given bank, the smaller is $x^2 + l^2$, the more loyal it is to it and the more costly it is to hire any other bank as its underwriter

Since the two imitators have the same quality, the imitator farthest from the issuer never obtains the deal. Thus, without any loss of generality (WLOG) we can always consider for any issuer only the competition for an underwriting mandate between the innovator and the imitator which is closest to the issuer. WLOG assume that the issuer happens to be located on the left part of the triangle, so the competition is between the innovator and the left-hand side (LHS) imitator. The values to the potential issuer located at $\mathbf{x} = (x, l)$ in the picture of choosing the innovator or the

left hand side imitator (LHS) as an underwriter are respectively given by:

$$\begin{aligned} u_0(x) &= q_0 - p_0 - sd(\mathbf{x}, 0), \\ u_1(x) &= q_1 - p_1 - sd(\mathbf{x}, 1), \end{aligned}$$

where d is the euclidean distance, so

$$\begin{aligned} d(\mathbf{x}, 0) &= \sqrt{x^2 + l^2} \\ d(\mathbf{x}, 1) &= \sqrt{(1-x)^2 + l^2} \end{aligned}$$

Let $\hat{x}(l)$ be the location of the client indifferent between choosing either whenever both offer a deal at marginal cost; i.e., $p_0 = p_1 = c$. The indifferent client $\hat{x}(l)$ solves the equation

$$\sqrt{x^2 + l^2} - \sqrt{(1-x)^2 + l^2} = \frac{\Delta q}{s}.$$

After rearranging and squaring both sides twice we obtain the standard equation of the hyperbola with vertex on $\left(x = \frac{1}{2} + \frac{\Delta q}{2s}, l = 0\right)$, namely

$$\frac{\left(x - \frac{1}{2}\right)^2}{\left(\frac{\Delta q}{2s}\right)^2} - \frac{l^2}{\left(\left(\frac{1}{2}\right)^2 - \left(\frac{\Delta q}{2s}\right)^2\right)} = 1, \quad (8)$$

with $\frac{\Delta q}{2s} < \frac{1}{2}$. Indeed, by definition the hyperbola is the locus of all points in the plane that have the same difference between the distance to two fixed points. In our case, the foci are $(0, 0)$ and $(1, 0)$ and the difference: $\frac{\Delta q}{s}$.

We need to show now that the innovator's region of influence, i.e., the region of clients that will choose the innovator, is increasing in the innovator's relative advantage, Δq . All the results of the two competitors case follow in this case if the previous statement is true. To see that it is, note that an increase in Δq in equation (8) while l is kept constant required a higher $\hat{x}(l)$ to preserve the equality. Hence, a higher expertise advantage implies a larger region, i.e. a higher probability to get the next deal. Figure A2.2 below shows this comparative statics in Δq by representing the half equilateral triangle with the innovator at $(0, 0)$ and the LHS imitator at $(1, 0)$.

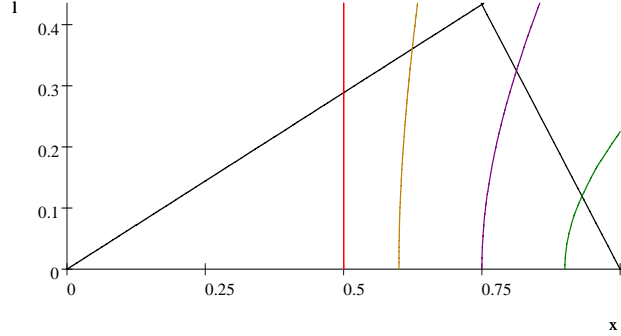


Figure A2.2.: This figure shows the thresholds locations for different values of the innovator’s advantage relative to the magnitude of switching costs. $\frac{\Delta q}{s}$ is set to 0, 0.2, 0.5, and 0.8. Issuers to the right of each line choose the innovator as their underwriter and issuers to the right hand side choose the imitator. The innovator’s clientele increases with the size of his initial expertise advantage.

The hyperbola arches in the picture represent the thresholds to the left (right) of which the innovator (the LHS imitator) can obtain the deal. These thresholds move to the right the higher is the quality advantage of the innovator, that is, for any given l , a larger advantage implies a larger x . The benchmark case of no expertise advantage ($\Delta q = 0$) gives the constant solution or degenerate hyperbola $\hat{x} = \frac{1}{2}$. If the innovator’s quality advantage is high relative to the clientele effect, i.e. $\frac{\Delta q}{s} > 1$, then the “indifferent” client curve lies outside the unit interval in the x axis, which means that the innovator gets the next deal for sure. Still, the presence of the potential imitator puts a bound on the markup that the innovator can obtain.

Figure A2.3 below shows how the three firms have the same market share equal to $1/3$ if there is no quality advantage (dotted lines). Whenever the innovator has a quality advantage (continuous lines), then $x = \frac{\Delta q}{2s} + \frac{1}{2} > \frac{1}{2}$ and the innovator has the largest market share.

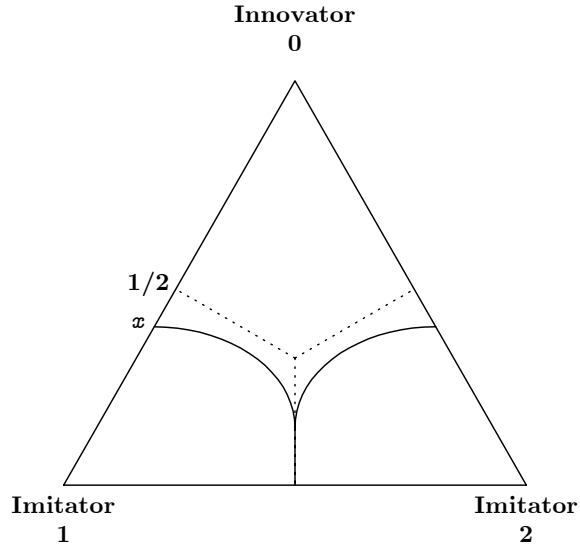


Figure A2.3.: This Figure shows the areas of potential issuers that each bank can attract. The dotted lines denote the areas with equal sharing of the market, when the innovator does not have an advantage over the imitators. The solid lines delimit these areas when the innovator has an expertise advantage.

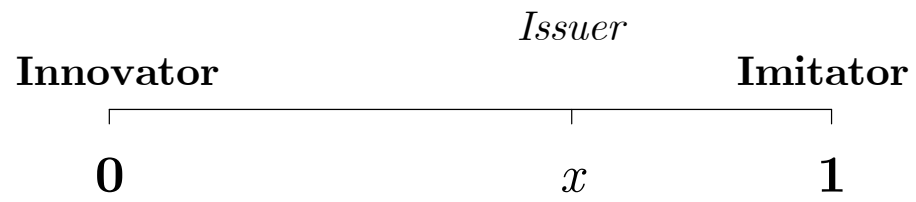
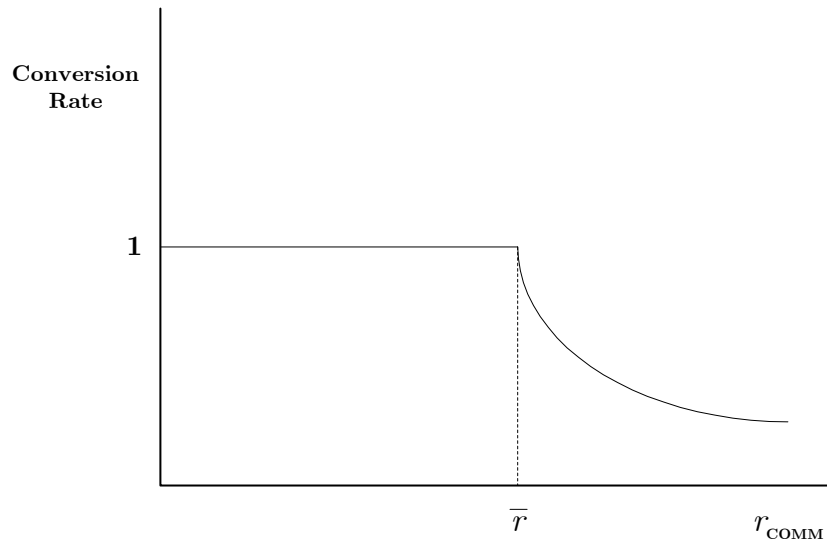


Figure 1: Location of banks and potential clients



PERCS (Preferred Equity Redemption Cumulative Stock)

- Innovator: Morgan Stanley, 6/1991.
- Imitators: Merrill Lynch, Dean Witter.
- Mandatory Conversion in 3 years.
- High dividend yield (>8%)
- \bar{r} between 25 - 40%.

Figure 2: The conversion ratio of Preferred Equity Cumulative Stock (PERCS) as a function of the returns of the underlying common stock.

Figure Captions

Figure 1: This figure illustrates the type of imperfect competition in our model of the market of corporate underwriting. Issuers lie along a unit interval according to their degree of loyalty to both banks. The two underwriters are located at the extremes, and the closer is an issuer of type x to a given bank, the more loyal it is to it, i.e., the more costly it is for the firm to hire the other bank as its underwriter.

Figure 2: This figure plots the conversion rate of a Preferred Equity Redeemable Stock (PERCS), as a function of the returns of the underlying common stock. Each unit of this preferred stock converts mandatorily after 3 years to one unit of common stock unless the stock appreciates above a cap of \bar{r} percent. If after 3 years the common stock appreciates above the cap, PERCS convert to less than one unit of common

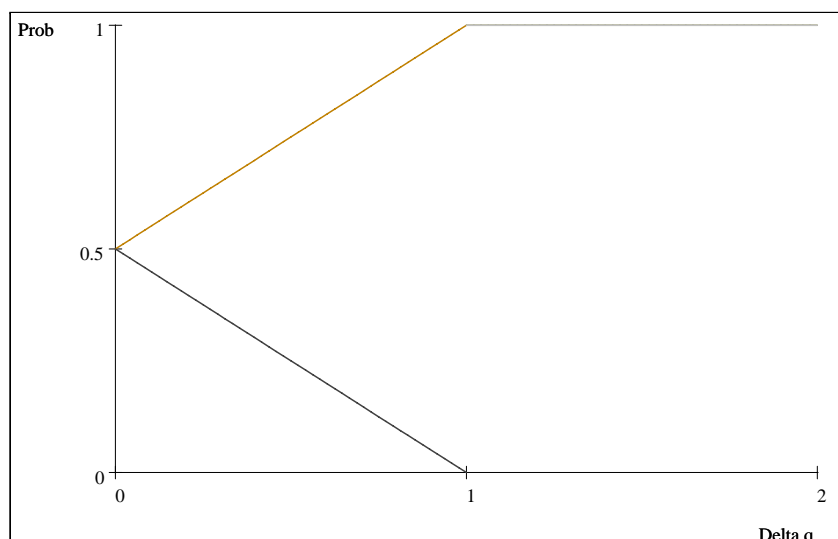
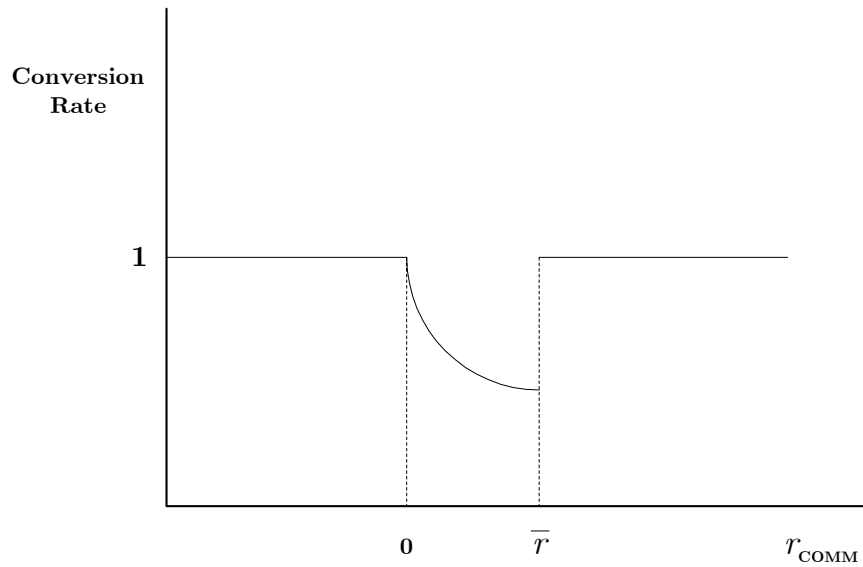


Figure 3: Probability of obtaining the deal of Innovator and Imitator

such that their conversion value is that of a stock that has appreciated by \bar{r} percent.

Figure 3: This figure plots the probabilities that the next underwriting deal will be signed by either the product innovator or its imitator. The next client that will be in the market seeking to sign an underwriting deal is drawn a random from a uniform distribution. The solid line plots the probability that the client chooses to deal with the innovator, as a function of the difference between the quality of the underwriting provided by the innovator or the imitator. The dashed line plots the probability that the client chooses the imitator as its underwriter. *Ceteris paribus*, if the innovator and the imitator can offer the same quality underwriting then the probability that either gets the next deal is 0.5. If the quality differential is higher then the probability that the innovator gets the next deal increases (and the imitator's probability decreases). If the quality differential is high enough, any client will prefer the innovator, and the probability that he gets the next deal is one.

Figure 4: This figure plots the conversion rate of a Dividend Enhanced Convertible Stock (DECS), as a function of the returns of the underlying common stock. Each unit of this preferred stock converts mandatorily after 3 years to one unit of common stock unless the stock appreciates within 0 and \bar{r} percent.



DECS (Dividend Enhanced Convertible Stock)

- **Innovator: Salomon Brothers, 6/1993.**
- **Imitators: Lehman Brothers.**
- **Mandatory Conversion in 3 years.**
- **Lower dividend yield.**
- **\bar{r} between 20 and 22%.**

Figure 4: The conversion ratio of Dividend Enhanced Convertible Stock (DECS) as a function of the returns of the underlying common stock.

If the common stock appreciates within these boundaries in 3 years, then DECS convert to less than one unit of common such that their conversion value is that of the stock's price at the issue date.

Figure 5: This figure plots the empirical cumulative function of the speed at which a security is imitated. The speed of imitation is measured by the number of the deal, in a chronological ordering, at which a given security was imitated. A security is said to be imitated if a banker different from the innovator also underwrites corporate issues using the same product structure. The dotted line is the CDF corresponding to those imitated securities that were first generation products, i.e., the first product in a sequence of related innovations. The solid line is the CDF of the speed of imitation of products that appear in the sequence after the first generation.

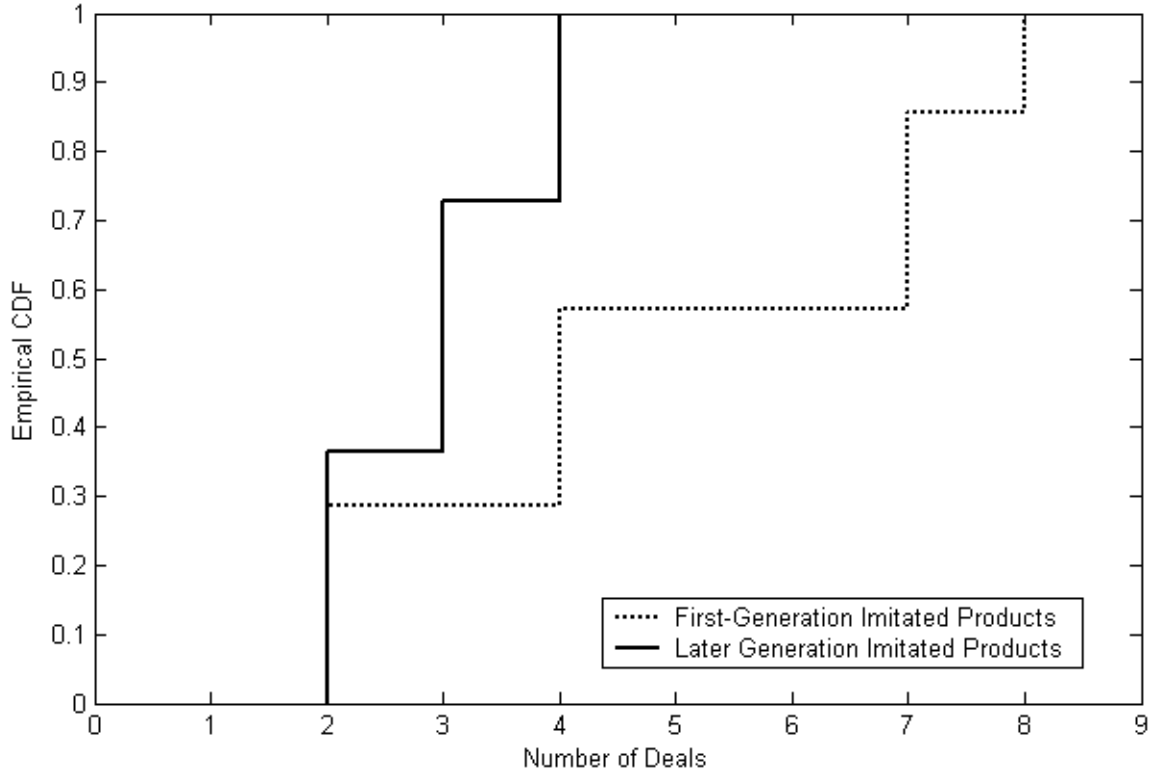


Figure 5: Plot of Empirical Cumulative Density Functions for the Speed of Imitation Conditional on the Generation Number of the Product

Figure 6: This Figure shows the probabilities that a security is not imitated before t days as from the date of its first issue. The probability that imitation time, N , occurs after t , i.e., the survival rates $S(t) = \Pr(N > t)$, are measured in the vertical axis and shown as a function of time, which is shown in the horizontal axis. These are given by $S(t) = \Phi(-\frac{1}{\hat{\sigma}} \ln(\hat{\lambda}t))$, where $\hat{\lambda}$ is the estimated imitation hazard rate, which is itself obtained from the estimated hazard rate model $\hat{\lambda} = \exp(-6.297 + 0.133 * \text{generation} + 0.002 * \text{mean size of prior deals})$, and $\hat{\sigma} = 1.273907$. The thick solid plot corresponds to first generation securities. The thin solid plot corresponds to second generation securities. The dashed plot corresponds to 5th generation securities and the dotted plot to 10th generation securities.

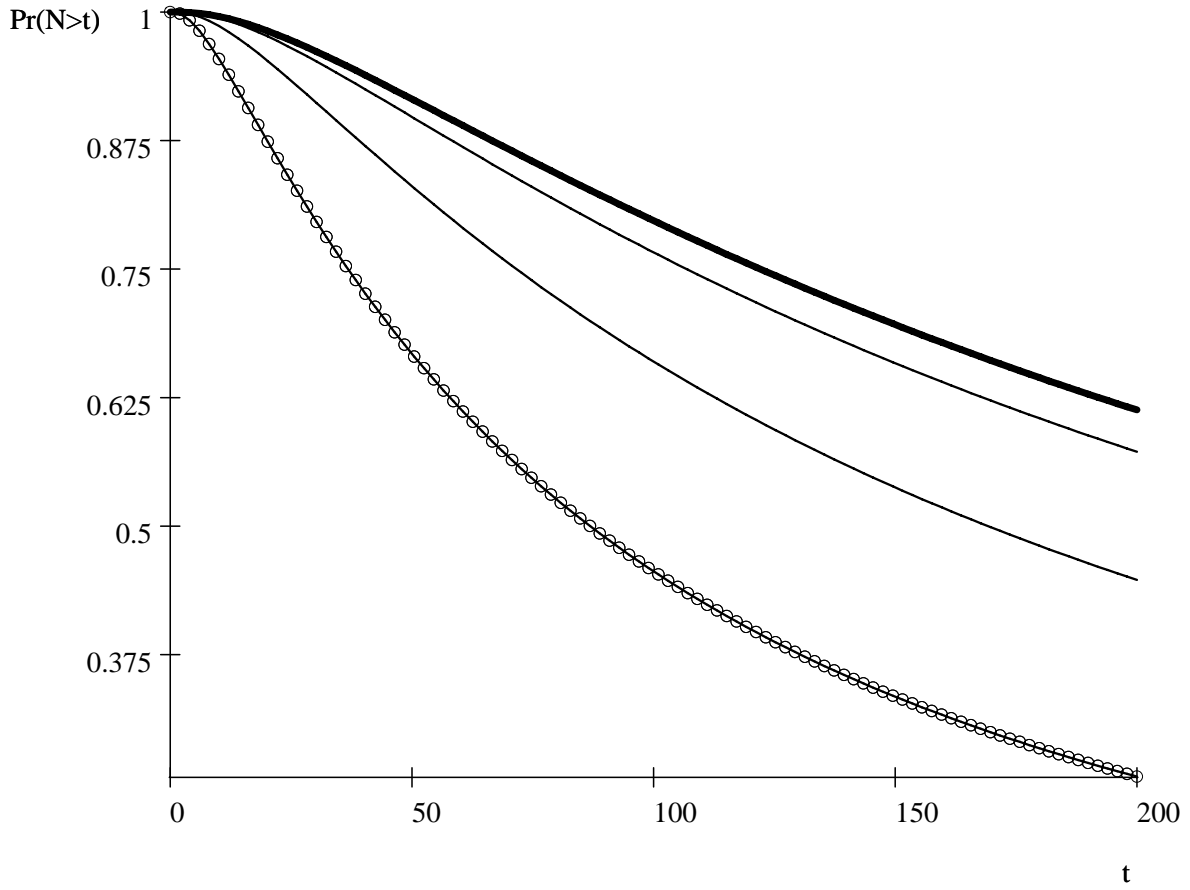


Figure 6: The probability that a security is not imitated before date t after its first issue, conditional on the generation number of the innovation. The thick solid plot corresponds to first generation securities. The thin solid plot corresponds to second generation securities. The dashed plot corresponds to 5th generation securities and the dotted plot to 10th generation securities.

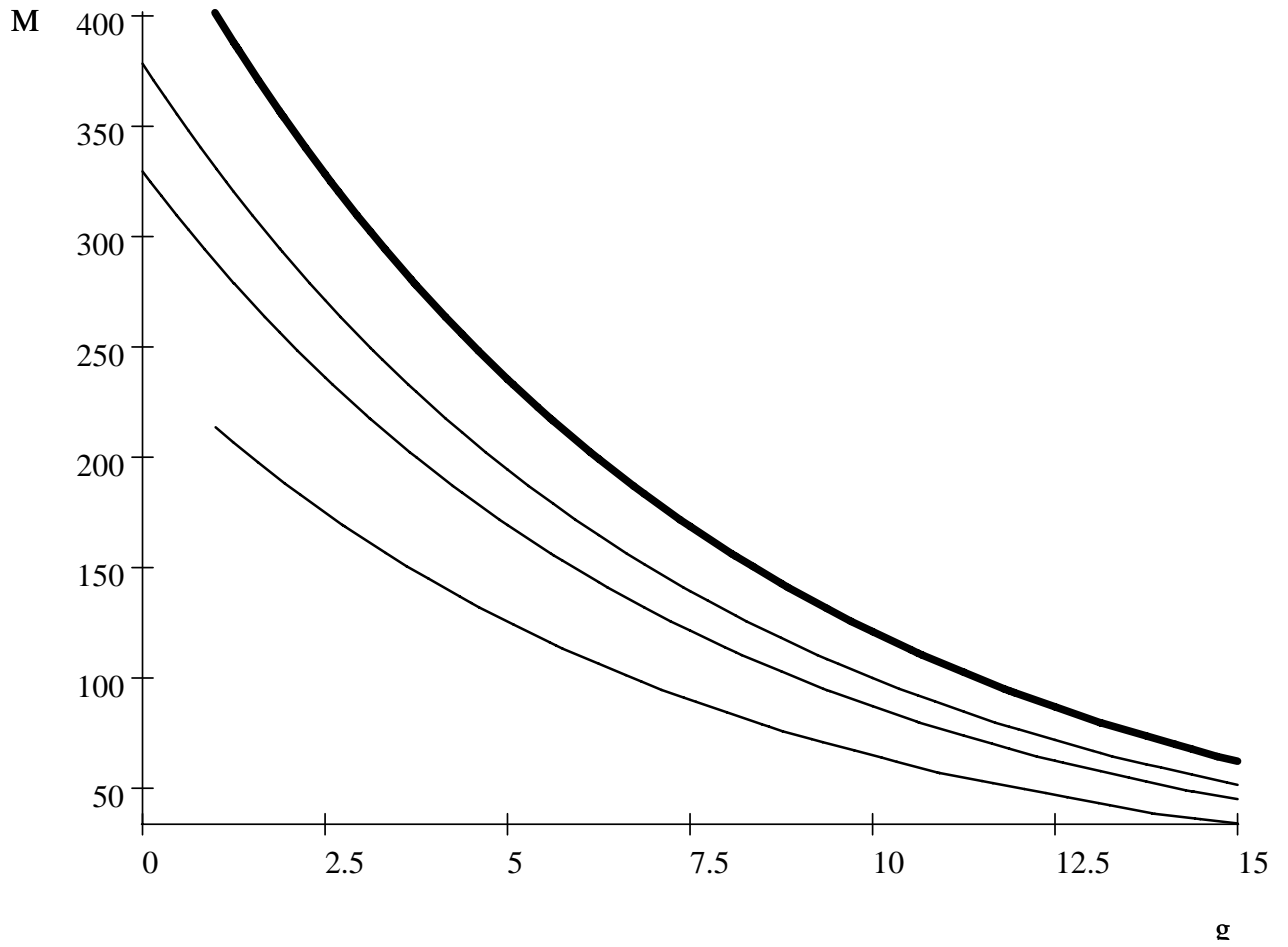


Figure 7: The median times of imitation as a function of the generation number of the security, estimated using a lognormal duration model. The thick solid plots the median times when we use the 1st quartile of the sample distribution of the size of the first deal of the security. The thin solid plot uses the median. The dashed plot uses the mean and the dotted plot uses the 3rd quartile.

Figure 7: This Figure plots the estimated median times of imitation as a function of the generation number of the security. The median predicted time is shown on the vertical axis (Med t) and the generation number, g , on the horizontal axis. The estimated median times, \widehat{M} , are given by $\frac{1}{\widehat{\lambda}}$, where $\widehat{\lambda}$ is the estimated imitation hazard rate, which is itself obtained from the estimated hazard rate model $\widehat{\lambda} = \exp(-6.297 + 0.133 * \textit{generation} + 0.002 * \textit{mean size of prior deals})$. The thick solid plots the median times when we use the 1st quartile of the sample distribution of the size of the first deal of the security. The thin solid plot uses the median. The dashed plot uses the mean and the dotted plot uses the 3rd quartile.

Table 1: Evidence on Innovation and Imitation from Previous Studies

Panel A: Summary of the Classification of Equity -Linked and Other Corporate Derivatives (source: Schroth, 2003).			
Product Group	Number of Products (i.e., Generations)	Number of Distinct Innovators	Number of Underwriters
1. Debt Products	1	1	2
2. Convertible Debt (Zero Coupon)	1	1	3
3. Convertible Debt (Dividend Paying)	2	2	2
4. Convertible Preferred Stock	15	8	15
5. Short-term, Income-Deferring Products	1	1	4
6. Perpetual, Income-Deferring Products	9	4	9
7. Convertible, Income-Deferring Products	7	5	10
8. Index-Tied Principal Appreciation	8	5	8
9. Stock Tied-Principal Appreciation	4	4	6
10. Privatization Exchangeable Debt	1	1	1
11. Corporate Pass-throughs	1	1	1

Panel B

Summary of Evidence on First-Mover Advantages by Corporate Securities Innovators

Study	Description and Methodology	Effect of Selected Variables
1. Tufano (1989)	Regressions of underwriters historical product market shares on reduced form exogenous variables, including: - a dummy variable for the innovator	Positive and Significant
2. Schroth (2003)	Instrumental Variables estimations of the issuers choice for an underwriter. Choice depends on controls and: - an innovator dummy 47 - innovator dummy interacting with time - innovator dummy interaction with time and generation number	Positive and Significant Negative and Significant Negative and Significant

Table 2: Regression Analysis of the Duration before Securities are Imitated

Panel A

The dependent variable is the time in days elapsed after the first deal of the security was made.

The estimated Hazard Rate Model is: $-\ln \lambda = \beta_0 + \beta_1 * generation + \beta_2 * expected\ market\ size + \varepsilon$

All estimates are computed with 48 observations of 18 imitated securities.

The distribution of ε is Lognormal.

Base Model	Proxy for expected market size: proceeds (in US\$ M) of					
	First deal	Previous deal	Mean of all prior deals	Total of all deals	Total number of deals	
Constant	5.822 (0.192)***	6.296 (0.433)***	6.299 (0.240)***	6.297 (0.385)***	5.895 (0.223)***	6.079 (0.314)***
Generation Number	-0.130 (0.047)***	-0.140 (0.061)**	-0.130 (0.062)**	-0.133 (0.063)**	-0.137 (0.045)***	-0.142 (0.039)***
Expected Market Size		-0.001 (0.000)**	-0.002 (0.000)***	-0.002 (0.001)***	-5.75e-06 (6.45e-06)	-0.009 (0.007)
χ^2 Wald statistic	7.57***	7.88**	41.63***	9.64***	12.00***	33.010***

Panel B

Hazard Rate: $\lambda = \exp(-6.297 + 0.133 * generation + 0.002 * mean\ size\ of\ prior\ deals)$

Median time of First Imitative Deal: $M = \frac{1}{\lambda}$

Generation Number	Median Time, evaluated at size of average prior deal =			
	1st quartile	Median	Mean	3rd quartile
1	400.97	331.15	288.38	213.48
5	235.28	194.31	169.21	125.26
10	120.83	99.79	86.90	64.33
15	62.05	51.25	44.63	33.04

Each observation in Panel A consists of the time in days after innovation at every deal before the first imitation paired with the generation number and the proxy for the expected market size. The panel includes all 18 imitated equity-linked and derivative securities between 1985 and 2002. The parameters are estimated by maximum likelihood, and the estimates of their standard errors are shown below them in brackets. These estimates are consistent in the presence of heteroskedasticity and within product group correlation. Estimates followed by *** are significant to the 0.01 level, by ** to the 0.05 level, and by * to the 0.1 level.

Table 3: Summary Statistics of the Speed, Duration and Innovator’s Market Shares for the Innovative Equity-Linked Securities

		Observations	Mean	Std. Dev.	Median
1. Market Share of Product Innovator	First Generations	7	0.74	0.27	0.82
	Later Generations	11	0.55	0.27	0.60
2. Speed of Imitation (deal number of first imitation)	First Generations	7	4.86	2.48	4
	Later Generations	11	2.91	0.83	3
3. Product Life (measured in deals)	First Generations	11	19.81	28.16	9
	Later Generations	39	11.39	18.47	5

The Market Share of the Product Innovator is the number of deals underwritten by the security innovator divided by the total number of deals underwritten with that security. The Speed of Imitation is the issue number (in chronological order) of the first deal by an imitator of a given security. The Product Life is the total number of issues (underwriting deals) of a given security.

Table 4: Distribution of Imitated and Non-Imitated Products, Conditional on their Generation

	Non-Imitated Products	Imitated Products	Total
1. First Generation Securities	4	7	11
	36.36	63.64	100.00
2. Later Generations	29	11	39
	72.50	27.50	100.00

Pearson $\chi^2 = 4.9332$; P-value: 0.026

There are 51 securities, 11 of which are a first generation product (one per group). The Pearson χ^2 statistic corresponds to the test whose null hypothesis is that there is no statistical association between the two binary variables. The first row at each numeral shows the data counts; the second the row percentages.

Table 5: Distribution of the Number of Innovations by Banks Competing in Each Product Group

Type of Banks	Number of Innovations in the Product Group					Total
	0	1	2	3	4	
1. Not the Group Innovator	28	16	3	1	2	50
	56.00	32.00	6.00	2.00	4.00	100.00
2. Group Innovators	0	8	1	1	1	11
	0	72.73	9.09	9.09	9.09	100.00

There are 61 bank-product group observations. 11 correspond to those banks that developed the first generation in each group. The rest correspond to any other bank that competed in the group, either as an imitator or an innovator of later generations. The first row at each numeral shows the data counts; the second the row percentages.