

oper's Expertise and ynamics of Financial ation:

y and Evidence

ERRERA

SCHROTH

per N° 124

onal Center for Financial Asset Management and Engineering



ME

F



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

Helios Herrera

ITAM[†] HEC - University of Lausanne and FAME[‡]

Enrique Schroth

First Draft: September 2003 This Draft: October 2004

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.

"There is at least a perception that the first mover is more familiar with the product he issues than the imitator."

(William Toy, Managing Director at CDC Capital Inc.)

*We are grateful to Hendrik Hakenes, Erwan Morellec, Preston McAfee, Pascal St-Amour and Elu Von Thadden for their very useful comments. Enrique Schroth acknowledges gratefully the financial support provided by the National Center for Competency in Research (FINRISK, Switzerland).

[†]Centro de Investigacion Economica, Av. Camino a Santa Teresa No.930, 10700 Mexico D.F.Tel: +52-55-5628-4000 ext.2961. Fax: +52-55-5628-4958. E-mail: helios@itam.mx.

[‡]Route de Chavannes 33, 1007 Lausanne, Switzerland. Tel.: +41 21 692 3346. Fax.: +41 21 692 3435. E-mail: enrique.schroth@hec.unil.ch

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). Similarly, in their study of the investment 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.

As we mentioned before, to our knowledge very few researchers have proposed reasons why firms may have a preference for the innovator rather than the imitator.³Our 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 provides a characterization of 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. In particular, 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 life cycle of new financial products usually involves the gradual erosion of the innovator's profits (see Van Horne, 1985).

The theoretical part of the paper is a duopoly model, i.e., an innovator vs. an imitator, that gives the comparative statics result that the smaller the innovator's quality advantage the faster the imitators 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 comparative statics 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. Interestingly, we find that the theoretical predictions on the speed of imitation match the empirical evidence on equity linked securities that we present.

Our paper is closely related to the work by Sugato Bhattacharyya and Vikram Nanda (Bhattacharyya and Nanda, 2000). They highlighted first the role of switching costs in financial innovation: if a potential innovator has already clients for whom switching banks is costly, then he can serve this clientele with prices over marginal costs, even when competitors are in the market offering identical products. Then, a broader or more loyal client base increases the incentives to innovate. Switching costs are certainly important to promote innovation because they imply that underwriting provided by two different banks cannot be substituted perfectly, thus ensuring that even perfect imitation does not eliminate the innovator's profits. However, switching costs alone do not eliminate the free-rider problem: regardless of the profitability of innovation, there is no advantage of being an innovator. Ceteris paribus, any potential innovator would rather be an imitator than develop the product himself. In other words, the advantage belongs to the second mover rather than to the first in their model. 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. Namely, 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, 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. Within this group of large banks, most banks do appear sometimes as leaders and other times as followers, and rarely do we observe small banks participating as either innovators or imitators. Thus, while it is clear that innovators could appear to have large market shares because of their large initial clienteles, it is not clear why large banks also have small market shares as imitators. Recently, Enrique Schroth estimated the demand function for a given underwriter and found that the leadership is systematic to the innovator, even after controlling for the size of the clientele of the bank (Schroth, 2003). An implication of the size of initial market shares is that a bank may appear in the long run to monopolize the introduction of future generations. As his market share grows due to a successful *current* innovation, his client base for *future* innovations increases. Morgan Stanley's dominance in

convertible preferred stock in the early and mid nineties is a notable example consistent with this prediction.

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 will innovate more often in markets where inference about the optimal engineering by the imitator is clouded with a changing economic environment, i.e., a higher volatility. Innovation should be more frequent in volatile market not because in such a context firms demand new risk hedging products but because banks that innovate would expect larger first-mover advantages there. In the model by Bhattacharyya and Nanda, the higher frequency of innovation in volatile markets is also due to a supply factor (Bhattacharyya and Nanda, 2000). In their case, it is the increase in the cost to issuing firms of delaying the adoption of innovation.

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.

Our analysis also addresses the effects of the new legal environment that favors patents for business methods innovations in the US. Josh Lerner has showed that there has been a sharp increase since 1999 in the applications and awards for patents in the subclasses where financial innovations are filed (Lerner, 2000), so it is natural to ask how patenting is going to change the incentives to innovate and the overall pace of innovation in financial products. We use a simple model where two banks race to develop the next security, and the first to arrive at the discovery enjoys the first-mover advantage. The loser however, can share part of the rents of R&D because he can reverse-engineer the innovator's design. In this context, we find that some level of profit sharing is optimal, since this softens the competition to be the winner of the innovation race. The possibility to reverse-engineer perfectly makes free-riding too attractive and there is too little R&D in equilibrium. On the other extreme, a patent regime is not optimal because banks try to outspend their rivals since the patent race is a winner-take-all race, and there is too much R&D in equilibrium. We argue that the current situation may have driven the financial services industry further away from the optimal level of profit sharing than the previous situation without patents.

Robert Hauswald and Robert Marquez have asked what is the optimal disclosure level that regulators may want to force on the innovators of proprietary credit assessment technologies (Hauswald and Marquez, 2003). In their model, the innovator screens better loan applicants and is therefore more competitive in the credit market. They find then that it is also optimal to reveal some of the innovator's acquired skills to competitors so as to toughen competition for borrowing clients, and increase their surplus. However, while some degree of profit sharing is optimal, they argue that there is a rationale for patents if the regulator is able to choose the level of disclosure every time *after* an innovation occurs. In such case, a regulator cannot commit credibly not to force full disclosure *ex-post*, and this affecting severely the incentives to innovate *ex-ante*. A patent system then may be the only way to keep the regulator from forcing excessive disclosure. Our model differs from theirs because there is a race for the innovation, so an innovation of a given aggregate value can be achieved at different R&D costs in equilibrium. Also, their rationale for patent protection is ruled out in our setup because we do not have a regulator with full discretion to force any level of disclosure after innovation. We believe this assumption is reasonable for corporate products innovation given that the SEC does not have a policy of changing its rules of disclosure discretionally.

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 discusses the effects of foreclosing imitation with a patent regime on the speed of innovation and the expenditures in R&D. Section 9 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 = 0and the imitator is i = 1. 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 is 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, $F_1 = 0.4$ The 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).

	Issuer			
Innovator		Imitator		
Γ	1			
0	x	1		

Figure 1: Location of banks and potential clients

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:

$$u_0(x) = q_0 - p_0 - sx,$$

 $u_1(x) = q_1 - p_1 - s(1 - x),$

where s is the cost per unit of distance of choosing a variety located away from the preferred one.⁵

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 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.

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:

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

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

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 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, \overline{r} . The contract also must specify the dividends paid and the sale price (see Figure 2).⁶



- \overline{a} between 25 40%
- \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.

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.⁷Deal 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 equitylinked 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 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 $\Delta q \equiv q_0 - q_1 \ge 0$ be the quality differential.

3 Equilibrium

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.

3.1 Monopoly

Since the innovator has monopoly power on its first deal, the highest price it can charge is the one that makes the issuing firm indifferent between underwriting the deal or not. If the reservation value is zero then:

$$u_0(x) = 0$$

$$\Rightarrow p_0 = q_0 - sx$$

The monopolist expected profits are:

$$\pi_M^e = \int_0^1 [q_0 - sx - c] dx$$

= $q_0 - c - \frac{s}{2}.$

To guarantee that the ex-post profit of the innovator and the imitator from a deal with any potential client are positive, we need to assume that q_1 , which is smaller than q_0 , is large enough, that is,

$$q_1 \ge c + s. \tag{2}$$

3.2 Oligopoly

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, for

$$p_0 = p_1 = c,$$

 \hat{x} solves $u_0(\hat{x}) = u_1(\hat{x}).$

Solving, we obtain

$$q_0 - c - s\hat{x} = q_1 - c - s(1 - \hat{x})$$

 $\hat{x} = \frac{1}{2} + \frac{\Delta q}{2s}.$

Note that if 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, which means that 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.

Define \overline{x} such that, for any client $x < \overline{x}$, the innovator can undercut the imitator below marginal cost and still get the deal (while making a profit). Thus, the range of clients of the innovator for the next deal is:

$$x \in [0, \overline{x}) \tag{3}$$

$$\overline{x} = \min(1, \hat{x}) \tag{4}$$

Figure 3 illustrates the probability of obtaining the deal of the Innovator and of the Imitator as a function of the quality advantage Δq when s = 1.

Consider a client in the region $x \in (0, \overline{x})$. Due to the preference for the innovator's variety, the innovator can undercut the imitator to its marginal marginal cost and attract the client with a price p_0 such that:

$$q_0 - p_0 - sx = q_1 - c - s(1 - x)$$

 $(p_0 - c) = (1 - 2x)s + \Delta q$



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

Given that in every period one firm is drawn uniformly, the expected one period profits of the innovator are:

$$\pi_0^e = \int_0^{\overline{x}} (p_0 - c) dx = \int_0^{\overline{x}} [(1 - 2x)s + \Delta q] dx$$
$$= (\overline{x} - \overline{x}^2)s + \overline{x}\Delta q = \overline{x} ((1 - \overline{x})s + \Delta q)$$

And the imitators expected one period profits are:

$$q_0 - c - sx = q_1 - p_1 - s(1 - x)$$

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

$$\pi_1^e = \int_{\overline{x}}^1 (p_1 - c)dx = -\int_{\overline{x}}^1 [(1 - 2x)s + \Delta q] dx$$

$$= (\overline{x} - \overline{x}^2)s + \overline{x}\Delta q - \Delta q = (1 - \overline{x})(\overline{x}s + \Delta q)$$

The innovator's expected profits are higher, due to the higher quality. In fact, the difference in profits is equal to the quality differential:

$$\pi_0^e - \pi_1^e = \Delta q$$

so for any positive quality differential the innovator is always ex-ante better off than the imitator.

In the case of high advantage relative to switching costs:

$$\Delta q > s \implies \overline{x} = 1$$

$$\begin{cases} \pi_0^e = \Delta q \\ \pi_1^e = 0 \end{cases}$$
(5)

In the case of low advantage:

$$\Delta q < s \implies \overline{x} = \frac{1}{2} + \frac{\Delta q}{2s}$$

$$\begin{cases} \pi_0^e = s \left(\frac{1}{2} + \frac{\Delta q}{2s}\right)^2 \in \left[\frac{s}{4}, s\right] \\ \pi_1^e = s \left(\frac{1}{2} - \frac{\Delta q}{2s}\right)^2 \in \left[0, \frac{s}{4}\right] \end{cases}$$

$$(6)$$

The profits in both cases can be summarized in Figure 4 for the case where s = 1.



Figure 4: One Period Profits of Innovator and Imitator as a function of quality advantage

Recall that if an imitator cannot reverse engineer the innovation perfectly, then $\Delta q > 0$. In this case the innovator's expected profits per deal despite competition are greater than $\frac{s}{4}$. As the Δq decreases we move on the above figure from right to left and the per deal expected profits of the innovator decrease while the per deal expected profits of the imitator increase. As Δq converges to zero, both profits converge to $\frac{s}{4}$ which is the a positive value, due to the client loyalty that allows above marginal cost pricing. As the quality advantage vanishes for uniformly distributed clients the probabilities of obtaining the deal converge to one half for both competitors.

Note that assumption (2) guarantees the expected profit of the developer in the first deal is bigger than in the following deals:

$$\pi_M^e = q_0 - \left(c + \frac{s}{2}\right) > q_0 - q_1 = \pi_0^e > 0 \tag{7}$$

Even if the innovator makes the upcoming deal, the presence of imitation brings downward pressure on prices and lowers the profit for the innovator.

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. In first place, the underwriter must learn how to choose the right parameters that are best for different issuers. In second place, 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. In third place, 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.⁸Thus, 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 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:

$$z_0 = a + \varepsilon_0,\tag{8}$$

where the noise component ε_0 is a normally distributed variable with:

$$E(\varepsilon_0) = 0, \tag{9}$$

$$Var(\varepsilon_0) = \tau^{-1}.$$
 (10)

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.

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:

$$\Delta q(t) = \tau \frac{1}{1 + \frac{\Sigma}{\tau}t}.$$
(11)

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:

$$\Delta q(t) = \Delta q(0) \left(1 + \frac{\Sigma}{\Delta q(0)} t \right)^{-1}.$$
(12)

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. From (6), 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 (12) 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:

$$t_{\frac{1}{2}} = \frac{\Delta q(0)}{\Sigma} \tag{13}$$

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.



Figure 5: The conversion ratio of Dividend Enhanced Convertible Stock (DECS) as a function of the returns of the underlying common stock.

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, second generation process of learningby-doing can start. The only difference with the framework presented above is that 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 5), are a second generation innovation derived from the PERCs, which have one degree of freedom less (Figures 7 and 8 illustrate generations following PERCS and DECS).

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:

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

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

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

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:

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

Define N as the first deal in which the imitator has a positive probability of obtaining it, that is given (12):

$$\begin{array}{ll} N & : & \Delta q \left(N - 1 \right) > s > \Delta q \left(N \right) \\ N & = & 1 + \operatorname{Int} \left[\frac{\Delta q(0)}{\Sigma} \left(\frac{\Delta q(0)}{s} - 1 \right) \right] \end{array}$$

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 (7) are:

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

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_0^9 . 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 $\frac{1}{\delta}$.¹⁰We conclude that later generation products last less, i.e., that new products arrive faster or after less deals of the previous 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:

$$f_{\alpha}(x) = \alpha x^{\alpha - 1} \qquad 0 \le x \le 1$$

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 innovator (located at 1) 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 client $\overline{x} = \min\left(1, \frac{1}{2} + \frac{\Delta q}{2s}\right)$ who is indifferent in equilibrium between both banks does not change, what changes is how many (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:

$$\begin{aligned} \pi_0^e &= \int_0^{\overline{x}} \left[(1-2x)s + \Delta q \right] \alpha x^{\alpha-1} dx = (s+\Delta q) \, \overline{x}^\alpha - 2s \frac{\alpha}{\alpha+1} \overline{x}^{\alpha+1} \\ \pi_0^e - \pi_1^e &= \Delta q + s \frac{1-\alpha}{1+\alpha} \end{aligned}$$

In the case of high product expertise advantage:

$$\Delta q > s \implies \overline{x} = 1$$

$$\begin{cases} \pi_0^e = \Delta q + s \frac{1-\alpha}{1+\alpha} \\ \pi_1^e = 0 \end{cases}$$
(14)

In the case of low advantage:

$$\Delta q < s \implies \overline{x} = \frac{1}{2} + \frac{\Delta q}{2s}$$

$$\begin{cases} \pi_0^e = \frac{2s}{1+\alpha} (\frac{1}{2} + \frac{\Delta q}{2s})^{1+\alpha} \in \left[\frac{2s}{(1+\alpha)2^{1+\alpha}}, \frac{2s}{1+\alpha}\right] \\ \pi_1^e = \frac{s}{1+\alpha} \left(2(\frac{1}{2} + \frac{\Delta q}{2s})^{1+\alpha} - (1-\alpha)\right) - \Delta q \in \left[0, \left(\frac{2s}{(1+\alpha)2^{1+\alpha}} - s\frac{1-\alpha}{1+\alpha}\right)\right] \end{cases}$$

$$(15)$$

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 that 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.

7 Empirical Evidence Related to the Model

7.1 Summary of Predictions

Now we address how the predictions of the model are consistent with the evidence found in the issues of corporate derivatives. Let us start by summarizing briefly the empirical implications of the model that can be tested with existing evidence and that we verify with our additional evidence:

- **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 will 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. The issues are done using 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 argued 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 found by Schroth, in his empirical study that uses the same database (Schroth, 2003). We also use Tufano's results as a benchmark to compare our predictions (Tufano, 1989).

Other prediction of the model was that later generation products should be shorter lived. Additionally, we will verify empirically if imitation is less frequent or not due to shorter living securities, or if early generation innovators are able to maintain their status as innovators or later generation products.

7.2 Product Groups and Sequences of Innovations

Equity-Linked securities were classified into product groups (or "families") and generations within these groups by Schroth (Schroth, 2003). Each one of the 51 different corporate derivatives in the SDC database is considered 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 defined as the lead underwriter of the first offer ever. Any other bank underwriting deals using the same security is called an imitator.¹¹

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 taxsaving 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

The studies by Tufano (Tufano, 1989) and Schroth (Schroth, 2003) have characterized the product market share dominance of corporate securities innovators. From Tufano's study, it is clear that the innovator's average share of all the deals done with any given innovation is larger than any of the imitators' average share. Thus, our first prediction is verified for all corporate securities innovations between 1971 and 1989. Using the equity-linked and corporate securities data, Schroth 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 market share advantage over time periods. Our second and

third predictions are also verified in the results of 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. Table 2 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 6 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.

Figure 6 compares the speed of imitation measured by the number of deals. 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 form a panel consists of all the deals from the second to the first imitation of each imitated security. The time elapsed between each deal and the first one is coupled with the time invariant covariates that we include in the following specification:

$$\lambda_i = \exp\{-(\beta_0 + \beta_1 * generation \ number_i + \beta_2 * size \ of \ first \ deal_i + \varepsilon\};$$
(16)

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*. The size of the first deal is used as a control for the expected size of the market at the time the security is introduced. The larger the expected market, the larger the incentives that imitator 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. We compute these estimates using three different distribution functions for the error term, ε : Weibull, Exponential and Log-normal.

The first three columns of Panel A in Table 3 show the estimates for the parameters in (16), omitting the size of the first deal as a control. The fit for the Weibull or the Exponential versions



Figure 6: Plot of Empirical Cumulative Density Functions for the Speed of Imitation Conditional on the Generation Number of the Product

of the model is very poor. While higher generation numbers are associated with larger hazard rates, and thus, faster expected time of imitation, these are not significantly different from zero. Moreover, the joint hypotheses that all parameters are zero cannot be rejected. The Lognormal version, though, fits the data much better. We can reject the null that all parameters are zero, and in this case the an increase by one generation increases the speed at which the security is imitated. Augmenting the specification with the size of the first deal does no change the inference made in the previous case. We see that the Weibull and Exponential models are a poor fit, while the lognormal fits the data much better (the Wald statistic has a p-value of 0.02). In all three columns the sign of β_1 is negative, but it is only significantly different from zero (with more than 95% confidence) in the lognormal case. Figure 9 illustrates the survival probabilities, i.e., the probability that a security is not yet imitated implied by the estimates of the lognormal model. Note also that the larger the first deal of every imitated security, the shorter the time it will take, on average, for imitators to enter the market. The Exponential model may not fit the data well because it assumes the hazard function is constant. The assumption of monotonicity in the hazard function implicit in the Weibull seems to be not borne in the data also: the implied increase of the speed of imitation along the earlier generations may be too high relative to what we observe.

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 size of the first deal: all the quartiles and the mean. We use the parameters in the last column of Panel A. In this case, the estimated time of entry of imitation is given by $\frac{1}{\lambda}$. At the median first deal size, the median imitation time is almost a year. The median imitation time decreases to just over six-months by the fifth generation. For the third quartile of the first deal size, the times are just over a half of these. For example, a tenth generation security's predicted median time of imitation is less than two months. This is also depicted in Figure 10.

Table 4 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).¹²In consistency with the previous results, the second row of Table 4 shows that we see that later generation products, if imitated, are imitated on average much faster than early generation products.

In Section 5.1, we argued 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. We argued that 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 4 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. However, one may ask why are some products imitated and why are others not. In fact, only 18 of the 51 innovations in this sample were imitated. Table 5 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. It shows that 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, as we saw earlier, 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 6 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.

As we have seen, group innovators may innovate or imitate later generations. In Table 7 we show that the average group market share of the group innovator decreases sharply when the group has more than one generation. In other words, in longer sequences of innovations, the group innovator does not seem to dominate its competitors as well as when in shorter ones. This is also confirmed by the negative, albeit weak, estimated correlation between the group innovator's share and the length of the product chain (the number of generations).

We look further into the group market shares by fitting a linear regression of a given bank's share of deals in a given product group on a dummy variable that takes value of 1 if the bank is the group innovator, the number of generations in the group and the total number of innovations by the bank in that group. Table 8 reports the least squares parameter estimates and the Huber/White estimates of their standard errors. All of these parameters are statistically different from zero with a significance level of 0.01. All other things constant, developing the first product in the chain is associated with capturing a bigger share of the overall number of deals in the group. The number of accumulated innovations matters and is positively associated with market shares. The interpretation would be that in a product group with a given number of generations a bank that has been always an imitator has, on average, a smaller share of the deals than had it been the innovator of some of the products. This fact is consistent with Tufano's evidence and the first prediction of our model, i.e., being the product innovator matters (Tufano, 1989).

We also find that on a group that is larger by one generation, the group innovator has, on average, a share that is almost 2 percentage points smaller. This is an important fact that a potential group innovator may consider strategically in his innovation decision: he will expect his market dominance to weaken if the product introduced is likely to be followed by many improvements.

Tables 9 and 10 provide some additional illustration of the patterns of innovation summarized so far. Tables 9 lists the innovators and imitators in the family of Perpetual, Deferrable-Income Securities. The group innovator is Goldman Sachs, who managed to introduce three more innovations of this type. However, Lehman Brothers also introduced three new products later in the sequence. Morgan Stanley and Smith Barney appear only as imitators in this group, and Merril Lynch, who imitates frequently early in the sequence, finally appears with one innovation in the last generation of the group. In the group of Index-Tied, Principal Appreciation Securities (Table 10), we see that Morgan Stanley is the group innovator but does not innovate more in the same group. In this group we do see some persistence in the leadership by Merril Lynch, with four consecutive innovations. Lehman Brothers innovates late after being an imitator earlier, and Goldman Sachs appears rather frequently as an imitator but never an 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 have argued 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 662 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 Patents, Welfare and Innovation

The analysis developed in this paper is relevant to the discussion of the effects of the new legal environment that favors patents for business methods and formulas in the US. Innovations in financial products fall in this category and, as Josh Lerner showed, there has been a dramatic increase since 1999 in the applications and awards for patents in the subclasses where financial innovations are filed (Lerner, 2000). Clearly, there is not yet enough data available to analyze the effects of the *State Street* case ex-post.¹³However, it is interesting to try to predict how patenting is going to change the benefits of investment banks and the overall pace of innovation in financial products. Below we compare the scenario described in this paper, where the innovator maintains an advantage over his imitators without patents, to a scenario in which the innovator is allowed to protect his innovation with a patent and eliminate imitation completely.

To address this comparison we propose the following parsimonious model of an innovation race. Let time be continuous and start at date 0, when banks must choose the hazard rate of the next innovation, i.e., the probability that they will discover the next innovation immediately after time t, given that no bank has innovated yet. Let there be two banks, i = A and B and call $h_i \ge 0$ the hazard rate for bank i.¹⁴ Assume that the cost of racing is an increasing and convex function of the hazard rate, i.e., for i = A and B,

$$e(h_i) > 0,$$

 $e'(h_i) > 0,$
 $e''(h_i) > 0.$

Note that this specification produces results identical to one where the banks have to choose how much to spend on R&D at time zero, and the hazard rate is a concave function of the amount spent. The common essential aspect of either setup is that the returns to R&D are diminishing.

Let t_i be the (random) time at which bank *i* discovers a new financial product. The net expected gain of bank *i* is thus

$$U_i = E\left(V\right) - e(h_i)$$

where

$$V = \begin{cases} \left(\overline{V} - v\right) e^{-rt_i} & \text{if } t_i = \min(t_A, t_B), \\ v e^{-rt_j} & \text{if } t_i = \max(t_A, t_B). \end{cases}$$
(17)

and r is the time discount factor. Note that $t_A = t_B$ is a zero probability event.

The form of the payoffs in (17) summarizes the central idea of this paper. Imitation cannot be foreclosed with a patent, yet the innovator has a first mover advantage and makes positive profits The imitators can share part of the profits from innovation though. Thus, the highest possible gains to one bank from innovation are \overline{V} , and if we let $v \in \left[0, \frac{\overline{V}}{2}\right)$ then the bank that discovers the new product first has the bigger gain always. The variable v denotes then the amount of profit sharing between the innovator (the first developer) and the imitator (the late developer). Better quality of imitation is denoted by a higher v and thus imitators get a larger share of \overline{V} . In the extreme case of perfect and immediate imitation, v would be close to $\frac{\overline{V}}{2}$ and we would have symmetric sharing of the profits. In the other extreme case of patent protection we would have v = 0 and all the profits are reaped by the innovator, i.e., we would have a winner-takes-all race.

Note that we have assumed here that the cumulative revenue of both firms is the constant $\overline{V}e^{-rt}$, regardless of how it is shared. However, the total revenue may change depending on the legal environment. With patents, the net aggregate revenue is reduced by the significant legal costs of patent filing and enforcement, whereas in the case of free imitation the aggregate revenue is reduced because competition makes issuers share part of the revenue through lower underwriting fees. The latter reduction should be smaller, though, if banks' clienteles are more loyal. Indeed, according to Battacharyya and Nanda the effects of client loyalty or switching costs are significant enough to make some non patentable innovations profitable in the banking business (Bhattacharyya and Nanda, 2000). The upcoming results will still hold if the cumulative net revenue of both banks does not change drastically from the case of a patent regime to the case of a free imitation regime.

We analyze now how the profits from innovation change as the amount of profit sharing, v, changes, assuming that banks compete to be the innovator. Banks A chooses h_A at time zero to maximize

$$U_{A} = \int_{0}^{\infty} e^{-rt} \left(\left(\overline{V} - v \right) h_{A} e^{-(h_{A} + h_{B})t} + v h_{B} e^{-(h_{A} + h_{B})t} \right) dt - e(h_{A})$$

= $\frac{h_{A}}{r + h_{A} + h_{B}} \left(\overline{V} - v \right) + \frac{h_{B}}{r + h_{A} + h_{B}} v - e(h_{A})$

taking h_B as given. Bank B solves a symmetric problem.

Proposition 4 There is a unique equilibrium to this innovation race. The equilibrium is symmetric, i.e., $h_A = h_B = h$. In the equilibrium, a higher degree of profit sharing implies a smaller rate of innovation, i.e., h is decreasing in v.

The uniqueness of the equilibrium allows us to do comparative statics. The closer we are to a

winner-takes-all situation in the race, i.e., closer to the case of racing for a patent, the banks will spend more effort or R&D resources in equilibrium. If the degree of profit-sharing, v, increases then banks will invest less and the discovery will happen later on average. Hence, the introduction of patents predicts faster and, thus more frequent, innovations. But the other important question is whether banks are better off with patents and more innovation than without patents. To answer this question, note first that for any given v the value to either bank in equilibrium is

$$E(U_i) = \frac{h}{r+2h}\overline{V} - e(h)$$

Hence, there is a unique value of h that maximizes the indirect utility of either bank in the race. Let this value be h^* . Since h is monotone in v, then there is also a unique level of profit-sharing that maximizes the expected profits of the racing banks net of R&D. However, we can show that a patent regime, which is the case where v = 0, does not implement the optimal speed of innovation in the race. In fact, we can show that the winner-takes all aspect of the race makes banks too eager to win the prize, so they overinvest in R&D, i.e., $h(0) > h^*$.

Proposition 5 A patent regime does not implement the optimal speed of innovation in the race where the returns to R&D are decreasing. Moreover, it makes banks overinvest.

The proof is in the appendix. Intuitively, patent protection increases the prize of winning and decreases the price of losing the race to the extreme where the winner takes it all, i.e., \overline{V} . Off the equilibrium, each banks tries to spend more than the other to try to win the race. In equilibrium, both banks spend too much. As v increases and the gains are shared to a larger extent, spending decreases in the unique equilibrium and we get closer to the optimum. If the amount of profit sharing, v, is too large however, we get to the other side of the peak and we have suboptimal investment in R&D because of a free-riding problem. If imitation profits are high enough, then off the equilibrium, each bank would rather spend less than the competitor. In equilibrium, both banks spend the same but less than the optimal amount.

To summarize, the optimal level of innovation from the point of view of competing banks is obtained if the bank that happens to innovate gets most but not all the profits from innovation. Some level of profit sharing is desirable to soften the competition to be the first one to develop the innovation. This may have been the case for the securities we have described in this paper, which could not be patented until recently. According to this analysis the recent introduction of patent protection and the winner-takes-all aspect it brings creates incentives to overspend in R&D. In fact, it is possible that the current situation may be further away from the optimal level of profit sharing than the previous situation without patents.

In this setup the incentives provided by patent protection are too extreme and implement a suboptimal investment in R&D. But certainly, patents have other effects of uncertain magnitude and direction. On the demand side, more innovation may benefit issuing firms by making more varieties of securities available to them. However, issuing fees faced would be higher due to the legal monopolistic. The net effect on welfare is uncertain and would deserve closer scrutiny. In our model, this effect is absent because there perfect price discrimination is possible: the valuation of the client for the type of security is known by all banks. Thus, the value that each borrowing firms derives from issuing any variety is equalized in equilibrium, so the speed of innovation only affects the banks.

9 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. Our work has concluded that when banks race for the next innovation, a certain degree of spill-over to the losing rivals is desirable, and *State Street*

may have given incentives to spend excessively in R&D costs, besides raising significantly the legal costs of patenting every new product often for mere defensive reasons. Wether *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.

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 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 expected profit per-deal for the next underwriting deal. 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 expected profit per-deal for the product 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 imitator's expected per-deal profits. *Ceteris paribus*, if the innovator and the imitator can offer the same quality underwriting then each one gets the same expected profit. This is the lowest profit that the innovator can get, and the highest expected profit for the imitator. If the quality differential is higher then the expected per-deal profits of the innovator increase (and the imitator's profits decrease). If the quality differential is high enough, the imitator cannot offer an attractive deal to any client and his profits are zero. At this point, the innovator's profit increases much faster as quality increases.

Figure 5

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. 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 6

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 7

This figure plots the conversion rate of an Automatically Convertible Equity Securities (ACES), as a function of the returns of the underlying common stock. Each unit of this preferred stock converts mandatorily after 4 years to one unit of common stock if the common stock appreciates between a floor \underline{r} and a cap \overline{r} percent. If the common stock does not appreciate enough in 4 years, then ACES convert to more than one unit of common such that their conversion value is that of a stock that appreciated \underline{r} percent. If the common stock appreciates more than \overline{r} percent in 4 years, then ACES convert to less than one unit of common such that their conversion value is that of a stock that appreciated \overline{r} percent.

Figure 8

This figure plots the conversion rate of a Participation Equity Preferred Stock (PEPS), as a function of the returns of the underlying common stock. Each unit of this preferred stock converts mandatorily after 4 years to one unit of common stock only if the common stock depreciates. If the common stock appreciates, but less than \overline{r} percent, then a unit of PEPS converts to less than one unit of common such that their conversion value is that of the common stock at the date of issue. The conversion rate, however, is floored.

Figure 9

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 $\hat{\lambda} = \exp(-6.296 + 0.140 * generation + 0.001 * size of first deal)$, and $\hat{\sigma} = 1.3124$. 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 10

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 $\frac{1}{\lambda}$, where $\widehat{\lambda}$ is the estimated imitation hazard rate, which is itself obtained from the estimated hazard rate model $\widehat{\lambda} = \exp(-6.296 + 0.140 * generation + 0.001 * size of first deal)$. 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.

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 pulished 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.

Schroth, E., (2003), "Innovation and First-Move Advantages in Corporate Underwriting: Evidence from Equity-Linked Securities," Working Paper, International Center for Financial Asset Management and Engineering (FAME).

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

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.

Footnotes

- 1. It 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.
- 2. Thorough 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).
- 3. In a recent survey, Peter Tufano argues that many mechanisms that reward innovation still remain to be studied (Tufano, 2003).
- 4. We 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.
- 5. Note 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.
- PERCs are shares of preferred stock that are mandatorily convertible to common stock after 3 years.
- 7. A 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.
- 8. The case of the Put Warrants introduced by Goldman, Sachs & Co. in 1990 illustrates these factors very well. Goldman pioneered the underwriting of US traded put options on foreign market indices. Goldman started engineering put warrants on the Nikkei index to capitalize on the desire by American investors to bet on the Nikkei's fall. Later, deals included puts on the French CAC-40. The deal also included a swap of the risk of exercise for a fixed payment between the option writer and the underwriter, who had to hedge this risk himself.

- 9. The 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} .
- 10. We 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.
- 11. Innovative corporate products were 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.
- 12. The 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.
- 13. It is not clear yet who is patenting most intensively, i.e., top investment banks, small firms, independent software developers, and what is being patented, i.e., security designs, pricing algorithms or trading platforms, after *State Street*. However, some top Wall Street firms have accumulated patents on tradeable securities lately. For example, Morgan Stanley patented a financial instrument of the Equity-Linked class recently. A debt instrument called "OPALS", which is an equity-linked obligation secured by an underlying basket of stocks of a chosen performance, and which provides tax advanatages to investors was awarded a patent in July 2000 (see "Data processing system and method for financial debt instruments," US Patent 6'092,056). Also, Goldman Sachs obtained a patent in 1999 for a marketable security that offers selective exposure to the real estate market by backing its payments with real estate accounts swapped from selected REITs (see "System for managing real estate SWAP accounts", US Patent 5'950,175).
- 14. Note that the model can be extended to allow for the case with a given larger number of banks in a straightforward way, without a change in the results.

15. Note too that having the banks spend R&D once and for all in time zero is equivalent as letting them choose a continuous flow of expenditures along the race.

Appendix

Lemma 1 The difference in quality after t deals is:

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

Proof. 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:

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

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.

Proof. 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 :

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

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

Proposition 4 There is a unique equilibrium to this innovation race. The equilibrium is symmetric, i.e., $h_A = h_B = h$. In the equilibrium, a higher degree of profit sharing implies a smaller rate of innovation, i.e., h is decreasing in v.

Proof. The first-order conditions for both banks are

$$\frac{r(\overline{V}-v) + h_B(\overline{V}-2v)}{(r+h_A+h_B)^2} = e'(h_A)$$
$$\frac{r(\overline{V}-v) + h_A(\overline{V}-2v)}{(r+h_A+h_B)^2} = e'(h_B)$$

Hence, defining the function H as

$$H(h_A) \equiv (r(\overline{V} - v) + h_A(\overline{V} - 2v)) e'(h_A),$$

$$H(h_B) = (r(\overline{V} - v) + h_B(\overline{V} - 2v)) e'(h_B).$$

>From the first-order conditions, $H(h_A) = H(h_B)$. Moreover, the function H is injective because $H'(h_A) > 0$. Thus, it must be that $h_A = h_B$, and the symmetric equilibrium is the unique candidate.

The values of $h_A = h_B = h$ must satisfy

$$r(\overline{V} - v) = e'(h)(r+2h)^2 - h(\overline{V} - 2v)$$

$$\overline{V}(r+h) - v(r+2h) = e'(h)(r+2h)^2$$

$$\overline{V} = \left(\frac{r+2h}{r+h}\right)[(r+2h)e'(h) + v]$$

The right hand side of the last expression is the product of two positive and increasing functions so it is increasing in h (namely, (fg)' = f'g + fg' > 0). Hence we have a unique value of h that satisfies the equality, and a unique equilibrium where $h_A = h_B$. The second order condition for either bank is always satisfied

$$-2\frac{r(\overline{V}-v) + h(\overline{V}-2v)}{(r+2h)^3} - e''(h) < 0.$$

To derive the comparative statics define the function

$$F(h(v),v) = \left(\frac{r+2h}{r+h}\right)\left((r+2h)e'(h)+v\right) - \overline{V} = 0$$

Finally, by the implicit function theorem we have

$$h'(v) = -\frac{F_v}{F_h} < 0.$$

Proposition 5 A patent regime does not implement the optimal speed of innovation in the race where the returns to R&D are decreasing.

Proof. The optimal rate of innovation h^* , is given by the best equilibrium value of h, i.e., the value that maximizes the indirect utility

$$E(U_i) = \frac{h}{r+2h}\overline{V} - e(h) \,.$$

Thus, h^* equates the first derivative of the previous expression to zero. It solves therefore

$$\overline{V} = \frac{\left(r+2h^*\right)^2}{r} e'(h^*).$$

But for a given v, the equilibrium value of h equates to zero the derivative of

$$E(U_i) = \frac{h_i}{r + h_A + h_B} (\overline{V} - v) + \frac{h_{j \neq i}}{r + h_A + h_B} v - e(h_i), \ i, j = A, B$$

with respect to h_i and imposes the symmetry condition. Thus, it solves

$$\overline{V} = \frac{(r+2h)^2}{r+h}e'(h) + v\frac{r+2h}{r+h},$$

which, for v = 0 becomes

$$\overline{V} = \frac{(r+2h(0))^2}{r+h(0)} e'(h(0)).$$

Clearly $\overline{V} = \frac{(r+2h(0))^2}{r+h(0)} e'(h(0)) < \frac{(r+2h(0))^2}{r} e'(h(0))$. Since $\overline{V} = \frac{(r+2h^*)^2}{r} e'(h^*)$, then $\frac{(r+2h^*)^2}{r} e'(h^*) < \frac{(r+2h(0))^2}{r} e'(h(0)).$

Moreover, both sides of the inequality are increasing in h, so $h^* < h(0)$.





- Mandatory Conversion in 4 years.
- High dividend yield (>7%).

Figure 7: The conversion ratio of Automatically Convertible Securities (ACES) as a function of the returns of the underlying common stock.



Figure 8: The conversion ratio of Participation Equity Preferres Stock (PEPS) as a function of the returns of the underlying common stock.



Figure 9: 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 10: 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.

Product Group	Number of Products	Number of	Number of	
	(i.e., Generations)	Distinct Innovators	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	

 Table 1: Number of Innovators and Competing Underwriters in All Product Groups of Equity

 -Linked and Derivative Securities

Source: Schroth (2003).

 Table 2: Summary of the Empirical 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	Positive and Significant
	- innovator dummy interacting with time	Negative and Significant
	- innovator dummy interaction with time	
	and generation number	Negative and Significant

Table 3: 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.

	Weibull	Exponential	Lognormal	Weibull	Exponential	Lognormal
Constant	6.374	6.467	5.822	6.735	6.776	6.296
	$(0.243)^{***}$	$(0.298)^{***}$	$(0.192)^{***}$	$(0.385)^{***}$	$(0.001)^{**}$	$(0.433)^{***}$
Generation Number	-0.097	-0.082	-0.130	-0.882	-0.0648	-0.140
	(0.108)	(0.129)	$(0.047)^{***}$	$(0.131)^{***}$	(0.155)	$(0.061)^{**}$
Size of First Deal				-0.001	-0.001	-0.001
(US\$ Millions)				(0.001)**	$(0.001)^{**}$	$(0.000)^{**}$
No. of Imitated Securities	18	18	18	18	18	18
No. of Observations	48	48	48	48	48	48
χ^2 Wald statistic	0.81	0.40	7.57***	4.18	3.88	7.88**

Hazard Rate Model: $-\ln \lambda = \beta_0 + \beta_1 * generation + \beta_2 * size \ of \ first \ deal + \varepsilon.$

Panel B

Hazard Rate: $\lambda = \exp(-6.296 + 0.140 * generation + 0.001 * size of first deal)$

Median time of First Imitative Deal: $M = \frac{1}{\lambda}$						
Generation	Median Time, evaluated at size of first deal $=$					
Number	1st quartile Median Mean					
1	415.90	341.36	278.77	193.36		
5	237.60	195.01	159.26	110.47		
10	118.01	96.86	79.10	54.87		
15	58.61	48.11	39.29	27.25		

Each observation in Panel A consists of the time in days after innovation, paired with the indicator of whether the deal is by the innovator or not, the generation number and the size of the first deal. 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 005 level, and by *.to the 0.1 level.

		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	First Generations	7	4.86	2.48	4
(deal number of first imitation)	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

 Table 4: Summary Statistics of the Speed, Duration and Innovator's Market Shares for the Innovative Equity-Linked Securities

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.

	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-v	value: 0.026	

Table 5: Distribution of Imitated and Non-Imitated Products, Conditional on their Generation

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.

Number of Innovations in the Product Group						
Type of Banks	0	1	2	3	4	Tota
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

Table 6: Distribution of the Number of Innovations by Banks Competing in Each Product Group

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.

Table 1. Summary Statistics for the Share of Dears 0	nuci wi iuch Dy	une or	oup mnovator
	Observations	Mean	Std. Deviation
1. In groups with one generation only	5	0.74	0.36
2. In groups with more than one generation	6	0.41	0.23
Correlation with the Number of Generations in the Group	11	-0.624	

Table 7: Summary Statistics for the Share of Deals Underwritten by the Group Innovator

The Group Market Share by the Group Innovator is the number of deals underwritten by the developer of the first generation product, divided by the total number of deals in the Product Group.

underwritten deals of all securities in a given group.				
Regressors	Parameter estimates			
Underwriter is Group Innovator	0.333***			
(Yes = 1)	(0.096)			
Number of Generations in Group	-0.019^{***}			
	(0.005)			
Total number of innovations by Underwriter	0.058***			
in that Group	(0.022)			
Intercept	0.219***			
	(0.052)			
Observations	61			
F statistic	17.41***			
R-squared	0.59			

Table 8: Bank's Group Market Shares, Length of Sequence and Innovative Activity

The dependent variable is the share of a banks's

The market shares are calculated using all the offerings of equity-linked and derivative securities between 1985 and 2001. Each observation consists of a bank's share of deals in the group, paired with the identity of the banker as the innovator of the group (Yes or No), the number of innovations to follow in the group (generations) and the number of times the bank innovated in the same group. The parameters are estimated using Ordinary Least Squares and their standard errors are obtained by computed the Huber-White hesteroskedasticity-consistent covariance matrix estimator. Estimates followed by *** are significant to the 0.01 level, by ** to the 0.05 level, and by * to the 0.1 level.

Innovations in Perpetual, Deferrable-Income, Securities									
Generation	1	2	3	4	5	6	7	8	9
Security	MIPS	EPICS	MIDS	TOPRS	QIDS	QUIPS	QUICS	RST-CAPS	COPRS
$Innovator^a$	GS	BEAR	GS	LEH	GS	GS	LEH	LEH	ML
Imitators	ML		ML	SAL	ML			ML	
	SB		MS	SB	MS			SB	
								UBS	

 Table 9: List of Innovating and Imitating Banks in the Class of Perpetual, Deferrable-Income

 Securities

^a The bank acronyms above are: GS = Goldman Sachs, ML = Merrill Lynch, SB = Smith Barney,
BEAR = Bear Stearns, MS = Morgan Stanley, LEH = Lehman Brothers, SAL = Salomon Brothers,
UBS = UBS Securities.

The data is found within the issues of equity-linked and derivative corporate securities between 1985 and 2002.

Index-Tied Principal Appreciation Securities								
Generation	1	2	3	4	5	6	7	8
Security	PERLS	SIRS	SMARTS	MITTS	EPS	CPNS	SUNS	CUBS
Innovator ^a	MS	UNK	ML	ML	ML	ML	LEH	BEAR
Imitators	GS	\mathbf{PW}		GS	GS			
	OP			LEH				

 Table 10: List of Innovating and Imitating Banks in the Class of Index-Tied Principal Appreciation

 Securities

 a The bank acronyms are: GS = Goldman Sachs, ML = Merrill Lynch, PW = Paine Webber, BEAR = Bear Stearns, MS = Morgan Stanley, LEH = Lehman Brothers, OP = Operaciones,

UNK = unknown.

The data is found within the issues of equity-linked and derivative corporate securities between 1985 and 2002.

The FAME Research Paper Series

The International Center for Financial Asset Management and Engineering (FAME) is a private foundation created in 1996 on the initiative of 21 leading partners of the finance and technology community, together with three Universities of the Lake Geneva Region (Switzerland). FAME is about **Research**, **Doctoral Training**, and **Executive Education** with "interfacing" activities such as the FAME lectures, the Research Day/Annual Meeting, and the Research Paper Series.

The **FAME Research Paper Series** includes three types of contributions: First, it reports on the research carried out at FAME by students and research fellows; second, it includes research work contributed by Swiss academics and practitioners interested **in** a wider dissemination of their ideas, in practitioners' circles in particular; finally, prominent international contributions of particular interest to our constituency are included on a regular basis. Papers with strong practical implications are preceded by an Executive Summary, explaining in non-technical terms the question asked, discussing its relevance and outlining the answer provided.

Martin Hoesli is acting Head of the Research Paper Series. Please email any comments or queries to the following address: Martin.Hoesli@hec.unige.ch.

The following is a list of the 10 most recent FAME Research Papers. For a complete list, please visit our website at www.fame.ch under the heading 'Faculty and Research, Research Paper Series, Complete List'.

- N° 123 A Double-Sided Multiunit Combinatorial Auction for Substitutes: Theory and Algorithms Henry SCHELLHORN, HEC-University of Lausanne and FAME, *December 2004*
- N° 122 Investment Under Uncertainty and Incomplete Markets Julien HUGONNIER, HEC-University of Lausanne and FAME and Erwan MORELLEC, HEC-University of Lausanne and FAME, *May 2004*
- N° 121 On the Debt Capacity of Growth Options Michael BARCLAY, William E. Simon School of Business Administration; Erwan MORELLEC, HEC-University of Lausanne and FAME and Clifford W. SMITH, Jr., William E. Simon School of Business Administration, January 2003
- **N° 120 Omega Portfolio Construction with Johnson Distributions** Alexander PASSOW, Gottex and FAME, *November 2004*
- N° 119 A Simple Alternative House Price Index Method Steven C. BOURASSA, School of Urban and Public Affairs; University of Louisville, Martin HOESLI, HEC, University of Geneva and University of Aberdeen, Business School and Jian SUN, School of Urban and Public Affairs, University of Louisville, November 2004
- N° 118 Real Asset Returns and Components of Inflation: A Structural VAR Analysis Matthias HAGMANN, HEC-University of Lausanne and FAME and Carlos LENZ, University of Basel, Department of Economics; October 2004
- N° 117 Equity Returns and Integration: Is Europe Changing? Kpate ADJAOUTE, HSBC Private Bank (Suisse) and FAME and Jean-Pierre DANTHINE, HEC-University of Lausanne, CPER & FAME; October 2004
- N° 116 The Price Impact and Survival of Irrational Traders Winner of the 2004 FAME Research Prize Leonid KOGAN, Sloan School of Management, MIT and NBER; Stephen ROSS, Sloan School of Management, MIT and NBER; Jiang WANG, Sloan School of Management, MIT, CCFR and NBER; Mark WESTERFIELD, the Economics Department, MIT; *August 2004*
- N° 115 Predicting Tail-related Risk Measures: The Consequences of Using GARCH Filters for Non-GARCH Data Amine JALAL, HEC - University of Lausanne and FAME and Michael ROCKINGER, University of Lausanne, FAME AND CEPR; June 2004

N° 114 Further Evidence on Debt-Equity Choice

Philippe GAUD, HEC-University of Geneva; Martin HOESLI, HEC-University of Geneva, FAME and University of Aberdeen (School of Business) and André BENDER, HEC-University of Geneva and FAME; *May 2004*







International Center FAME - Partner Institutions

The University of Geneva

The University of Geneva, originally known as the Academy of Geneva, was founded in 1559 by Jean Calvin and Theodore de Beze. In 1873, The Academy of Geneva became the University of Geneva with the creation of a medical school. The Faculty of Economic and Social Sciences was created in 1915. The university is now composed of seven faculties of science; medicine; arts; law; economic and social sciences; psychology; education, and theology. It also includes a school of translation and interpretation; an institute of architecture; seven interdisciplinary centers and six associated institutes.

More than 13'000 students, the majority being foreigners, are enrolled in the various programs from the licence to high-level doctorates. A staff of more than 2'500 persons (professors, lecturers and assistants) is dedicated to the transmission and advancement of scientific knowledge through teaching as well as fundamental and applied research. The University of Geneva has been able to preserve the ancient European tradition of an academic community located in the heart of the city. This favors not only interaction between students, but also their integration in the population and in their participation of the particularly rich artistic and cultural life. *http://www.unige.ch*

The University of Lausanne

Founded as an academy in 1537, the University of Lausanne (UNIL) is a modern institution of higher education and advanced research. Together with the neighboring Federal Polytechnic Institute of Lausanne, it comprises vast facilities and extends its influence beyond the city and the canton into regional, national, and international spheres.

Lausanne is a comprehensive university composed of seven Schools and Faculties: religious studies; law; arts; social and political sciences; business; science and medicine. With its 9'000 students, it is a mediumsized institution able to foster contact between students and professors as well as to encourage interdisciplinary work. The five humanities faculties and the science faculty are situated on the shores of Lake Leman in the Dorigny plains, a magnificent area of forest and fields that may have inspired the landscape depicted in Brueghel the Elder's masterpiece, the Harvesters. The institutes and various centers of the School of Medicine are grouped around the hospitals in the center of Lausanne. The Institute of Biochemistry is located in Epalinges, in the northern hills overlooking the city. *http://www.unil.ch*

The Graduate Institute of International Studies

The Graduate Institute of International Studies is a teaching and research institution devoted to the study of international relations at the graduate level. It was founded in 1927 by Professor William Rappard to contribute through scholarships to the experience of international co-operation which the establishment of the League of Nations in Geneva represented at that time. The Institute is a self-governing foundation closely connected with, but independent of, the University of Geneva.

The Institute attempts to be both international and pluridisciplinary. The subjects in its curriculum, the composition of its teaching staff and the diversity of origin of its student body, confer upon it its international character. Professors teaching at the Institute come from all regions of the world, and the approximately 650 students arrive from some 60 different countries. Its international character is further emphasized by the use of both English and French as working languages. Its pluralistic approach - which draws upon the methods of economics, history, law, and political science - reflects its aim to provide a broad approach and in-depth understanding of international relations in general. *http://heiwww.unige.ch*

UNIVERSITÉ DE GENÈVE







40, Bd. du Pont d'Arve PO Box, 1211 Geneva 4 Switzerland Tel (++4122) 312 09 61 Fax (++4122) 312 10 26 http: //www.fame.ch E-mail: admin@fame.ch



