Innovation, Differentiation, and the Choice of an Underwriter: Evidence from Equity-Linked Securities

Enrique Schroth
HEC-University of Lausanne and FAME

Investment banks imitate other bank's innovative corporate securities and compete with the innovator to underwrite new issues. This article uses data of all the corporate offerings of equity-linked and derivative securities in the Securities Data Company (SDC) to estimate the issuer's demand of underwriting services provided by investment banks across different varieties of securities. It finds that the demand for the innovator's variety is larger than the imitators'. This demand advantage decreases with time and faster for securities that appear later in a sequence of innovations. Imitation becomes less attractive later in the sequence as information from earlier deals spills-over to all banks.

The last 20 years have witnessed the introduction of a remarkable number of innovations in corporate securities.[1] Most of these have been brought about by investment banks through the business of underwriting new corporate issues. It is also remarkable that investment banks have found it profitable to develop new securities even when their competitors have been able to imitate them almost immediately and at significantly smaller development costs. The empirical evidence so far has suggested that, despite these disadvantages, innovators are compensated with the largest share of the underwriting market. In this article, I estimate the demand of firms for the underwriting services of investment banks that use innovative corporate products. This estimation allows the measurement of the difference in value to firms from raising money using a security engineered by an innovator or an imitator and thus explains part of the innovator's market share advantage. The dynamic setup of the econometric model will allow a characterization of this advantage

I am grateful to the members of my thesis committee, Douglas Gale, Boyan Jovanovic, and Marti Subrahmanyam. I also acknowledge helpful comments by Franklin Allen, Chris Flinn, Helios Herrera, Eugenio Miravete, Franco Peracchi, Michael Rockinger, Peter Tufano, the Referee and the Editor, Maureen O'Hara. Any errors in the article are mine. I also thank Ingo Walter, Iftekhar Hasan, and the Berkley Center for Entrepreneurial Studies who made possible my access to the data. Address Correspondence to Enrique Schroth, Route de Chavannes 33, 1007 Lausanne, Switzerland, e-mail: enrique.schroth@unil.ch.

[1] For a list and a description of innovations in corporate securities until 1991, see Finnerty (1992). All innovations in equity-linked and derivative corporate securities until March 2001 are listed here.

doi:10.1093/rfs/hhj023 Advance Access publication February 17, 2006

over time and shed light on the question of what makes innovators have a demand advantage over imitators.

Product innovation in finance is particularly interesting, because innovators face many disadvantages. Tufano (1989) provided evidence showing that imitation occurs shortly after the first issue of a new security, leaving virtually no time for innovators to be the sole underwriters for that new product. He also observed that the development cost is significantly smaller for imitators than for innovators. Furthermore, the design of new securities is rapidly revealed to competing banks because of the Securities and Exchange Commission (SEC) rules of disclosure. Most strikingly, imitation cannot be foreclosed legally with patents.

In his seminal contribution, Tufano (1989) observed that innovators are compensated on average with the largest share of the market for underwriting corporate new issues: the bank that develops a new security is able to underwrite more capital than its largest imitator over all the security's history. Why innovators are able to have an unchallenged lead in these markets is perhaps the first question that this evidence raises. This article takes a first step toward answering the question empirically. For that purpose, I model the choice of a firm that needs to raise money externally through the issue of a security. This firm has to choose the type of security to be used and the bank that will underwrite it. Aggregating the choices of all the firms that need to make an offering, the model predicts the market shares of underwriting by different banks using different securities, conditional, among other things, on the characteristics of the banks (e.g., if they are imitators or innovators). With estimates of this model, we can test if innovators have an advantage, because their security is more valuable to the issuers.

This article introduces two features that allow a better description of the facts of financial innovation. One is the use of a framework of differentiated products to model and estimate the demand for underwriting services. An inspection of recent innovations in corporate products clearly suggests that different securities are created to target different types of issuers or investors. For example, two similar debt products that tie the repayment of the principal to the performance of other indices provide different hedging devices to investors: the Stock Market Annual Reset Term Notes (SMARTs) are corporate bonds that pay a capped floating rate that is tied to the American Stock Exchange Oil Stocks Index, whereas the Currency Protected Notes (CPNs) are bonds that pay a floating rate that is inversely tied to the Canadian six-month bankers acceptance rate. By taking into account the location of securities in a product space, it is possible to identify consistently a demand function for underwriting. This is possible because we can associate the variation in market shares with the variation in its underwriting prices (i.e., spreads) of varieties of the same security by different banks and the variation in underwriting spreads of similar varieties which can be close substitutes.[2]

The other feature is that this study focuses also on the dynamic pattern of market shares. Instead of comparing the market shares of innovators and imitators over the whole product history, I observe them over time and estimate the innovator's demand advantage accordingly. This reveals whether the innovator's advantage is preserved steadily through the life cycle of the security or whether imitators catch up with innovators. This dynamic setup also allows a comparison between the demands along sequences of securities. Most securities appear sequentially, and the life a security ends shortly after the next innovation is introduced.

The empirical finance literature has not yet addressed extensively the question of why innovators have advantages over imitators. Most authors have examined extensively the *causes* of the demand for innovative securities. The focus has been on trying to explain what made each particular innovation attractive to investors or issuers but not on why it is privately profitable to develop such instruments. Miller (1986), for example, argued that the major impulse to financial innovations between the sixties and the eighties came from incentives to circumvent ever-changing tax codes and regulations and pay less taxes (Tufano, 1995; Finnerty, 1992).

Not much work has been done to try to solve the puzzle of why an unpatentable innovation is worth its R&D expenditure if imitation is less costly and immediate. In particular, not much has been said about what gives innovators an advantage over imitators. Some authors have tried to explain Tufano's stylized facts at the theoretical level by arguing that innovators may face lower search costs of identifying potential issuers and investors (Allen and Gale, 1994, chap. 4) or lower average marketing costs if there are fixed costs to setup marketing networks (Ross, 1989). Battacharyya and Nanda (2000) argued that innovators can appropriate a larger value of their innovations despite being imitated if it is costly for their clients to switch to other underwriters. In contrast with these views, I analyze the possibility that the asymmetry between innovators and imitators is on the information they have about structuring the new security for their clients. If innovators can keep private some information about the new security, then they can appropriate a larger proportion of the value added. In other words, the innovator can profit, because he is imitated imperfectly.

This possibility was explored in a theoretical paper by Herrera and Schroth (2000): innovators of derivatives that move one period in advance receive private informative signals about their clients or the market. This allows them to offer more attractive deals than imitators.

[2] Tufano's study compared spreads only between banks issuing the same security, so it is not a demand estimation for the underwriting services of banks across different securities.

For recent innovations in corporate securities, it is possible that imitation is imperfect. Equity-linked securities and other derivative corporate products are sophisticated securities, specified by many parameters, some of which vary from deal to deal. Thus, it is possible that imitators cannot reverse-engineer perfectly the observed new securities from only a few deals. For example, the Equity-Linked Note (ELK) was the first debt product to tie the repayment of principal to the stock price of another publicly traded company. The optimal choice of which stock to tie the notes to is observable, but the knowledge of what stocks are optimal for different issuers or investors is a private component of R&D. Imitators who want to underwrite ELKs issues for potential clients may know the structure of such instruments, but may not know exactly, from what is observable, what stock to tie to the repayment of client's debt.

Using data on all the new issues of equity-linked and corporate derivative securities, the (qualitative) results show that (1) the demand for underwriting services using this type of securities is sensitive to the underwriting spread (i.e., its price); (2) on average, the issuers' demand for the underwriting service by an innovator of a given security is bigger than the demand for an underwriter that uses an imitative variety; (3) the difference in demand disappears gradually during the security's life cycle; and (4) imitators catch up with innovators faster if securities are later improvements on past innovations. Thus, this article provides the first empirical test that the advantage to innovators may come from a bigger demand for the innovators' products. Although the empirical work directly shows how the innovator's advantage evolves over time, the particular dynamic patterns found point strongly to the fact that innovators have superior knowledge on the product's engineering and hence provide more value to the issuers. In fact, the dynamic patterns of competition it identifies are consistent with the predictions of VanHorne (1985) when imitators enter the market.[3]

This article also explores why and when are securities imitated, if so. I propose a simple model where imitation is increasingly profitable over generations (because the demand for imitator's underwriting services converges faster to the innovator's) and increases the chances of innovating later but also allows outsiders to learn from observing the imitators complete deals. If imitation spills-over some information to nonimitators, then imitation can become less frequent along the sequence of new securities. This view is supported with evidence from regressions of the probability that a given security is imitated on the generation number of the security.

In the next section of the article, I describe the data set and present some preliminary results that motivate the assumptions of the demand model to estimate. In Section 2, I characterize the decision of issuing firms to choose

[3] Kanemasu, Litzenberger, and Rolfo (1986) observed the same pattern for the case of stripped treasury securities.

different types of underwriters (innovators and imitators). I show that firms should derive additional value if they chose an innovator who is more informed about deal structuring. I explain how this conjecture is tested with nested logit models of demand. In Sections 3 and 4, I present the estimation results and test their robustness to different definitions of what can be considered as innovative or as imitative securities, respectively. In Section 5, I derive the rule that rival banks follow to decide to imitate a new security or not, which depends on the security's generation and how much imitation there has been in earlier generations. The predictions are tested using probabilistic regressions. In Section 6, I summarize the main claims of the article.

1. Some Preliminary Evidence

This research uses data from the Securities Data Company's on-line databases of financial transactions. I use all the private and public offerings of equity-linked and derivative corporate securities in the New Issues database and record all available details of the offer: the issuer's name, the proceeds, the underwriter's name, the underwriting spread, and details of the issue like offered yield to maturity, average life, spread of coupon over treasury notes, call options, and so on. I merge this data set with the quarterly COMPUSTAT database using six digit CUSIP numbers to have financial information about the issuer.

Tufano (1989) used all types of securities between 1974 and 1986. Here, I restrict the sample to equity-linked securities, because this type fits better the motivation that first-mover advantages may come from information asymmetries between underwriters. These securities are complex, and underwriters have to choose many parameters to engineer each deal (Table 1 summarizes the relative size of this market). Variations on mortgage-backed products, for example, may be already too familiar to investment banks to hide something in their structure to imitators.

The first thing to realize about equity-linked and derivative securities is that they can be classified into groups. The SDC database identifies 50 different types of them but a closer look indicates that some of these securities are related to each other in terms of their structure and purpose. For example, MIPS and TOPRS are instruments used by issuers to deduct debt-interest payments from their taxable income, but the former are issued by a limited liability company, whereas the latter by a specially conformed business trust (Pratt, 1995). To classify all the 50 corporate products in the SDC database, I have relied on the experts' opinion about the uses of these products for corporations and investors.[4]

[4] For every product, I have compiled articles in Investment Dealers' Digest, American Banker, Dow Jones Newswires and others found using the ABI Search Engine. For every one, I was able to find a description of the product and a reference to an older product that was similar to it. Tom Pratt wrote descriptive article of almost every corporate security invented in the Investment Dealers' Digest.

Table 1
Total principal underwritten in the new issues markets using standard equity and equity-linked products between April 1985 and March 2001

	Total underwritten principal		Average deal size	
Period	Standard equity	Equity-linked and derivative products	Standard equity	Equity-linked and derivative products
1	64,477	3915	38	489
2	61,034	590	36	84
3	35,801	4120	48	589
4	45,123	9314	46	388
5	61,216	19,517	57	887
6	99,507	10,750	57	336
7	129,073	22,106	56	287
8	90,397	22,414	56	270
9	139,518	32,357	61	225
10	169,569	35,463	83	269
11	143,016	19,616	112	265
12	230,734	20,680	381	440

Figures shown are expressed in millions of current US$. The source of the data is the Securities Data Company Database of New Issues. Each one of the 12 periods is an equal-sized interval between April 1995 and March 2001.

I found 10 distinct categories summarized in Appendix B. I refer to these categories as “groups” or “families” of securities interchangeably hereafter.

I will refer to each security as an innovation because for each one there is a unique feature that distinguishes it from the already existing securities. I assign to each security a generation number depending on the order in which they appeared in their group. For example, MIPS were the first tax-advantage preferred note, so I call them the first generation of this family. I call TOPRS the fourth. Like Tufano (1989), I define the innovator of a security as the underwriter of its first offer. For equity-linked corporate securities, it is also the case that innovators lead imitators in terms of market shares. Eighteen of the 50 products are imitated. Innovators lead in principal underwritten in 13 of these and lead in number of new issues in 15. Imitation here is fast too: for 10 of these securities, the second underwriting deal was made by an imitator (Table 2).

I want to test here whether issuers have stronger preferences for innovators than imitators as underwriters of their offer. To choose an appropriate framework to study the demand for the underwriter, it is worth examining the dynamics of the market for a new issue within each security and within each family. In Figures 1 and 2, I show the total quarterly principal underwritten by investment banks using the most important securities of two families. Each security seems to be popular for a certain period until a next generation appears and leads the market for issues in

Table 2
Innovators' market shares for imitated securities

Product	Innovator	Number of rivals	Share of deals (%)	Leader	Share of principal (%)	Leader
RISRS	Kemper securities	1	43	No	50	Yes (tied)
SIRS	Paine Webber	1	67	Yes	90	Yes
DECS	Salomon-Smith Barney	1	94	Yes	79	Yes
Equity participation securities	Merrill Lynch	1	94	Yes	50	Yes (tied)
MIDS	Goldman, Sachs	2	29	No (2)	30	No (2)
SMART notes	Merrill Lynch	2	60	Yes	41	Yes
X-CAPs	Merrill Lynch	2	67	Yes	92	Yes
ELKS	Salomon-Smith Barney	2	82	Yes	90	Yes
PERCS	Morgan Stanley	2	82	Yes	88	Yes
QIDS	Goldman, Sachs	2	82	Yes	76	Yes
PERLS	Morgan Stanley	2	93	Yes	97	Yes
LYONS	Merrill Lynch	2	97	Yes	95	Yes
FRAPS	Merrill Lynch	3	22	No (2)	23	No (2)
RST-CAPS	Lehman Brothers	3	25	Yes (tied)	22	No (2)
MIPS	Goldman, Sachs	3	89	Yes	96	Yes
TOPRS	Merrill Lynch	3	95	Yes	92	Yes
TRUPS	Salomon Brothers	4	71	Yes	72	Yes
Convertible TOPRS	Merrill Lynch	6	47	Yes	33	No (2)
Trust-originated convertible preferred	Robertson	7	6	No (5)	5	No (4)
Total number of leads (of 18 cases)			15		14	

An entry of "Yes" in the fifth or last columns (Leader) means that the innovator had the biggest market share. A "No" entry is followed by the innovator's ranking in brackets. A tie for the lead is also shown in brackets.

the group. Also, for most families, the market-share advantage of innovators over imitators seems to be bigger in the early generations. On average, the average ratio of principal underwritten by imitators relative to principal underwritten by innovators is 0.34 for first generation securities only and 0.68 for all imitated securities (see the supplement to this article for a more detailed picture of the innovator's advantage over different product groups).

The evidence above suggests that innovators may have an advantage over imitators that is stronger for earlier generations. Further inspection suggests that the appropriate framework to analyze the securities in this sample may be one of the differentiated products: it is clear from the definition of the 50 securities that these corporate products offer different benefits to issuers or investors. Some provide a tax advantage, and others provide a hedge against the risk of defaulting on debt. Interestingly, within imitated securities, there seems to be differentiation across underwriters. For example, the within-variety variation for some of these characteristics is smaller for imitators than for the innovators (see supplement). The price of

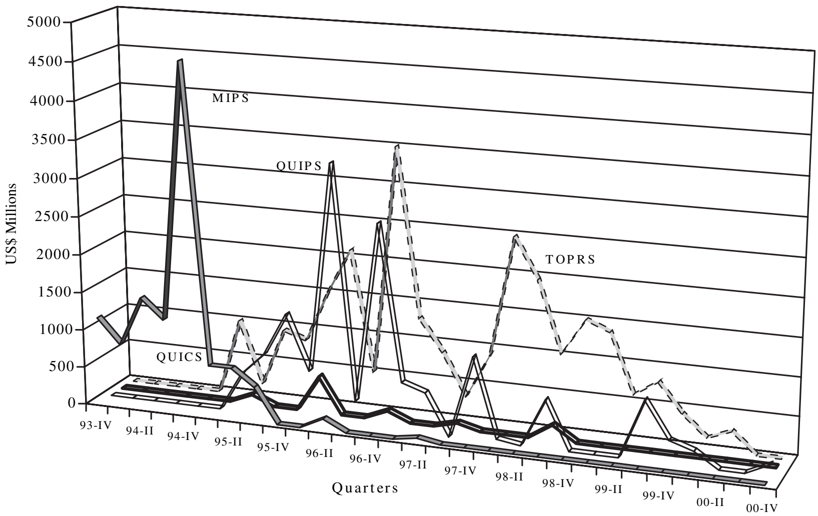

Figure 1
Principal underwritten using tax-deductible preferred products
This figure shows the plots of the total principal underwritten using different types of perpetual preferred stock whose dividend payments are used by the issuer as deductions to taxable earnings. Underwritten totals are aggregated on quarterly bases and expressed in millions of current US dollars. MIPS are cumulative guaranteed monthly income preferred. QUIPS are quarterly version of MIPS, and QUICS are debt convertible to preferred at the investor's option. TOPRS pay cumulative monthly cash distributions and are issued by trusts, not limited liability corporations. MIPS are the family's first generation. Differentiated versions of MIPS appear later as next generations, and their use as corporate finance instruments also peaks later.

underwriting, that is, the underwriting spread, does not seem to differ significantly, although the within-security sample mean of the underwriting fee for innovators is larger than for imitators (Table 3).

I will interpret this evidence as I present the econometric model and as I discuss the estimates. What I conclude from this evidence is that a model of differentiated products in an oligopolistic industry is suitable to describe competition for underwriting deals between investment banks using equity-linked and corporate securities. Differentiation occurs at the underwriters level, where innovators are distinguished from their imitators. Henceforth, I call a distinct combination of a security and an underwriter a *variety*.

2. The Demand Model

2.1 The econometric approach

Below I illustrate the decision of firms that want to raise capital and have to choose an appropriate security and the best underwriter. The underwriter or "book-manager" will engineer the security and sell it. The choice of a bank is derived from the issuers' preferences for the banker's quality

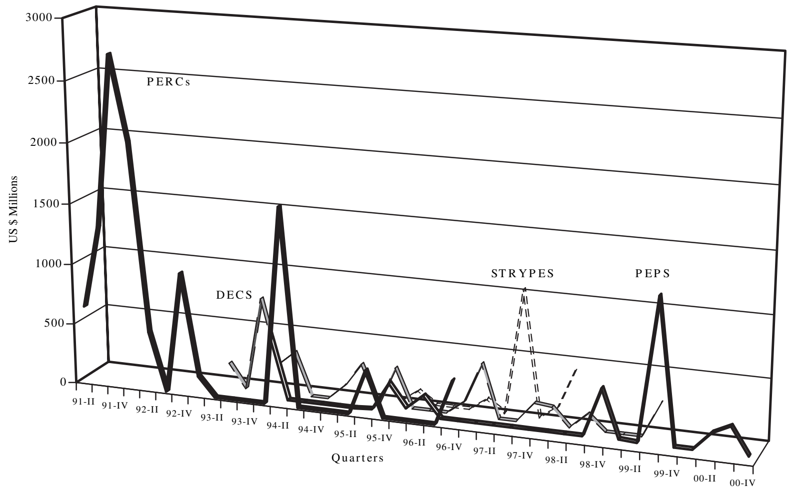

Figure 2
Principal underwritten using convertible preferred products
This figure shows the plots of the total principal underwritten using different types of mandatorily convertible preferred stock. Underwritten totals are aggregated on quarterly bases and expressed in millions of current US dollars. PERCs are premium-dividend preferred stock convertible in three years, callable anytime before conversion. For DECS, the issuer gives up the call option before conversion. STRYPES are a later generation of this family and may also convert to cash. PEPS are four-year convertibles that are also putable. Differentiated versions of PERCS appear later as next generations, and their usage peaks later.

Table 3
Univariate tests and summary statistics

Panel A: Paired difference of means test for the underwriting spreads of innovators and imitators

$\triangle$ = Underwriting spread by innovator – underwriting spread by imitator

Periods of aggregation	Observations	Mean	Standard error	p-value ($\Delta > 0$)
16	137	0.058	0.0537	0.139
12	113	0.063	0.0573	0.137
11	110	0.051	0.0577	0.189
8	79	0.083	0.0706	0.121

Panel B: Data summary statistics

Variable	Mean	Standard deviation	Median	Minimum	Maximum
Offer					
Principal (US$ millions)	306.160	446.515	175.50	10.00	4050.00
Underwriting spread (%)	2.442	1.135	3.00	0.085	12.60
Private placement (yes = 1)	0.138	0.345	0	0	1
Days before first imitative deal	484.440	642.875	214	7	2408
Issuer information					
Total capitalization (US$ millions)	15,668.700	44,750.980	4115.70	1.20	742,536.60
Common stock (% of capitalization)	43.436	22.513	44.97	0.27	100

Table 3
(continued)
Panel B: Data summary statistics

Variable	Mean	Standard deviation	Median	Minimum	Maximum
Preferred stock (% of capitalization)	7.574	11.501	5.29	0.01	99.31
Long-term debt (% of capitalization)	43.522	19.546	42.98	0	100
Short-term debt (% of capitalization)	15.991	16.551	10.14	0.01	52.99
Bank's market share (share of total deals, by period)					
Variety's share of all equity underwriting	0.0008	0.0019	0	0	0.0148
Variety's share of the product market	0.6857	0.3787	0.9621	0.0256	1
Standard equity market share	0.9535	0.0149	0.9454	0.9395	0.9959
Other variety indicators (by period)					
Innovations accumulated by bank in category	1.172	1.293	1	0	4
Innovations accumulated by bank (total)	5.023	5.198	3	0	14
Product innovator dummy (yes = 1)	0.585	0.493	1	0	1
Group innovator dummy (yes = 1)	0.328	0.470	0	0	1
Years or protection from call	3.968	2.645	3.963	0	30
Yield to maturity (%)	6.809	5.827	7.292	0	73.094
Yield advantage (%)	2.151	3.424	0	0	31
Instruments (average over other varieties or other banks in same product market)					
Innovations accumulated by bank in category	0.503	0.743	0	0	4
Innovations accumulated by bank (total)	2.566	3.435	1	0	14
Years or protection from call	1.395	3.477	0	0	19.583
Yield to maturity (%)	2.772	4.903	0	0	24.133
Yield advantage (%)	0.618	2.0888	0	0	11.998

In panel A, the means of the underwriting spreads for innovators and imitators at a given period for a given security are compared using a paired difference of means test. Four different frequencies are used to define a period: one year (t = 16), 17 months (t = 12), 18 months (t = 11) and 2 years (t = 8). The alternative hypothesis to compute the p-value is that the difference of means is positive. Data are for 662 issues of equity-linked and derivative securities between April 1985 and March 2001. Market shares and variety-specific data statistics are computed after the aggregation required by the estimation period.

as an underwriter. If the innovating bank has superior information over other banks on how to structure deals, then this builds into a larger demand for his variety. Thus, the identity of the underwriter summarizes the quality of the engineering choices he makes for a given deal, so the estimation of a demand function that depends on the bank's characteristics identifies the bank's superior knowledge underwriting that security.[5]

[5] See Herrera and Schroth (2000), Black and Silber (1986), or Allen and Gale (1994) for models where the advantage of an innovator over its imitators is developed endogenously as a result of moving first.

2.1.1 The setup. Banks compete by offering differentiated products: debt or equity types of deals that offer investors different payoff schedules, horizons, call options, convertibility possibilities, and so on. The issuer is indexed by $i \in \mathcal{I}$, and it chooses a financing instrument from a set of varieties $\mathcal{J} = \{1, 2, \ldots, J\}$. Let g be an index for groups of varieties, $g = 1, \ldots, G$, that describes a partition $\mathcal{G} = \{\mathcal{J}_1, \mathcal{J}_2, \ldots, \mathcal{J}_G\}$ of $\mathcal{J}$, so that each set $\mathcal{J}_g$ contains those varieties that are closer to each other in terms of their characteristics. In this setup, for example, if consumers were choosing models and brands of a car, then a set $\mathcal{J}_g$ would contain all brands of, say, SUVs, and some other set, $\mathcal{J}_h$, would contain all brands of compact cars. In our case, a group in $\mathcal{G}$ contains all securities with the same name, for example, PERCS, LYONS, or TOPRS, and each variety is given by the name of the bank that underwrote the issue, for example, PERCS by Morgan Stanley or PERCS by Merrill Lynch.

Let b represent an underwriter in the set of banks, $\mathcal{B}$. The firm chooses one security and one underwriter among those that offer that security. A variety $j \in \mathcal{J}$ is thus a unique (b, g) pair, and u_{ij} is the value to firm i of choosing j. The setup for this decision is illustrated in Figure 3. The value function u_{ij} is central to this article, because it is the function whose parameters I will estimate empirically. I will characterize this function when I derive the preferences of issuers for underwriters who are innovators and underwriters who are imitators. The empirical literature that deals with the estimation of preference parameters in models of discrete choice uses special cases of the general specifications of linear preferences by Caplin and Nalebuff (1991) or Anderson, DePalma, and Thisse (1989). In these studies, agents value a product according to a weighted sum of its

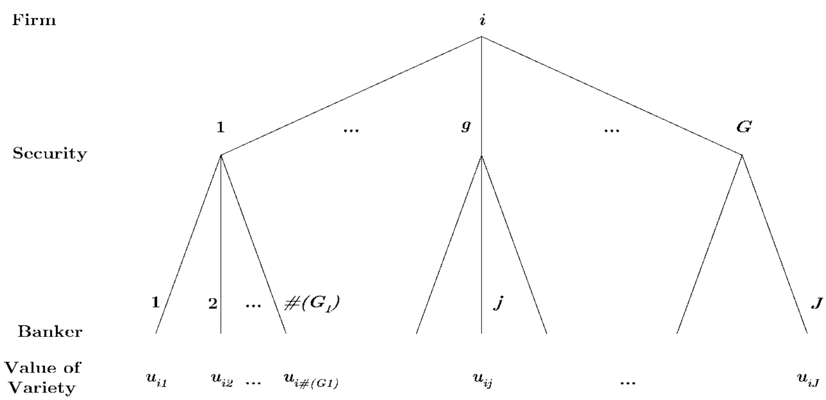

Figure 3
The decision tree of an issuer of a corporate security
This figure illustrates the decision tree faced by an issuer of a corporate security, indexed by i. It chooses one variety, $j \in \{1, 2, \ldots, J\}$, which is valued as u_{ij}. A variety is determined by a combination of a security type, $g \in \{1, \ldots, G\}$ and a banker within the type.

components. These components, in general, are functions of the observable characteristics of the product. We shall see below how this structure is particularly appropriate for this study.

2.1.2 The value of a deal. An underwriting deal is defined fully by a combination of a vector of characteristics, $\boldsymbol{\theta}$, an underwriting fee, p, a variety, j, and an issuer, i. The vector $\boldsymbol{\theta}$ includes, for example, the premium over dividend paid by common stock, the date of maturity, the number of periods of a call protection, or, more broadly, discrete variables that determine whether the security is convertible or not, convertible to common stock or debt, and so on. A general random value function for the ith issuer that chooses some variety j is denoted by

$$u_{ij} = u(y_i - p_j, \boldsymbol{\theta}_j) + \varepsilon_{ij}, \tag{1}$$

where y_i measures the proceeds of i, and ε_{ij} is assumed to be an additive random component that captures the random preferences of a given firm for a particular variety. It is unobservable to underwriters and distributed independently with a continuously differentiable cumulative distribution function, $G(\varepsilon)$.[6]

Underwriters choose a price and an engineering vector $\boldsymbol{\theta}_j$ for their variety that maximizes the likelihood of getting each given deal. One can then eliminate $\boldsymbol{\theta}_j$ from the issuers' value function, by substituting the *optimal* choice of $\boldsymbol{\theta}_j$. Therefore, banks can also be differentiated if they pick $\boldsymbol{\theta}_j$ from different sets. In fact, an issuers' indirect utility can then be written as a function of the bank's identity.

Financial innovation and imitation can then be adapted to this setup. If an underwriter spends resources on R&D to come up with a new security, then it will discover new possible combinations of characteristics a firm may find valuable, possibly changing zero entries in $\boldsymbol{\theta}$, to add new dimensions to the structure of a security. The PERC, for example, was the first issue of preferred stock convertible to common stock with capped and floored appreciation.[7] The first underwriting deal made by innovators using a new security reveals the new security design and a particular engineering vector. The whole set of possible engineering choices for this particular security is kept private. If innovators can pick deal structure from larger sets than generic imitators, then the value derived by a given issuer should

[6] McFadden (1981) showed that this function is continuous and twice continuously differentiable in $y - p$, and $\boldsymbol{\theta}$ provided that the function is continuous and twice continuously differentiable in other goods and in $\boldsymbol{\theta}$.

[7] Clearly, the set of new possible engineering choices is tied to a discovery of firms' preferences for some yet inexistent feature of securities. This was certainly the case of the PERCs, where its innovator, Morgan Stanley, figured out that in the uncertain environment for investment in late 1991, an issue of a PERC would allow firms to attract investors with stable high yields, pricing the offer better than common stock while capping the security's appreciation potential (Pratt, 1991).

be larger if it chooses the innovator as its banker. Formally, we can rewrite the issuers' value, maximized in $\boldsymbol{\theta}$, as $v(y_i - p, b)$. If we denote the identity of an innovating investment bank by $b = 1$ and any imitator by $b = 0$, then for any couple of varieties j, l of the same security g, then

$$v_j(y_i - p, 1) \geq v_l(y_i - p, 0).$$

A formal proof is provided in the supplement to this article.

2.1.3 The demand function. The innovator's advantage can be measured by the additional value to issuers if they were to choose an innovators' variety. Call the value added $\Delta v_j \equiv v(y_i - p, 1) - v(y_i - p, 0)$.

Working with a value function, where $\boldsymbol{\theta}$ has been eliminated, is convenient because what I want to capture is exactly the difference in preferences for the different underwriters. Also, it avoids losing a large proportion of observations for which the full $\boldsymbol{\theta}$ is not available. In other words, my interest is to distinguish preferences for these two types of banks rather than to estimate the marginal valuation (and the derived elasticities) for a given characteristic of a security, for example, years of call protection, yield advantage.

Another reason is that preferences for the choices of each attribute in $\boldsymbol{\theta}$ may be complicated functions that make the estimation difficult. Thus, using a simpler function that summarizes all the attributes seems reasonable. This approach has been used by Caplin and Nalebuff (1991), who use a utility function that is linear in generic functions of all the product characteristics. Using their own example, the value of a car is linear in comfort and speed, but these may be more complicated functions of the physical attributes.

Let $q_j(\mathbf{p}, .)$ be the market demand for variety j.

Proposition 1. *If $\Delta v_j \geq 0$, then for two varieties j,k of the same security g and for a fixed vector of underwriting spreads, $\mathbf{p} \in \mathbb{R}_+^J$, then*

$$q_j(\mathbf{p}, .) \geq q_k(\mathbf{p}, .)$$

if j is the innovator's variety. If $\Delta v_j > 0$, then $q_j(\mathbf{p}, .) > q_k(\mathbf{p}, .)$.

I prove this proposition in the appendix.[8]

2.1.4 The demand advantage through time. Proposition 1 has established that, if part of the full set of engineering choices for a given security is

[8] Note that, in equilibrium, innovators need not have the biggest within-securities market shares even if the demand for their variety is the largest. Under mild regularity conditions, that is, $v(.)$ is not too concave, then the game is one of the strategic complementarity. Moreover, if marginal underwriting costs are similar across banks, then the innovator's equilibrium market shares and prices are increasing in his first-mover advantage. The supplement to this article formalizes these conditions.

kept private, then the demand for his variety will be larger than the demand for any imitator's variety. Clearly, though, as the innovator completes more deals, more possibilities for engineering are disclosed to imitators. With time, the set from which the imitator picks $\boldsymbol{\theta}$ becomes larger and eventually gets arbitrarily close to the innovator's set. The imitator may learn the innovator's set element by element through observation, or it may reverse-engineer this set much faster using also some prior knowledge of the issuer's preferences. Anyway, it is clear that with time the imitator's choice set gets larger.

Lemma 1. *If Δv_j is decreasing in time, then the demand for the imitator's variety converges monotonically to the demand for the innovator's variety.*

As a corollary, if next generation innovations are marginally decreasing improvements of the innovators, then convergence should be faster for later generations. Namely, if the initial innovator's advantage Δv_j gets smaller as new products are introduced, we would expect the advantage to be shorter lived, which would show up empirically by a demand function for the imitative varieties that, *ceteris paribus*, converges more rapidly to the innovator's.

2.2 The estimable model

Below I present the equation that is obtained from aggregating the issuers' demands for the different varieties and how it is estimated.

2.2.1 The general setup. I consider each time period $t = 1, \ldots, T$ a different market in which an issuer i chooses its desired variety j. As the standard of the empirical literature of discrete models of demand, I will specify the value of this issuer as a function of observed and unobserved characteristics of the issuer and of the product offered by the underwriter and of the relevant parameters.

Let

$$u_{ijt} = \alpha(y_{it} - p_{jt}) + \mathbf{x}_{jt}\boldsymbol{\beta} + \boldsymbol{\xi}_{jt} + \varepsilon_{ijt}. \tag{2}$$

The first term on the right is the net proceeds of the issue, which are assumed to be separable and linear.[9]

The second term captures the indicators that distinguish an innovator from an imitator. The vector $\mathbf{x}_{jt}$ contains Δv_j, which itself includes all the

[9] The quasi-linear form rules out wealth effects. This assumption is tractable and convenient for the estimation but also is reasonable for this particular data set: the amount paid in underwriting fees is small relative to the value of the outstanding equity of issuers (the 95th percentile of the ratio of fees to equity is 0.02) or to their yearly earnings (the median proportion of fees of the yearly earnings is 0.06; the third quartile is 0.21). Nevo (2000), for example, used a quasi-linear form for the demand of ready-to-eat cereal, because its price represents a small share of household expenditures, but Berry, Levinsohn, and Pakes (1995) used a nonlinear net income term, $\alpha \log(y_{it} - p_{jt})$ to model the demand for cars.

variables that reflect quality differences between underwriters over time and over generation within groups. It also contains measures of banks' acquired reputation based on their history as innovators or successful underwriters within the family or the whole equity-linked class.

The value of choosing alternative j to raise capital can be decomposed into its mean, $\delta_{jt} = -\alpha p_{jt} + \mathbf{x}_{jt}\boldsymbol{\beta} + \boldsymbol{\xi}_{jt}$ and a deviation from it, $\alpha y_{it} + \varepsilon_{ijt}$. The unobservable (to the econometrician) characteristics of the security j itself are captured by ξ_j, whereas ε_{ijt} accounts for the heterogeneity of firms' preferences.

2.2.2 The nested logit demand. Berry (1994) showed that we can identify the parameters of (2) using the observed market shares of all varieties if we assume that ε_{ijt} is drawn independently and identically from a zero-mean Weibull distribution, that is, with density $f(\varepsilon) = \exp[-\exp(-\varepsilon)]$. The data are generated starting from decision to choose a variety, which is represented by a tree as in Figure 3. Issuer's valuations are allowed to be correlated within groups so that different varieties of the same securities offered by different investment banks would be closer substitutes of each other than any other security. Therefore, the utility of a given choice, j, can be modeled as a restricted version of (2), allowing for random coefficients, ζ_{ig}, on security-specific dummies:

$$u_{ijt} = \delta_{jt} + \sum_{g} d_{jg}\zeta_{ig} + (1-\sigma)\varepsilon_{ijt}, \tag{3}$$

where $d_{jg} = 1$ if $j \in \mathcal{J}_g$.

If the random coefficients are also assumed to be drawn from a Weibull distribution, then so is $\zeta + (1-\sigma)\varepsilon$. The parameter σ measures the within-security correlation. If σ approaches one, then so does the within-security correlation of value, and if it approaches zero, then there is no correlation.[10] The overall and within-group market shares predicted by the model, $\widehat{s}_{jt}$ and $\widehat{s}_{j/gt}$, are obtained by integrating all the realizations of unobservables that lead to a choice of j over all other varieties. The closed form solution is

[10] For $\sigma = 0$, we obtain the multinomial logit model. In this case, the odds ratios of choosing one variety over another do not depend on the value of *other* varieties. This can be problematic: if the demand for MIPS were evenly split between Goldman, Sachs, and Merrill Lynch, and each were half the demand for Salomon Brother's ELKS, then the ratio of the market share of Goldman's MIPS to Salomon's ELKS would still be one half, even if Merrill increases its underwriting fee by any magnitude. Implicitly, the business lost by Merrill Lynch would be absorbed by both Goldman and Salomon so as to preserve the ratio, even if Morgan Stanley and Merrill Lynch offer close varieties and Salomon offers a different product. With $\sigma > 0$, we have more reasonable substitution patterns. Cross-price elasticities for two different varieties within the same group will be different than cross-price elasticities for varieties across different groups, even if they have the same prices and market shares. The restricted versions of the model were estimated too and are included in the supplement to this article.

$$\widehat{s}_{j/gt}(\boldsymbol{\delta}_t, \sigma) = \frac{e^{\delta_{jt}/1-\sigma}}{D_{gt}}, \tag{4}$$

where

$$D_{gt} = \sum_{h \in \mathcal{J}_g} e^{\delta_{ht}/1-\sigma}, \tag{5}$$

and

$$\widehat{s}_{jt}(\boldsymbol{\delta}_t, \sigma) = \frac{e^{\delta_{jt}/1-\sigma}}{D_{gt}^{\sigma} \sum_{f \in \mathcal{G}} D_{ft}^{1-\sigma}}. \tag{6}$$

Normalizing the average utility of the outside alternative to zero ($\delta_0 = 0$) and noting that the term αy_{it} drops out because it is common to all the choices, then (6) can be written as:

$$\ln s_{jt} - \ln s_{0t} = \sigma \ln s_{j/gt} - \alpha p_{jt} + \mathbf{x}_{jt}\boldsymbol{\beta} + \xi_{jt}. \tag{7}$$

Hence, $\alpha, \boldsymbol{\beta}$, and σ can be estimated with an Instrumental Variables estimator from the regression of the log-difference of the market share of j and the outside alternative on $\mathbf{x}_{jt}$, p_{jt}, and $s_{j/gt}$.[11]

The challenge of estimating consistently the parameters of (7) is in finding appropriate instruments for the price because it is surely correlated with the unobservables, ξ_{jt}. This is a typical problem of using product characteristics as regressors. Here, $\mathbf{x}_{jt}$ uses issuers' characteristics that summarize the full description of the product, so it is less likely that ξ_{jt} still unobservable correlated with the price. The within-security market share, $s_{j/gt}$, is also endogenous, so we must use additional instruments.

2.2.3 Issuers heterogeneity. To enrich the substitution patterns in the demand model, I incorporate data from COMPUSTAT about all the firms that did new issues using equity-linked security in the sample. This introduces heterogeneity that makes the cross-price elasticity depend

[11] In some periods, market shares of existing varieties are zero. The estimation of the nested logit requires taking the logarithm of these shares. To avoid the indeterminacy problem, I compute

$$\ln s^*_{jt} = \ln(s_{jt} + 0.00001),$$
$$\ln s^*_{j/g} = \ln(s_{j/gt} + 0.00001),$$

which does not alter the distribution of s_{jt} significantly. Excluding these observations would not only bias the sample selection but also imply a loss of 191 observations that actually reveal that the demand was zero.

on the issuers' characteristics and relaxes significantly the distributional assumptions on ε_{ijt}. I consider a vector of f firm characteristics $\mathbf{f}_t$, each one interacted with the price to obtain the specification:

$$\ln s_{jt} - \ln s_{0t} = \sigma \ln s_{j/gt} - \alpha p_{jt} + \mathbf{x}_{jt}\boldsymbol{\beta} + p_{jt}\mathbf{f}_t\boldsymbol{\phi} + \xi_{jt}. \tag{8}$$

The cost of this specification is a loss of a significant proportion of the observations in the sample, because many issuers of equity-linked securities have no COMPUSTAT entry.

2.3 The data

The SDC Database of New Issues records all the public and some private offerings made since 1962. There are 662 issues until March 2001 defined in SDC as equity-linked or derivative corporate securities (the first issue, a LYON, was made in April of 1985). There are 50 securities and a total of 98 varieties. I compute the varieties' market shares over the whole market of new issues and over the varieties within the security at different time periods.

I divide the whole sample in time periods rather than aggregate the data by varieties over the whole time span. Overall aggregation would reduce significantly the number of observations (to 98) and would also eliminate the time variation of market shares and underwriting fees, compromising seriously the consistency of the estimators. I treat each period as an independent market, so that there is a demand function for each variety at each time. The parameters of this function are identified by cross-sectional variation in prices, in the identity of the underwriter, and the issuer's characteristics and by the time variation in prices and issuers' "experience." The panel structure of the data is crucial to study the dynamics of the advantage to innovators. To form the panel, I must choose the length of each period. The number of usable observations is larger for shorter chosen period lengths, but this increases the risk of aggregating very few deals per period, increasing dramatically the variation in the market shares. To avoid arbitrariness in the choice of length, I do the estimations at four different levels of aggregation: using 16 periods (annually), 8 periods (biannually), 11 periods (18 months), and 12 periods (16 months). Thus, we can also assess the robustness of the results to this choice.

The panel is unbalanced because not all securities are offered at each period. Only two varieties are offered in the first period and 98 in the last. I consider standard equity as the outside option to issuers, that is, the variety $j = 0$. The total size of the market for new issues in t is approximated with

$$M_t = q_{0t} + q_{1t} + \ldots + q_{Jt}, \tag{9}$$

and demand, q, is measured in number of deals, not dollars. This assumes that firms set *ex ante* the amount of cash they want to raise in the offer, and the choice modeled here is of the security and the underwriter.

2.3.1 Variables. Market shares. Overall market shares are $s_j = q_{jt}/M_t$, whereas within-security market shares are $s_{j/gt} = q_{jt}/\sum_{k\in\mathcal{J}_g} q_{kt}$.[12]

Prices. Prices of underwriting are the fees charged by the investment bank that leads the syndicate of book managers of an offer. They are expressed as a percentage of the principal underwritten and called underwriting spreads. Usually this spread can be disaggregated in the underwriting fees and management fees, but this disaggregation is seldom observable, so the price variable I will use is the total spread.[13]

Demand shifters in $\mathbf{x}_{jt}$. The demand shifters that do not capture the information asymmetry between underwriters are variables about the underwriter's experience and reputation issuing this type of securities. I use the total capital underwritten in the equity-linked class, the total number of innovations in equity-linked products, and innovations within the particular family of the security accumulated by the underwriter. I use period dummies to control for observable economy-wide shocks and group dummy variables.

Advantage to innovators. To test whether innovators have advantages on the demand side, I included a dummy variable that equals one for innovator's varieties. A positive estimate of the coefficient of this variable implies that, on average, firms have stronger preferences for innovators.

If the advantage diminishes as more deals are completed by imitators, then we would expect a negative estimate of the coefficient of the innovator dummy interacted with time after the security was imitated. Moreover, if the security is a late generation of a given group, then more information about this type of securities would have been aggregated, and we would expect imitators to learn the innovators' private information faster. Thus, I also interact this dummy with the generation number to get a richer characterization of the dynamics of learning by doing. Formally, to model these dynamics, I specify Δv_j as:

$$\Delta v_{jt} = \gamma_0 i_j + \gamma_1 i_j \times gen_j + \gamma_2 i_j \times gen_j \times \widetilde{t}, \tag{10}$$

where the dummy variable $i_j = 1$ if j is the innovator's variety, gen is the generation number of the security, and $\tilde{t}$ is the number of periods since imitation.

[12] For periods of zero market shares, no information of the bids made by potential underwriters is observed. Because these varieties were available to firms, although none chose to use them, we assume there is a going price for them. We approximate this price with the last observed price for that variety.

[13] The covariance matrix estimator is robust to possible correlation of errors within the same security. While errors of varieties of the same bank may also be correlated, this may not be significant if they are in different product groups. Also part of this effect should be picked up by the bank-specific regressors. The inference made on the innovator's advantage index or the bank's characteristics parameters using an estimator of the covariance matrix that considers within-bank correlation rather than within-security does not change at all. The estimate for α becomes insignificant when $t = 12$, and the estimate for σ becomes insignificant when $t = 16$.

Issuers data. I use financial data from COMPUSTATs quarterly database that matches the period of the offer. I use the total market capitalization to measure the size of the firm. I also use indicators of common equity, preferred equity, short-term debt, long-term debt, and subordinated debt all expressed as percentages of capitalization.

2.3.2 Instruments. I need to use instruments to obtain consistent estimates of the parameters of the model, because prices and within-securities market shares are endogenous. To choose appropriate instruments, I follow the suggestions of Berry, Levinsohn, and Pakes (1995) and Berry (1994). Instruments for the underwriting spread (price) include the averages of characteristics of the security over the *competing* varieties, like years of call protection, years before call at par, percentage yield, which should not be correlated with the error term because the advantage term summarizes all characteristics of the security. By the same, token instruments for the within-security market shares include characteristics of *other* underwriters in the same group (e.g., total and within family accumulated innovations by the other underwriters of the same security). To test if these instruments over-identify the parameters of the models, I perform a Hausman test.

3. Results

3.1 Benchmark model

Table 4 reports the estimates of the parameters of (7) using an instrument for the underwriting spread and the within-security market shares (the standard errors were estimated using the Huber–White variance estimators, allowing for heteroskedasticity and serial correlation within securities). I use period-specific dummies and fixed effects for the security group. This model was fitted for the four different aggregations of data: 8, 11, 12, and 16 periods.

The estimated coefficient of price has the expected negative sign for all the aggregations, that is, the demand for a bank's underwriting service of a given variety slopes downward. The estimate is significant at the 95% level for $t = 16$ and $t = 12$ and at the 90% level for $t = 11$ but insignificant for $t = 8$ (note that I report $-\alpha$, to be consistent with notation above). If, as assumed, the data of the underwriting markets were generated by a model of oligopoly, then the elasticity of demand at the fees observed should be bigger than 1 as, in equilibrium, banks should price their deals in the inelastic portion of the curve. Table 4 also reports the number of estimated demand curves that are elastic at the observed fee and violate this condition. In all columns, the average elasticity is well above 1, and the number of violations is always relatively small (never more than 3.6%).

The estimated coefficient for the substitution parameter is significant in all cases, and it is within the appropriate bounds of 0 and

Table 4
IV estimates for the base nested logit demand model

Regressors (parameter name)	Parameter estimates			
	$t = 16$	$t = 12$	$t = 11$	$t = 8$
Underwriting spread (α)	−1.359** (0.594)	−1.139** (0.524)	−1.049* (0.542)	−1.794 (1.183)
Product market share, in logarithms (σ)	0.371* (0.228)	0.489** (0.243)	0.494** (0.228)	0.584** (0.289)
Total principal underwritten in the equity-linked class by the underwriter (US$ M)	0.079*** (0.016)	0.078*** (0.019)	0.079*** (0.019)	0.078*** (0.025)
Innovator dummy[1]	2.428** (1.053)	2.218** (1.153)	1.889* (1.149)	2.780* (1.714)
Innovator dummy interacted with the security's generation number	0.151*** (0.037)	0.220*** (0.055)	0.154*** (0.039)	0.170*** (0.049)
Innovator dummy interacted with imitation time and generation numbers	−0.191*** (0.071)	−0.276*** (0.102)	−0.269** (0.114)	−0.506** (0.231)
Group dummies[2]	Yes	Yes	Yes	Yes
Time dummies[3]	Yes	Yes	Yes	Yes
Instrumented regressors: underwriting spread and product market shares (in logarithms)				
χ^2 statistic for OIR test[4]	5.138	12.968	10.324	4.168
p-value	0.526	0.073	0.109	0.526
Average elasticity of market share to the underwriting spread	5.522 (2.360)	5.569 (2.420)	5.230 (2.261)	10.646 (4.842)
Median elasticity of market share to the underwriting spread	6.131	6.125	5.849	12.181
Number of predicted inelastic demands	15	10	11	4
Number of observations	418	323	312	225
R^2	0.246	0.308	0.308	0.127
F-statistic	12.95***	14.95***	15.35***	9.75***

The dependent variable is the difference of the logarithms of the market shares of a banks' underwriting of a given security and of the market share of equity underwriting: $\ln(s_j) - \ln(s_o)$. The sample includes all offerings of equity-linked and derivative securities between 1985 and 2001. The parameters of a theoretical logit demand model are obtained using an Instrumental Variables estimator. The standard errors are obtained using a Huber–White covariance estimator, robust for heteroskedasticity, and within-security correlation. They are shown below the estimate, in parenthesis. Each observation consists of the market share of a banks' variety of a given security at a given period, associated with the characteristics of the security and the banker, in the same period, and their instruments. The sample time span has been split in 16, 12, 11, and 8 periods.

[1]Equals one when the underwriter is the innovator of that security.

[2]A dummy for each one of the 10 categories. Equals one when the security belongs to that category.

[3]A dummy for each one of the 12 periods. Equals one when the observation corresponds to that period.

[4]Test of over-identified restrictions. The null hypothesis is that the model is correctly specified and that the instruments over-identify the variation in the endogenous regressor. The p-value is shown below the χ^2 statistic.

Estimates followed by ***, **, and * are significant to the 0.01 level, 0.05 level, and 0.1 level, respectively.

1: $0.371 < \hat{\sigma} < 0.584$. This result is consistent with a setup where varieties within a security type are closer substitutes than varieties of other types. Issuers switch bankers before switching security structures.[14]

The innovator dummy has a positive coefficient, significant to the 95% level of confidence when $t = 16$ and $t = 12$. The coefficient is significant only to the 90% level for the other two aggregations. This evidence suggests that issuers have stronger preferences for the innovator's variety, *ceteris paribus*. The parameters associated to the innovator dummy interacted with the generation number and to the innovator dummy interacted with the generation number and time after imitation are all significant to the 99% level of confidence. The estimate of the coefficient of $i \times gen$ is positive, revealing that the later the generation, the higher the average *initial* advantage of the innovator. The other interaction, $i \times gen \times \tilde{t}$, has a negative estimated coefficient, showing that this advantage decreases with the number of periods that imitators have been in the market but *faster* for each next generation.

Figures 4 and 5 illustrate better what these estimates mean in terms of the time during which issuers value innovator's varieties more than imitators'. I plot the estimated advantage of the innovator, that is, the predicted ratio of the innovator's market share to the imitator's (s_{In}/s_{Im}), in the vertical axis against time, measured in years, in the horizontal axis. In all cases, the innovator's advantage decreases monotonically, and it disappears slowly for first generations: on average, in 12 years. However, an innovator of, say, a seventh generation has a larger initial advantage, but it is gone shortly after the product's second year after having been imitated (which happens when both market shares are predicted to be equal, *ceteris paribus*). One possible interpretation is that late generations are often very complicated modifications of existing securities. At first, it may be difficult for imitators to learn how innovators are engineering the deals, but in time they should learn faster given that more information has been aggregated about the security type or about the family of securities.

The benchmark specification includes as a regressor the cumulative principal by the bank using equity-linked securities until the time of the observation. The estimates for the associated parameter are strongly significant and positive for all aggregations, suggesting that issuers derive more value from hiring a bank with a larger placement ability in the

---

[14] Alternatively, one can think of using well-known underwriter reputation measures, such as those found in Loughram and Ritter (2004), which update the original prestige rankings by Carter and Manaster (1990) and Carter, Dark, and Singh (1998) to the year 2000. These measures base reputation on the number of times that a bank appears higher in the list of syndicated underwriters in the prospectus list. The reputation ranking from 1980 to 2000 is of little use to this study, because it exhibits very little time and cross-sectional variation for most of the banks in this sample. Of 662 deals, 650 are underwritten by banks that scored 8.75 or 9.0 between 1985 and 1991 and 8 or 9 between 1992 and 2000. Thus, a measure that distinguishes these banks in terms of their placement ability is a measure related directly to the placement capacity in this product class.

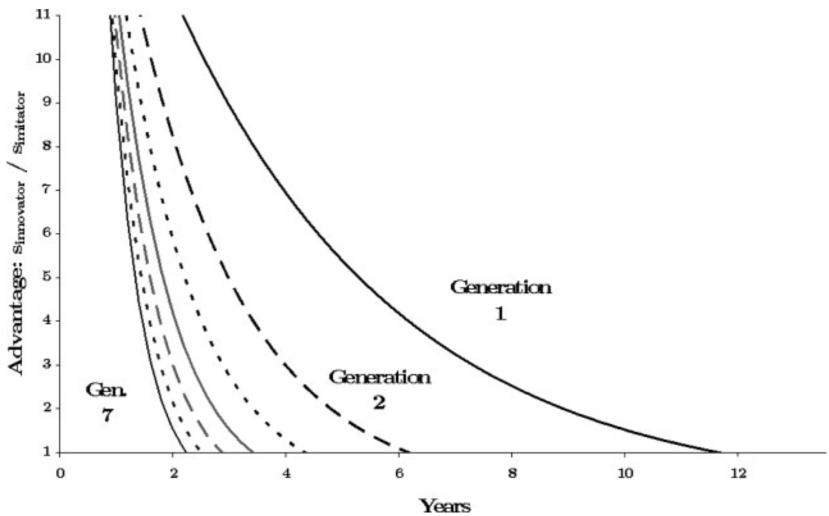

Figure 4
Estimated advantage to the innovator over time and generations (base model, for $t = 8$)
This figure plots the estimated dynamics of the demand advantage that an innovator has over its imitators. The vertical axis measures the parametric shift in the demand function for an innovator-engineered security relative to an imitator-engineered security. It is expressed as the predicted ratio of the innovator's market share to the imitator's market share: $s_{innovator}/s_{imitator}$. The horizontal axis measures the time elapsed after the first imitative deals any given security, in years. Each plotted line shows the estimated advantage dynamics for a given generation. The predicted advantage to the innovator decreases monotonically in the generation number. The advantage estimates are derived from the parameter estimates of the base nested logit demand model, aggregating the data through eight time periods.

equity-linked class.[15,16] These estimates are also economically significant for all aggregations. They imply, for example, that if the imitator had underwritten US$10 million more in equity-linked securities previous to a given generation, then the innovator's advantage would be gone by the fourth year rather than the sixth for second-generation securities and shortly after the first year rather than the second for seventh-generation securities.

I test whether the restrictions imposed by the chosen instruments over-identify the parameters of the model or not with a Hausman test. The test statistic is χ^2 under the null hypothesis that the model is correctly specified and the instruments over-identify the parameters. Table 4 reports this statistic and its p-value. The null is rejected with 92% confidence for $t = 12$, but it cannot be rejected with more than 90% confidence

[15] I tried using also the cumulative *number* of underwriting deals rather than the principal, but its estimated coefficient was never statistically significant.

[16] Notable is the case of YES shares, a mandatorily convertible preferred stock product introduced by Goldman Sachs seven months after Morgan Stanley introduced its PERCs. These two products are virtually identical in design, and the only difference between them seems to be their brand name.

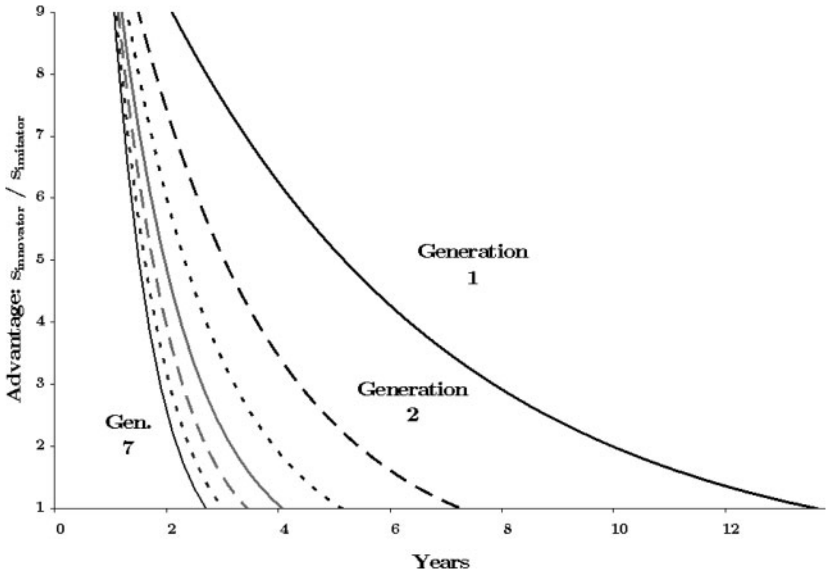

Figure 5
Estimated advantage to the innovator over time and generations (base model, for $t = 16$)
This figure plots the estimated dynamics of the demand advantage that an innovator has over its imitators. The vertical axis measures the parametric shift in the demand function for an innovator-engineered security relative to an imitator-engineered security. It is expressed as the predicted ratio of the innovator's market share to the imitator's market share: $s_{innovator}/s_{imitator}$. The horizontal axis measures the time elapsed after the first imitative deals any given security, in years. Each plotted line shows the estimated advantage dynamics for a given generation. The predicted advantage to the innovator decreases monotonically in the generation number. The advantage estimates are derived from the parameter estimates of the base nested logit demand model, aggregating the data through 16 periods.

otherwise. Rejection for $t = 12$ may be because of the fact that the instruments chosen do not introduce sufficient independent variation themselves to account for the variation in all the instrumented endogenous variables or may also indicate that the model is not fully specified, and the instruments themselves are correlated with other excluded exogenous variables. However, the rejection of over-identification at $t = 12$ is not a strong evidence against our choice of instruments, because almost all of the estimates are consistent across all the aggregations.

3.2 Issuers heterogeneity

The estimation described above may have allowed us to obtain consistent estimates of the own-price elasticity, but it may still yield implausible cross-price elasticities for varieties in different groups. Also, the test of over-identifying restrictions for the nested logit specification failed for $t = 12$. In Table 5, I show the results after adding the characteristics of the issuer to the model via interactions with the price variable. Thus,

Table 5
IV estimates for the base nested logit demand model with firm heterogeneity

Regressors (parameter name)	Parameter estimates			
	t = 16	t = 12	t = 11	t = 8
Underwriting spread (α)	−0.605 (0.579)	−1.041* (0.539)	−0.987* (0.581)	−1.782* (1.027)
Product market share, in logarithms (σ)	0.644* (0.396)	0.664* (0.347)	0.821* (0.461)	0.558 (0.383)
Underwriting spread interacted with market capitalization (US$ millions)	1.11*** (0.415)	1.91*** (0.500)	1.26* (0.743)	3.07*** (0.836)
Underwriting spread interaction with proportion of preferred stock	0.022 (0.015)	0.023* (0.015)	0.018 (0.014)	0.056 (0.028)
Innovator dummy[1]	4.980*** (1.370)	4.679** (1.625)	3.841** (1.698)	7.022*** (2.066)
Innovator dummy interacted with the security's generation number	0.111** (0.047)	0.188*** (0.065)	0.120** (0.053)	0.199*** (0.069)
Innovator dummy interacted with imitation time and generation number	−0.359*** (0.075)	−0.467*** (0.137)	−0.466*** (0.149)	−1.001*** (0.258)
Group dummies[2]	Yes	Yes	Yes	Yes
Time dummies[3]	Yes	Yes	Yes	Yes
Instrumented regressors: underwriting spread and product market share (in logarithms)				
χ^2 statistic for OIR test[4]	2.556	15.660	11.975	2.974
p-value	0.995	0.154	0.366	0.992
Average elasticity of market share to the underwriting spread	4.341 (1.939)	7.717 (3.490)	13.879 (6.296)	9.876 (4.722)
Median elasticity of market share to the underwriting spread	4.666	8.260	15.061	10.751
Number of predicted inelastic demands	18	10	7	3
Number of observations	332	261	250	177
R^2	0.354	0.369	0.304	0.290
F-statistic	113.19***	41.71***	55.23***	10.19***

The dependent variable is the difference of the logarithms of the market shares of a banks' underwriting of a security and of the market share of equity underwriting: $\ln(s_j) - \ln(s_o)$. The sample includes all offerings of equity-linked and derivative securities between 1985 and 2001. The parameters of a theoretical logit demand model are obtained using an Instrumental Variables estimator. The standard errors are obtained using a Huber–White covariance estimator, robust for heteroskedasticity, and within-security correlation. They are shown below the estimate, in parenthesis. Each observation consists of the market share of a banks' variety of a given security at a given period, associated with the characteristics of the security and the banker, in the same period, and their instruments. The sample time span has been split in 16, 12, 11 and 8 periods.

[1]Equals one when the underwriter is the innovator of that security.

[2]A dummy for each one of the 10 categories. Equals one when the security belongs to that category.

[3]A dummy for each period. Equals one when the observation corresponds to that period.

[4]Test of over-identified restrictions. The null hypothesis is that the model is correctly specified and that the instruments over-identify the variation in the endogenous regressor. The p-value is shown below the χ^2 statistic.

Estimates followed by ***, **, and * are significant to the 0.01 level, 0.05 level, and 0.1 level, respectively.

adding issuers' heterogeneity allows for different own-price elasticities for each issuer.

Despite the loss of observations, Table 5 summarizes results that do not differ qualitatively to the previous ones. The same dynamic behavior of the innovators' advantage is observed in all cases. For all aggregations, the initial advantage is bigger than in the previous specification, but it also decreases at a much faster rate. On average, for all cases of the model with firm-specific interactions, the advantage of each generation would be gone almost by the time predicted in the nested logit model without them.

Of the five issuers variables that I interact with price, only market capitalization and preferred stock as a percentage of market cap were found to be significant at a level higher than 90%. Their estimated coefficients were both positive. One possible explanation is that market capitalization is approximating the sources of finance available to the issuer. Similarly, because most equity-linked varieties are forms of preferred stock or convertible to preferred stock, firms with a larger proportion of this type of stock have more available instruments to raise capital, and thus their demands are more elastic to underwriting spreads.

Adding firm-specific data has increased the average estimated elasticities of demand through a higher $\hat{\sigma}$ but has a very similar relative number of violations in the elasticity test. With the added firm-specific regressors, the null hypothesis of over-identification and correct specification cannot be rejected even with 90% confidence.

3.3 Underwriter placement ability interactions

The cumulative equity underwritten with equity-linked securities by the bank has a very significant effect on the issuer's value. Moreover, the specification of the model improved considerably when this variable was included. Table 6 summarizes the estimates when I add as an explanatory variable its interaction with the innovator dummy. This variable loads significantly, suggesting that banks that have been stronger with these products earlier can profit more from innovating later. It is possible that successful underwriters in the past might have discovered characteristics from the issuers that allow them to perfect the structure of securities, targeted to a client population that will adopt it more easily. Also, one would expect that at some point in the innovation sequence, a small group of banks specialize in innovation in the given group. Morgan Stanley, for example, dominates as an innovator of mandatorily convertible preferred stock.

However, these results should be interpreted with care, because the cumulative volume itself does not load significantly anymore. It is possible that the interaction term must be picking up all the bank-specific variation itself, leaving no more for the former.

Table 6
IV estimates for the base nested logit demand model (II)

Regressors (parameter name)	Parameter estimates			
	t = 16	t = 12	t = 11	t = 8
Underwriting spread (α)	−1.224** (0.556)	−1.000** (0.503)	−0.858* (0.515)	−1.657 (1.102)
Product market share, in logarithms (σ)	0.443** (0.221)	0.460** (0.234)	0.449** (0.221)	0.674** (0.277)
Total principal underwritten in the equity-linked class by the underwriter (US$ millions)	−0.034 (0.025)	−0.033 (0.030)	−0.020 (0.028)	−0.067 (0.041)
Innovator dummy[1]	2.570** (1.007)	2.614** (1.097)	2.370** (1.096)	2.902* (1.652)
Innovator dummy interacted with the bank's equity-linked underwritten principal	0.130*** (0.029)	0.127*** (0.036)	0.115*** (0.034)	0.165*** (0.048)
Innovator dummy interacted with the security's generation number	0.152*** (0.036)	0.223*** (0.054)	0.155*** (0.039)	0.169*** (0.049)
Innovator dummy interacted with imitation time and generation numbers	−0.248*** (0.067)	−0.355*** (0.096)	−0.360*** (0.107)	−0.620*** (0.222)
Group dummies[2]	Yes	Yes	Yes	Yes
Time dummies[3]	Yes	Yes	Yes	Yes
Instrumented regressors: underwriting spread and product market shares (in logarithms)				
χ^2 statistic for OIR test[4]	4.781	10.091	7.585	2.352
p-value	0.572	0.121	0.270	0.799
Average elasticity of market share to the underwriting spread	5.616 (2.400)	4.625 (2.010)	3.929 (1.699)	12.553 (5.709)
Median elasticity of market share to the underwriting spread	6.236	5.086	4.393	14.370
Number of predicted inelastic demands	15	11	20	4
Number of observations	418	323	312	225
R^2	0.292	0.346	0.351	0.205
F-statistic	13.39***	15.64***	14.64***	11.01***

The dependent variable is the difference of the logarithms of the market shares of a banks' underwriting of a security and of the market share of equity underwriting: $\ln(s_j) - \ln(s_o)$. The sample includes all offerings of equity-linked and derivative securities between 1985 and 2001. The parameters of a theoretical logit demand model are obtained using an Instrumental Variables estimator. The standard errors are obtained using a Huber–White covariance estimator, robust for heteroskedasticity, and within-security correlation. They are shown below the estimate, in parenthesis. Each observation consists of the market share of a banks' variety of a given security at a given time period, associated with the characteristics of the security and the banker, in the same period, and their instruments. The sample time span has been split in 16, 12, 11, and 8 periods.

[1] Equals one when the underwriter is the innovator of that security.

[2] A dummy for each one of the 10 categories. Equals one when the security belongs to that category.

[3] A dummy for each one of the 12 periods. Equals one when the observation corresponds to that period.

[4] Test of over-identified restrictions. The null hypothesis is that the model is correctly specified and that the instruments over-identify the variation in the endogenous regressor. The p-value is shown below the χ^2 statistic.

Estimates followed by ***, **, and * are significant to the 0.01 level, 0.05 level, and 0.1 level, respectively.

4. Robustness Analysis and Further tests

In this section, I analyze the data set further to discuss the appropriateness of some of the assumptions made earlier and test empirically if the results presented above are robust to changes in some of these assumptions.

4.1 Trademarks on securities names

It is possible that some securities considered here as innovations may only be replications of existing designs but marketed with different names. If innovators register a trademark for the original name of the security, then its imitators could be confused as innovators of later generations. It may be incorrect to consider such varieties as innovative and their first underwriters as innovators rather than imitators. An inspection of all the 50 securities in the data set identifies differences in the design of almost all securities with respect to their previous generations.[17] Moreover, the setup of the estimation allows for intergenerational substitution effects. However, these effects are by construction smaller than intragenerational substitution effects, and the fact remains that the aggregation used probably defines too many varieties as innovative. It is worth noting, though, that labeling more banks as innovators rather than imitators in later generations is likely to bias the innovator's advantage estimate *downward*, because late generations have typically small market shares and are not as imitated as early generations. The current definition of an innovator associates observed small market shares to innovators, some of which could be imitators. Thus, because the possibly downward-biased estimate of the innovator's demand advantage over its imitators is already positive and significant, our inference from these results should not change much qualitatively.

To assess the robustness of the results, I propose two alternative exercises. First, I estimate the same parameters having excluded from the data set all those securities that were not imitated and whose name was trademarked. Second, I estimate the same parameters after redefining the varieties as banker-product group pairs rather than banker-security name pairs.

4.1.1 Estimation using only non-trademarked securities. I searched for all the 50 security names in the US Patent and Trademark Office records to see if there is a trademark for any. This criterion eliminates seven securities and

[17] Merrill Lynch (innovator) and Salomon Smith Barney co-managed the 73rd TOPRs issue, and Goldman Sachs (innovator) and Morgan Stanley co-managed the 16th QIDS issue. In both cases, I granted each co-manager with one full deal for the computation of its market share. Note too that the data set used here overlaps little with Nanda and Yun (1995), but the results here can very well be consistent with their hypothesis. They argue that the incentives to co-manage the first deals are smaller if innovators expect a large market for their innovation and a more dominant position in the market. This data set exhibits almost no co-management, deal sizes that are significantly larger than common stock underwriting deals, and dominant positions by innovators. The fact that 14 of the 17 co-managed deals happen in fourth or later generation securities could be also indicative that co-management is more likely when in late generations, which is when the evidence shows that the dominant position of innovators is smaller.

77 deals. The remaining securities either were or could have been imitated. Table 7 summarizes the new estimates, which are very similar than the ones obtained before. The estimated coefficient of price still has the correct sign for all the aggregations. The estimates of the price elasticity of underwriting demand are similar and so is the estimated coefficient for the substitution parameter, which is still between the correct bounds: $0.382 < \hat{\sigma} < 0.564$. The estimates of the innovator's advantage index follow the same pattern as all the previous results. The estimated initial advantage of the innovators is slightly larger with respect to Table 4, but the estimated speed at which the advantage decreases is larger.

4.1.2 Product groups as observational units. The second approach that I use to overcome the problem of defining correctly the set of varieties is to recompute underwriter market shares, this time over product groups rather than over individual security names. I redefine a variety as a banker-product group pair, and compute market shares accordingly in all periods. This aggregation is the other extreme of the previous one: only the products that start a category are considered as innovations. Before, any security name was considered as such. This aggregation rules out the fact that some generations are innovations in their own right. Note that, with this aggregation, the generational effects cannot be assessed, because all subsequent generations within a group are considered imitative varieties. The number of total varieties falls from 98 to 60 and so does the number of usable observations.

The results change little with this aggregation (Table 8). This aggregation fits the data well too: the joint hypothesis of correct specification and overidentification through the instruments is only rejected with 95% confidence for $t = 12$. The same dynamic pattern concerning the innovator's advantage is observed: it disappears with time. In this case, the size of the advantage is initially smaller, but this is not surprising given that this estimate measures the innovator's advantage over a typical bank in the product group rather than over a bank underwriting the same security only.

For this aggregation, the cumulative equity-linked principal that the bank has underwritten so far also increases the demand for its underwriting services within the category of securities, as does the accumulated number of innovations. Controlling for these effects, the coefficient of the underwriting fee is also significantly different from zero as well as the substitution parameter.

4.2 Further comments

Another concern may be that security underwriters actually cooperate, so the demand for a given bank's variety of a security is not independent of the demand for other bank's varieties if these banks were part of the others' underwriting syndicates. It is well known that underwriters

Table 7
IV estimates for the Nested Logit Demand Model excluding trademarked securities

Regressors (parameter name)	Parameter estimates			
	t = 16	t = 12	t = 11	t = 8
Underwriting spread (α)	−1.191** (0.565)	−0.621 (0.471)	−0.573 (0.465)	−0.944 (0.898)
Product market share, in logarithms (σ)	0.462** (0.232)	0.382* (0.236)	0.405* (0.222)	0.564** (0.270)
Total principal underwritten in the equity-linked class by the underwriter (US$ millions)	−0.036 (0.025)	−0.036 (0.030)	−0.024 (0.028)	−0.070** (0.035)
Innovator dummy[1]	2.631*** (1.014)	3.053*** (1.113)	2.693** (1.103)	3.703** (1.455)
Innovator dummy interacted with the bank's equity-linked underwritten principal	0.138*** (0.029)	0.137*** (0.036)	0.125*** (0.034)	0.171*** (0.041)
Innovator dummy interacted with the security's generation number	0.146*** (0.039)	0.224*** (0.056)	0.153*** (0.041)	0.161*** (0.048)
Innovator dummy interacted with imitation time and generation numbers	−0.268*** (0.067)	−0.410*** (0.095)	−0.412*** (0.106)	−0.737*** (0.196)
Group dummies[2]	Yes	Yes	Yes	Yes
Time dummies[3]	Yes	Yes	Yes	Yes
Instrumented regressors: underwriting spread and product market shares (in logarithms)				
χ^2 statistic for OIR test[4]	3.466	6.016	6.554	2.101
p-value	0.749	0.364	0.474	0.910
Average elasticity of market share to the underwriting spread	5.697 (2.469)	2.528 (1.120)	2.441 (1.074)	5.411 (2.453)
Median elasticity of market share to the underwriting spread	6.309	2.866	2.745	6.165
Number of predicted inelastic demands	12	15	10	6
Number of observations	377	291	282	204
R^2	0.327	0.427	0.418	0.426
F-statistic	14.29***	19.24***	16.73***	17.69***

The dependent variable is the difference of the logarithms of the market shares of a banks' underwriting of a given security and of the market share of equity underwriting: $\ln(s_j) - \ln(s_o)$. The sample includes all offerings of equity-linked and derivative securities between 1985 and 2001, except for the securities whose name was registered as a trademark in the US Patents and Trademarks Office. The parameters of a theoretical logit demand model are obtained using an Instrumental Variables estimator. The standard errors are obtained using a Huber–White covariance estimator, robust for heteroskedasticity, and within-security correlation. They are shown below the estimate, in parenthesis. Each observation consists of the market share of a banks' variety of a given security at a given period, associated with the characteristics of the security and the banker, in the same period, and their instruments. The sample time span has been split in 16, 12, 11, and 8 periods.

[1] Equals one when the underwriter is the innovator of that security.

[2] A dummy for each one of the 10 categories. Equals one when the security belongs to that category.

[3] A dummy for each time period. Equals one when the observation corresponds to that period.

[4] Test of over-identified restrictions. The null hypothesis is that the model is correctly specified and that the instruments over-identify the variation in the endogenous regressor. The p-value is shown below the χ^2 statistic.

Estimates followed by ***, **, and * are significant to the 0.01 level, 0.05 level, and 0.1 level, respectively.

Table 8
IV estimates for the Nested Logit Demand Model: demand for all banks' products within a group

Regressors (parameter name)	Parameter estimates $t = 16$	$t = 12$	$t = 11$	$t = 8$
Underwriting spread (α)	−0.818** (0.355)	−0.959** (0.442)	−0.995** (0.485)	−0.851** (0.403)
Product market share, in logarithms (σ)	0.736*** (0.200)	0.702*** (0.216)	0.733*** (0.197)	0.768*** (0.207)
Cumulative number of deals recorded by the bank in the equity-linked class	0.006** (0.003)	0.022* (0.012)	0.027** (0.013)	0.015 (0.011)
Innovator dummy[1]	2.441** (0.941)	2.507*** (1.066)	2.080* (1.100)	3.231*** (1.116)
Innovator dummy interacted with the bank's equity-linked underwritten principal	0.097*** (0.016)	0.079*** (0.021)	0.080*** (0.023)	0.0843*** (0.022)
Innovator dummy interacted with the time after imitation	−0.221*** (0.079)	−0.325*** (0.114)	−0.316** (0.129)	−0.613*** (0.180)
Group dummies[2]	Yes	Yes	Yes	Yes
Time dummies[3]	Yes	Yes	Yes	Yes
Instrumented regressors: underwriting spread and product market shares (in logarithms)				
χ^2 statistic for OIR test[4]	15.546	27.553	15.184	12.354
p-value	0.275	0.016	0.366	0.578
Average elasticity of market share to the underwriting spread	7.585 (3.173)	7.849 (3.170)	9.163 (3.732)	8.788 (3.687)
Median elasticity of market share to the underwriting spread	8.560	8.793	10.525	9.994
Number of predicted inelastic demands	10	7	7	3
Number of observations	278	211	205	147
R^{2s}	0.528	0.548	0.534	0.625
F-statistic	26.30***	28.26***	25.65***	21.86***

The dependent variable is the difference of the logarithms of the market shares of a banks' underwriting of all securities in a given group and of the market share of equity underwriting: $\ln(s_j) - \ln(s_o)$. The sample includes all offerings of equity-linked and derivative securities between 1985 and 2001. The parameters of a theoretical logit demand model are obtained using an Instrumental Variables estimator. The standard errors are obtained using a Huber–White covariance estimator, robust for heteroskedasticity, and within-security correlation. They are shown below the estimate, in parenthesis. Each observation consists of the market share of a banks' variety of a given security at a given time period, associated with the characteristics of the security and the banker, in the same period, and their instruments. The sample time span has been split in 16, 12, 11, and 8 periods.

[1] Equals one when the underwriter is the innovator of the first security in the group.

[2] A dummy for each one of the 10 categories. Equals one when the security belongs to that category.

[3] A dummy for each period. Equals one when the observation corresponds to that period.

[4] Test of over-identified restrictions. The null hypothesis is that the model is correctly specified and that the instruments over-identify the variation in the endogenous regressor. The p-value is shown below the χ^2 statistic.

Estimates followed by ***, **, and * are significant to the 0.01 level, 0.05 level, and 0.1 level, respectively.

frequently form syndicates to structure the deals and sell the securities. As Nanda and Yun (1995) argued, co-management of underwriting may be a form of cooperation that overcomes the free-rider problem in innovation: if all the competing banks underwrite the first issue of an innovative security jointly, then they can share the risks and the profits of the innovation minimizing the spill-overs. Thus, they argue, one would expect that underwriters would team up in syndicates more often when the market for the innovation is small or when a potential innovator does not expect to capture a large share of the new market. However, there are strong reasons to believe that the joint management of deals is not a common practice in the equity-linked and derivatives class. Of the 662 equity-linked deals observed, only 17 are co-managed. In 15 of these cases, one of the co-managers was indeed the innovator, but in only three cases the co-managed deal happened before the fifth deal. Only one first deal is co-managed, and only in two deals the innovator shared the management with a bank that also underwrote any other deal using that security.[18]

Another important point worth noting is that one cannot observe when each variety was made available by imitators, that is, since when rivals compete. The econometrician can only compute imitators' market shares after they have underwritten their first issue. For example, if an imitator appears on period 5, then there are no observations for this variety in period 4 or anytime before when the security was in the market but not yet imitated. It is possible that the bank was offering this variety well before we observe its first deal. The exclusion of these observations may bias the results but downward, decreasing the estimate of the innovator's advantage. Because entry of imitators can only happen sooner than we observe, the data set would have some more zero market share observations associated with imitators. Thus, the estimate we have now, which is already positive and significant, could even be larger. Note too that as we increase the time horizon to compute market shares, the effects of this problem should weaken, because the imitator appears, most likely, in the same period as the innovator, reducing the potential for bias. Precisely, we observe the highest estimates of the innovator's advantage at $t = 8$.

5. Patterns of Imitation in Equity-Linked Securities

One remaining concern is why are some securities imitated and others not. Understanding why and when these innovations are imitated sheds

[18] The grouping of generations captures the idea that some generations may actually belong to the same cohort of designs. I tried also grouping them in five categories: category 1 for first generations, 2 for second and third, 3 for fourth to sixth, 4 for seventh and eighth, and 5 for else. The association is still "negative" (i.e., later generations are imitated less frequently) but weaker (p-value is 0.183).

some more light into the nature of the game that investment banks play when underwriting issues of corporate derivatives.

5.1 The decision to imitate

There are 43 securities that could have been imitated, but only 18 of these were actually imitated. Even if imitation were costly and if imitators were at a disadvantage, it is clear that imitation should be less costly than innovation (e.g., free riding the costs of development and legal advice). Table 9 summarizes some preliminary findings that help to narrow our focus on possible explanations for why securities are imitated or not. Panel A groups generations in three categories: first-generation securities, securities between generations 2 and 6, and later generations. Imitation seems to be more frequent earlier: the null hypothesis that imitation is independent of the generation group can be rejected with 95% confidence. Panel B suggests that banks are more likely to be the innovators of a given generation if they had previously underwritten deals in the same group: we reject the independence between previous underwriting experience in the same group and innovating in the future with 96.6% confidence. Thus, it seems that doing deals helps banks to develop new products but that this helps less later in the product sequence.

If imitation is increasingly profitable as the demand for an imitator's variety increases faster for later generations, then why is it *less* frequent later in the sequence? A possible explanation is that the benefits of imitation spill-over to rival underwriters that do not imitate: if they observe more imitation by others, then it is more likely that they will be successful at attempting next generations' innovation. If it becomes easier to be successful with innovation later in the sequence, then the spill-overs increases in size and the incentives not to imitate increase.

Table 9
Summary of imitation and innovation facts

	No	Yes	Total
Panel A: Was the security imitated?			
First-generation securities	4	7	11
Generations 2–6	10	9	19
Generations 6–15	11	2	13
Panel B: Did the bank innovate previously in the product group?			
Bank is the imitator	18	14	32
Bank is the innovator	10	22	32

The sample includes all the 43 nontrademarked equity-linked and derivative securities between 1985 and 2001. Of these, 11 are a first-generation products (one per product group). There are 90 bank-security pairs (i.e., varieties) in the sample, of which 26 are of first generation. The Pearson χ^2 statistic is for the test with a null hypothesis that there is no statistical association between the two qualitative variables. The rows at each numeral show the data counts. In panel A, Pearson $\chi^2(2) = 6.124$; p-value = 0.047; and in panel B, Pearson $\chi^2(1) = 4.064$; p-value = 0.044.

For example, in a situation where all potential innovators are indifferent between imitating or not, the probability of observing imitation will be smaller later in the sequence, because it takes a smaller individual probability to have a significant spill-over (this claim is illustrated formally in the supplement).

Note that another conjecture might be that banks have dealt enough by the time they attempt at innovating late-generation securities, so there is little gain from imitating if it cannot increase enough anymore the future success probabilities. But even in that case, imitation profits are higher because of the higher demand. Moreover, a bank can always deal its innovations and imitate other bank's products simultaneously.

Thus, we predict that (1) if imitation does not spill-over to nonimitators, then imitation should be independent of the generation number; (2) if imitation spills-over, then imitation may be less frequent in later generations; and (3) the higher the expected profits from innovating generation $g + 1$, the more likely is the imitation of generation g.

5.2 The Evidence on imitation

To test these predictions, I fit probabilistic regressions of a security being imitated or not on the generation number of the security. I use all 43 securities whose names were not trademarked in the US Patent and Trademark Office. As controls, I use (1) the total number of deals for all previous generations in the same product group, as a proxy for the expected size of the market of the next innovation; (2) the number of previous generations in the same product group that were actually imitated, as an additional source that speeds up the spill-overs; and (3) dummy variables that indicate if any of the underwriters of the given security were also dealers of the first generation, the second, and so on.

Table 10 summarizes the probit estimates of these regressions. The covariance matrix has been estimated using Huber–White's estimator, which is robust to correlation of the securities within the group and robust to heteroskedasticity. Columns 1 and 3 use the generation number as the main regressor, and columns 2 and 4 use instead the same categorical variable used in panel A of Table 9. Columns 1 and 3 show that a given generation is imitated with a probability that is smaller by 0.08 or 0.07 than the probability that previous generation is imitated, respectively. These estimates are significantly different from zero with 99% confidence. Columns 2 and 4 show a similar result when we group generations as in Table 9.

As expected, in larger product groups, the next product is imitated more frequently, and this effect is also statistically significant to the 0.01 level always. Note that I control for the number of previous *imitated* generations, which is also statistically significant and, as

Table 10
Probabilistic regressions for the imitation of securities

Regressors	Coefficient estimates[1] 1	2	3	4
Generation number	−0.081*** (0.022)	−0.569** (0.205)	−0.070*** (0.019)	
Generation category (1–3)[2]				−0.527** (0.194)
Total deals underwritten by previous generations	0.014*** (0.005)	0.014*** (0.005)	0.020*** (0.003)	0.020*** (0.002)
Number of previous generations that were imitated	−0.261*** (0.080)	−0.217*** (0.061)	−0.346*** (0.068)	−0.302*** (0.064)
Were the underwriters in this generation also of the first generation? (1 if yes or 0 otherwise)	0.357 (0.232)	0.476* (0.230)	0.421** (0.193)	0.514*** (0.190)
Were the underwriters in this generation also of the second generation? (1 if yes or 0 otherwise)			−0.338** (0.155)	−0.320* (0.182)
Number of observations	43	43	43	43
χ^2 statistic for Wald test[3]	131.33***	271.90***	275.52***	563.85***
Pseudo R^2	0.185	0.263	0.207	0.281

This table summarizes the parameter estimates of a probabilistic regression of the probability that a security is imitated. The dependent variable is the binary indicator that takes value 1 if the security was imitated or zero otherwise. The sample includes all 43 nontrademarked equity-linked and derivative securities between 1985 and 2001.
[1] Estimates of the change in the probability that the security is imitated given an infinitesimal change in the associated continuous variable or a change of 0–1 in the associated dummy variable; estimated at the sample mean. The standard errors are shown below the estimate, in parenthesis, and are obtained using a Huber–White covariance estimator, robust to heteroskedasticity, and correlation within-product groups.
[2] Category is 1 if generation number is 1, 2 if generation number is between 2 and 6, and 3 otherwise.
[3] The null hypothesis is that all coefficients are equal to zero.
Estimates followed by ***, **, and * are significant to the 0.01 level, 0.05 level, and 0.1 level, respectively.

expected, reduces the probability that the subsequent products are imitated. Also, securities whose underwriters had also dealt the group's first generations are imitated with a higher probability, and all other things constant. In other words, securities that are imitated with a higher probability are those for which their innovators or imitators have had accumulated experience innovating or imitating the first generation. Having experience in the second generation has a negative effect on imitation (experience in further generations was rejected of all specifications). This result, though, may just be a reflection of the main hypothesis: that banks have less incentives to imitate later because of the spill-overs of other imitators.

6. Summary and Discussion

This article has provided new evidence of the sources of first-mover advantages in innovations in finance. This article argues that for innovations in equity-linked products, banks offer underwriting services that are differentiated within the same product type and across different generations of similar products. It estimates the preference parameters for the nested choice of a security type and an underwriter and finds that preferences for innovators are, on average, stronger than for imitators. These preferences are initially stronger for later generation innovations, but they diminish with time. Furthermore, they diminish faster for later generations. My interpretation is that late generations appear naturally when more information has been aggregated about the family of securities they belong to, making learning about the innovator's private information easier.

The results here have characterized the nature of the game played by underwriters of corporate derivatives. In these markets, the largest, most reputed investment banks compete to underwrite issues which are, on average, much larger than common stock deals. Banks differentiate from others using their innovative ability, their engineering skills, and their placement capacity. Success at innovation gives them the discretion to choose underwriting fees through a demand advantage. Since this advantage dissipates with time, it is not surprising that banks typically do not co-manage deals with others: co-management could disclose their private information to imitators faster and shorten his advantage horizon.

Banks also imitate their rivals at the early stages of innovation sequences. Imitating early generations seems to increase the chances that banks will develop innovations in the future. But even as imitation becomes more profitable along the innovation sequence, less imitation is observed later. Imitation becomes less frequent faster if there has been more of it in the earlier generations, possibly because banks that do not imitate also learn by observing deals completed by rival imitators. The evidence here supports this hypothesis. It is quite apparent too that the innovative process evolves to a situation, where the top-tier banks have all learnt enough about the type of products in a group, and most of them underwrite deals with their own differentiated securities.

It is shown here that underwriting of innovative corporate products is quite different to standard equity underwriting, where there is little room for differentiation structuring deals, and banks are differentiated rather by the coverage of their analysts or the quality of stocks they underwrite. Common stock underwriting also exhibits a great deal of cooperation between banks that need each other's clienteles to sell their issues successfully. Furthermore, deals are priced almost symmetrically. It is clear that the determinants of market shares in both lines of business are quite different. Still, it is likely that the two lines of business are related. To what extent the bank's success in one of them may affect the success in the other remains an open question. A fertile ground for future research may consist of the study of the investment bank as a multiproduct firm, where innovation in corporate products allows the bank to increase its profit margins by reaching larger deals but where cooperation with other banks in more conventional securities markets is crucial for reputation acquisition.

The findings here also motivate future research on how the market structure affects the speed of financial innovation. The benefits of innovation are decreasing, as imitators catch up faster along the sequence. Yet, more innovation feeds back into the demand for future generations. Are the incentives to innovate increasing or decreasing along the sequence? Future research could look into the determinants of the speed and duration of innovation across different product groups, including not only equity-linked products but also other classes of corporate securities.

This article has taken innovation as exogenous. The set of choices available to firms was taken as given at each time. Certainly, one interesting way to continue this line of research would be to identify the preferences of firms for new securities at each time they make their choices. It seems that banks learn about their client needs for given structures and offer more valuable securities in the future. If the choices of the firm were to choose a security of a set of already existing securities or to rather choose to be the first issuer of a new security, then the data in each deal could reveal what determines when an innovation is to be introduced.

Appendix A

Proof of Proposition 1 Let j and k, l be varieties of the same security g, and q an arbitrary variety of another arbitrary security. By Lemma 2

$$v_j(y-p,1) \geq v_k(y-p,0),$$
$$v_j(y-p,1) \geq v_l(y-p,0).$$

The aggregate demand for the innovator's variety, given than varieties of the same security are priced symmetrically, is

$$q_j(\mathbf{p},.) = \int_{\varepsilon} G(\Delta v_g + \varepsilon') \prod_{l\neq k, l\in \mathcal{J}_g} G[v(y-p,1) - v(y-p_l,0) + \varepsilon'] \times \prod\nolimits_{q\neq j, q\notin \mathcal{J}_g} G[v(y-p,1) - v(y-p_q,b) + \varepsilon'] dG(\varepsilon').$$

For the imitators' varieties, the demand is

$$q_k(\mathbf{p},.) = \int_{\varepsilon} G(-\Delta v_g + \varepsilon') \prod_{l\neq k, l\in \mathcal{J}_g} G[v(y-p,0,.) - v(y-p_l,0,.) + \varepsilon'] \times \prod\nolimits_{q\neq j, q\notin \mathcal{J}_g} G[v(y-p,0) - v(y-p_q,b) + \varepsilon'] dG(\varepsilon').$$

Clearly, since $G(.)$ is a strictly increasing function, by Lemma 2

$$G(\Delta v_g + \varepsilon') \geq G(-\Delta v_g + \varepsilon');$$
$$G[v(y-p,1) - v(y-p_l,0) + \varepsilon'] \geq G[v(y-p,0) - v(y-p_l,0) + \varepsilon'],$$

and

$$G[v(y-p,1) - v(y-p_q,b) + \varepsilon'] \geq G[v(y-p,0) - v(y-p_q,b) + \varepsilon'].$$

and thus

$$q_j(\mathbf{p},.) \geq q_k(\mathbf{p},.).$$

Note that if $\Delta v_g > 0$, then $q_j(\mathbf{p},.) > q_k(\mathbf{p},.)$.

Proof of Lemma 1. The proof follows directly from the fact that $G(.)$ is a strictly increasing function. If Δv_g is decreasing with time, then so is the difference between $G(\Delta v_g + \varepsilon')$ and $G(-\Delta v_g + \varepsilon')$. The aggregate demand $q_j(\mathbf{p},.)$ is also strictly increasing in $G(.)$, so $q_j(\mathbf{p},.) - q_k(\mathbf{p},.)$ is decreasing in Δv_g, and thus decreasing in time.

Appendix B: List and Brief Description of the Equity-Linked and Derivative Securities in the SDC Database from 1985 to 2001

Name	Description
Debt products	
RISRS	Rising Interest Subordinated Redeemable Securities Debt notes subordinated to existing and future debt notes
Convertible debt (zero coupon)	
LYONS	Liquid Yield Option Notes. Zero-coupon bonds, convertible to common, stock, temporarily noncallable, and nonputable
Convertible debt (dividend paying)	
SIRENS	Set-Up Income Redeemable Equity Notes. Low-dividend paying, subordinated debt, callable, convertible to common stock
ICONS	Investment Company Convertible Notes. Notes issued by closed-end mutual funds that are putable, conversion price adjusted yearly
Convertible preferred	
PERCS	Preferred Equity Redemption Cumulative Stock. Mandatorily convertible preferred, after three years, callable anytime
YES shares	Yield Enhanced Stock. High-dividend paying preferred, mandatorily convertible in three years, callable anytime
DECS	Dividend Enhanced Convertible Stock. Noncallable, mandatorily convertible preferred stock
X-Caps	Exchangeable Capital Securities. Fixed dividend paying, perpetual, callable, convertible to subordinated preferred stock
ACES	Automatically Convertible Equity Securities. Putable, mandatorily convertible to common stock in four years
PRIDES	Preferred Redeemable Increased Dividend Equity Securities. Callable, nonputable, four years conversion, premium dividend paying preferred stock
PEPS	Participating Equity Preferred. High-dividend yield, four years conversion, callable, and putable preferred stock
SAILS	Stock Appreciation Income Linked Securities. Three years mandatory conversion, converted at premium if common stock appreciates.

Appendix B
(continued)

Name	Description
AutoComEx	Automatic Common Exchange Securities. Treasury notes, mandatorily convertible to issuer's common in three years. Callable.
PEPS (2)	Premium Equity Participation Securities. Preferred Stock mandatorily convertible in three years, conversion ratio determined *ex post*
STRYPES	Structured Yield Product Exchangeable. Preferred stock convertible to common or cash. Initial 20% appreciation given up in exchange of higher yield
MARCS	Mandatorily Adjustable Redeemable Preferred. Callable, convertible in four-year preferred stock. Depreciation of convertible common bounded at 10%
MEDS	Mandatorily Exchangeable Debt Securities. Four-year mandatorily convertible preferred at a premium
Trust Convertible Preferred	Mandatorily convertible preferred, issued by a business trust, not a Limited Liability Corporation
TRACES	Trust Automatic Common Exchange Securities. Trust-originated preferred stock convertible to common stock. Conversion after three years.
Short-term tax advantage	
FRAPS	Floating/Adjustable Rate Preferred Stock. Preferred Stock that pays a fixed for five years. Adjustable rate after fifth year, if not called
Perpetual tax advantage	
MIPS	Cumulative Guaranteed Monthly Income Preferred. Perpetual preferred stock. Dividend payments used as tax deductible by the issuer
EPICS	Exchangeable Preferred Income Cumulative Shares. Perpetual preferred stock with the issuer's option to convert to common anytime
MIDS	Monthly Income Debt Securities. Perpetual preferred stock with the issuer's option to defer monthly dividend payments
TOPRS	Trust-Originated Preferred. Forty-nine-year preferred stock, issued by business trust. Monthly cumulative fixed cash distributions
QUIDS	Quarterly Income Debt Securities. Fifty-year preferred with bounded deferral horizon and initially noncallable.
QUIPS	Quarterly Income Preferred Securities. Forty-nine-year preferred with the issuer's option to shorten maturity
QUICS	Quarterly Income Capital Securities. Tax-deductible, thirty-year bonds with interest deferral option
RES-CAPS	Reset Capital Securities. Perpetual preferred stock, with an initial call protection
COPRS	Canadian Originated Preferred Securities. Nonconvertible, tax-deductible dividend-paying preferred, issued by Canadian LLCs.
Convertible tax advantage	
Convertible MIPS	Optionally convertible perpetual preferred stock that is tax deductible. Initially protected from a call, with optional interest deferral
TECONS	Term Convertible Shares. Mandatorily convertible, tax deductible, and callable only if price of common stock is high enough
Convertible TOPRS	Same as TOPRS (above) but mandatorily convertible to common stock
QDCS	Quarterly Debt Capital Securities. Twenty years or less securities, convertible to preferred stock at maturity. Tax-deductible coupon, callable
EPPICS	Equity Providing Preferred Income Convertible. Forty-year convertible bonds, convertible to cash, common, or preferred stock
TRUPS	Trust Preferred Stock. Callable notes, tax deductible, that investor has to buy jointly with preferred stock from a subsidiary of issuer

Appendix B
(continued)

Name	Description
Convertible QUIPS	Same as QUIPS (above), convertible to common, at different maturities
Index-tied principal	
PERLS	Principal Exchange Rate-Linked Securities. Dollar denominated bonds with redemption value tied to a "hard" currency
SIRS	Stock Index Return Securities. Bonds with face value adjusted upward if S&P MidCap 400 index appreciates enough
MITTS	Market Index Target-Term Securities. Zero-coupon bond, principal tied to the S&P 500 Index
SMARTS	Stock Market Annual Reset Term Notes. Variable-capped coupon paying bond, principal can be tied to S&P or Amex Oil indexes
EPS	Equity Participation Securities. Zero-coupon bonds. Appreciation of principal pegged to S&P 500, capped, and floored.
CPNs	Currency Protected Notes. Principal of bonds tied inversely to Canadian six-month banker's acceptance rate
SUNS	Stock Upside Note Securities. Principal tied to a basket of telecom stocks. The depreciation is floored
CUBS	Customized Upside Basket Securities. Zero coupon, downside protected bond with principal tied to a basket of stocks across industries
Stock-tied principal	
ELKs	Equity-Linked Securities. Annual interest paying notes. Principal tied to the performance of the stock of a chosen firm
YEELDS	Yield-Enhanced Equity-Linked Debt Securities. Three-year notes. Principal tied to the stock of a firm in the IT sector. Appreciation is capped
CHIPS	Common-Linked Higher Income Participation Securities. High-yield annual coupon debt with appreciation tied to stock options, floored, and capped
PERQS	Performance Equity-Linked Redemption Quarterly Pay Securities. Quarterly coupon paying notes. Principal tied to a given stock at investor's option
Privatization exchangeable debt	
PENs	Privatization Exchangeable Notes. Foreign currency and dollar denominated bonds exchangeable for a privatized firm shares

References

Allen, F., and D. Gale, 1994, *Financial Innovation and Risk Sharing*, MIT Press, Cambridge, MA.

Anderson, S., A. DePalma, and J.-F. Thisse, 1989, "Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach," *The Review of Economics Studies*, 56(1), 21–35.

Berry, S., 1994, "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242–262.

Berry, S., J. Levinsohn, and A. Pakes, 1995, "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Bhattacharyya, S., and V. Nanda, 2000, "Client Discretion, Switching Costs, and Financial Innovation," *The Review of Financial Studies*, 13(4), 1101–1127.

Black, D., and W. Silber, 1986, "Pioneering Products: Some Empirical Evidence from Futures Markets," working paper 395, Salomon Brothers Center for the Study of Financial Institutions, New York: New York University.

Caplin, A., and B. Nalebuff, 1991, "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1), 25–59.

Carter, R., F. Dark, and A. Singh, 1998, "Underwriter Reputation, Initial Returns, and the Long-Run Underperformance of IPO Stocks," *The Journal of Finance*, 53(1), 285–311.

Carter, R., and S. Manaster, 1990, "Initial Public Offerings and Underwriter Reputation," *The Journal of Finance*, 45(4), 1045–1067.

Finnerty, J., 1992, "An Overview of Corporate Securities Innovation," *Journal of Applied Corporate Finance*, 4(1), 23–39.

Herrera, H., and E. Schroth, 2000, "Profitable Innovation Without Patent Protection: The Case of Credit Derivatives," working paper 01–02, Berkeley Center for Entrepreneurial Studies, New York: New York University.

Kanemasu, H., R. Litzenberger, and J. Rolfo, 1986, "Financial Innovation in an Incomplete Market: An Empirical Study of Stripped Government Securities," working paper 26–86, Rodney White Center, Wharton School, Philadelphia: University of Pennsylvania.

Loughram, T., and J. Ritter, 2004, "Why has IPO Underpricing Increased Over Time," working paper, Gainsville, FL: University of Florida.

McFadden, D., 1981, "Econometric Models of Probabilistic Choice," in C. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, MA.

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, Ann Arbor, MI: University of Michigan Business School.

Nevo, A., 2000, "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9(4), 513–548.

Pratt, T., 1991, "How Percs Became the Year's Hottest Product," *Investment Dealers' Digest*, 57(48), 20.

Pratt, T., 1995, "SEC Issues No-Action Letters on Streamlined Trust 'MIPS'," *Investment Dealers' Digest*, 61(13).

Ross, S., 1989, "Institutional Markets, Financial Marketing, and Financial Innovation," *The Journal of Finance*, 44(3), 541–556.

Tufano, P., 1989, "First-Mover Advantages in Financial Innovation," *Journal of Financial Economics*, 3(2), 350–370.

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

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