The development and spread of financial innovations

Isaiah Hull
Research Division, Sveriges Riksbank

I study financial product innovation in a model with two classes of agents: "sophisticated" and "unsophisticated." Unsophisticated agents are hit with frictions that lower the return to a conventional asset they hold. Sophisticated agents construct financial innovations that are perfect substitutes for the conventional asset, but are not subject to the friction. In the absence of complete information, unsophisticated agents learn about innovations through a contagion process, as they encounter competitors who have already adopted them. The model yields two equilibria: in one, the innovation persists; in the other, it disappears. Only one equilibrium is stable, and this is determined by the strength of the contagion and by early strategic interactions between sophisticated agents. The model suggests mechanisms for several empirical regularities in the financial innovation literature. Additionally, two applications demonstrate how to estimate the contagion parameter with a short time series of data, and how to use it to predict whether a financial innovation will spread.

Keywords. Financial innovation, innovation, financial markets.

JEL classification. G10, G12, G14, O31, O33.

1. Introduction

When a financial product innovation is introduced, investors do not have access to a historical series of returns from which they can infer the properties of that innovation. Instead, they must rely on the analysis of ratings agencies and on the information revealed through the decisions of the individual and institutional investors holding it. This cross section of information about institutional and individual holdings is often incomplete, and will only slowly be revealed through a search process, since centralized information about new securities often emerges with a long lag. Thus, investors must choose whether or not to hold the security while information remains imperfect.

We analyze this process through a model of financial product innovation, where frictions arise that lower the relative return to an existing asset (hereafter, the "conventional" asset), making it profitable to construct innovations. The friction may take the form of the proposed categories in Miller (1986), Campbell (1988, Chapter 16), or Frame and White (2004). That is, it may be generated by changes in technological opportunity, market power concentration, firm size, product demand conditions, regulation, taxes,

Isaiah Hull: isaiah.hull@riksbank.se

or macroeconomic conditions. Each of these is an event that may lower the relative return to a conventional asset. A tax, for instance, lowers the return directly. A new technological opportunity, such as the development of computer systems that price exotic derivatives, improves the return of the financial innovation, lowering the relative return of the conventional asset.

In the model, these stylized facts are captured through interactions between two classes of agents. Unsophisticated agents are assumed to hold a single unit of currency and use it to purchase the conventional asset at the start of the model. Once the friction is introduced, sophisticated agents decide whether to pay a fixed cost to research and develop an innovation that resolves the friction. If the innovation is developed, the sophisticated agents will market it to their base of unsophisticated clients, a small number of whom will adopt it initially.

After the innovation has been developed, unsophisticated agents holding the conventional asset learn about it by searching over the portfolios of other unsophisticated agents to determine whether they hold it. This occurs empirically when firms acquire information about other firms, swap employees, discuss best practices, undertake syndication projects, and exchange information through third parties, such as common academic co-authors. A large empirical literature documents the impact of such spillovers. Jaffe, Trajtenberg, and Henderson (1993), Cockburn and Henderson (1998), Cohen and Levinthal (1989), and Lerner (2006) demonstrate the importance of such local spillovers of information in the context of financial and nonfinancial innovations. Hong, Kubik, and Stein (2004, 2005), Antweiler and Frank (2004), Ivkovic and Weisbenner (2007), and Kaustia and Knupfer (2012) document the importance of social learning in finance more generally.

The interactions between unsophisticated agents in the model can be aggregated into a susceptible–infected–susceptible (SIS) model from epidemiology. We exploit this property to make use of well developed tools from the epidemiology literature, including the concept of $\mathbb{R}_0$ and the associated threshold condition.[1] The benefits of this approach are threefold: first, it captures interactions that arise in new markets when centralized sources of information are not yet available—or at least when returns histories are short. Second, it permits us to model an important part of the nature of financial innovations: they are sometimes transitory and sometimes permanent (Graham and Dodd (1996), Tufano (2003)). And third, it allows us to make predictions about the nature of specific financial innovations, even if we only have a time series of aggregate data. This is useful because detailed panel data are rarely publicly available for new financial innovations, but time series volume data can often be obtained. Frame and White (2004) identify this as a problem that has hindered empirical work on financial innovation.

The model yields two categories of equilibria: one where the innovation disappears and one where it persists.[2] Only one of these equilibria will be stable. This will be de-

---

[1]See Keeling and Rohani (2008) for an overview of the SIS family of models and their associated tools.

[2]One benefit of modeling the persistence of innovations, rather than their profitability, is that the model does not place heavy demands on the data. The approach introduced in this paper will permit us to make predictions about financial innovations using the Federal Reserve Board's Flow of Funds data, which are publicly available and regularly updated. More detailed data about financial innovation—including the profitability of individual firms—are more difficult to obtain and often proprietary.

termined by the strength of contagion, captured by an analog of the basic reproductive number from epidemiology,[3] $\mathbb{R}_0$, and the choices that sophisticated agents make at the start of the model. The measure of contagion in the model depends on the liquidity of the innovation, the true size of the friction applied to the conventional asset, and the speed at which unsophisticated agents learn from observing peers.

The model suggests mechanisms for several stylized facts in the financial innovation literature. In particular, it explains why innovators might tend to be larger and less profitable (Lerner (2006)). It also suggests an explanation for why financial innovators not only fail to deter imitators, but even actively seek out syndicated release partners for new securities (Tufano (1989), Lerner (2006), Schroth (2006), Herrera and Schroth (2011)), rather than attempting to exclude them with barriers to entry.

This paper is not the first to apply an epidemiological model to an economic problem. Carroll (2003) uses an epidemiological model to describe inflation expectations formation. Additionally, Sommer and Carroll (2004), Hong, Hong, and Ungureanu (2011), Pfajfar and Santoro (2012), and Shive (2010) use epidemiological models in economic and financial applications. This paper makes further methodological contributions by applying the mathematical tool set developed around the SIS family of models, including $\mathbb{R}_0$ and the associated threshold condition.

The epidemiological component of this paper also borders on the literature initiated by Bikhchandani, Hirshsleifer, and Welch (1992), which considers social and economic "cascades." The concept of a cascade is broad and may be used to describe many phenomena. It may also be formalized using several different modeling techniques. Importantly, cascades typically spread to become the dominant social or economic mode, whereas, during an epidemic, a disease may grow and persist without infecting most of the population.

This paper will proceed as follows: In Section 2, I set up a model of financial product innovation with two classes of agents, describe how $\mathbb{R}_0$ arises from the model, and derive several theoretical results. In Section 3, I demonstrate how $\mathbb{R}_0$ can be estimated, how its standard error can be constructed, and how the related threshold condition may be checked. I also show how this approach could have been used to make predictions about the spread of asset-backed securities (ABS) exchange-traded funds (ETFs). Finally, in Section 4, I conclude.

2. The model

There are two types of agents, who are referred to as *sophisticated* (S) and *unsophisticated* (U). Unsophisticated agents initially hold a single unit of a conventional asset. This conventional asset previously generated a riskless return of r, but is now subject

---

[3]With respect to epidemiology, estimating $\mathbb{R}_0$ for different types of infectious diseases, in different regions, and on the most recent data plays an important part in effective disease control. A high estimate of $\mathbb{R}_0$ in a particular region may suggest that an outbreak will occur if the disease is introduced, and that control measures, such as vaccination or a public health initiative, need to be implemented immediately. For example, see Mukandavire, Smith, and Morris (2013), who use this approach to analyze the recent cholera outbreak in Haiti.

to a friction that lowers its return to $r - \zeta$. Since unsophisticated agents do not have the expertise to resolve this friction internally, they must instead rely on sophisticated agents, who produce financial innovations. Initially, the size of the friction, ζ, is small enough to ensure that the production of a financial innovation is unprofitable, satisfying the Modigliani–Miller (M&M) (1958) equivalent condition for financial innovations hypothesized by Tufano (2003).[4] At the start of the model ($t = 0$), a shock is introduced that increases the size of the friction, potentially making it profitable for a sophisticated agent (S) to resolve it by creating a financial innovation.

There are N sophisticated agents (S), who are assumed to each have some initial mass, μ_i, of unsophisticated (U) agent clients, each of whom uses his initial endowment of currency to purchase one unit of the conventional asset and pay an associated fee, P. After the friction is introduced, the return to the conventional asset is reduced by ζ. A sophisticated agent may choose to pay a fixed cost, c, to develop an innovation that is not subject to this friction and to introduce it to her clients, who may purchase it for a fee, P'.

2.1 Unsophisticated agents

Unsophisticated agents each hold a single unit of currency and must choose to invest it in either the conventional asset or the financial innovation in each period. These agents also lack information about the impact of the friction and must estimate its size with some idiosyncratic error after it is issued ($t = 0$): $\zeta_{i0} = \zeta + \varepsilon_{i0}$.[5] Since unsophisticated agents are myopic, they will adopt the financial innovation if doing so maximizes expected profits in the following period:[6]

$$r - P' > r - \zeta - \varepsilon_{i0} - P. \tag{1}$$

This condition implies a threshold value for the idiosyncratic error's estimate, $\hat{\varepsilon_{i0}}$. Above this value, the agent will adopt the innovation at $t = 0$. Let χ denote the mass of

---

[4]Tufano (2003) suggests that an M&M-style condition would hold with respect to financial innovation in the absence of changes in frictions. The model's initial equilibrium is intended to replicate this equilibrium state, where financial innovations are not introduced, unless a shock that increases the size of the friction is applied to the conventional asset.

[5]I assume that the sophisticated agent can communicate some information about the innovation's value to the unsophisticated agent, but cannot remove all uncertainty. This is at least partly driven by the sophisticated agent's own uncertainty about the product, which induces the unsophisticated agent to collect information from the innovator, from public sources, and from sources idiosyncratic to the firm. Importantly, the sophisticated agent does not need to know the exact size of the financial friction, ζ, to solve its choice problem. She only needs to know that the friction exists and that it has generated a wedge between P and P'.

[6]Note that we implicitly assume that unsophisticated agents are risk-neutral. If, instead, they were risk-averse, then the specification used would make financial innovations less attractive by incorporating some uncertainty into their return, but not the conventional asset's return. The assumption of risk neutrality should not affect the qualitative results; however, if the model were extended to cover the case of risk-averse unsophisticated agents, then it might be necessary to incorporate uncertainty into the conventional asset's return, too, to avoid a counterfactual bias toward the conventional asset.

agents with values of ε_{i0} above the threshold, and let μ_i denote the mass of unsophisticated agents who are clients of sophisticated agent i, who chooses to issue the innovation. We may write the initial mass of agents who adopt the financial innovation as $I_0 = \chi\mu_i$, where I_t denotes the mass of unsophisticated agents who hold financial innovations. This also denotes the proportion of funds invested in the innovation, since each unsophisticated agent holds one unit of currency worth of one asset.

Unsophisticated agents do not make their internal estimates of the friction's value public. Consequently, each agent may only infer something about the value of another agent's estimate by observing his/her choice to hold the innovation or the conventional asset. During periods $t = 1, \ldots, \infty$, unsophisticated agents engage in a search process to determine whether other agents have adopted the innovation. If agent j encounters agent i, who has adopted the innovation, then he may infer that one of two things is true: (i) that agent i has a high value of ζ_{i0}, which increases agent j's subjective probability that ζ is positive, or (ii) that agent i has encountered other agents with a high value of ζ_{i0}, which also increases j's probability that ζ is positive. Thus, conventional asset holders learn about the true size of the friction—and, therefore, the value of the innovation—by simply observing that other unsophisticated agents hold it. Since agent i cannot observe the private information of other unsophisticated agents she encounters, she cannot discern the reason why an innovation holder holds the innovation. Instead, it treats each innovation holder as if he/she provides the same information. If j does not hold the innovation, but encounters i who does, then j will switch to the innovation with probability ψ, which is j's subjective probability that ζ is positive, conditional on observing that i holds it.

Unsophisticated agents do not improve their estimates of ζ over time. Instead, they learn in a cross-sectional sense: when returns seem good (i.e., ζ_{it} is high) or when many other agents think they are good (i.e., I_t is high), they switch to the innovation. This permits unsophisticated agents to behave as if they learn over time (i.e., with the evolution of I_t), even though all learning is cross sectional. It also permits assumptions about information to be less demanding: agents can identify other agents who hold the innovation by searching, but cannot directly view private estimates of the friction's size. Importantly, this learning does not depend on which sophisticated agent the unsophisticated agent purchases assets from; rather, it arises from interactions with other unsophisticated agents, and all unsophisticated agents are assumed to interact with each other with equal probability.

Unsophisticated agents may also switch back to the conventional asset when it performs well. If, for instance, the conventional asset generates an idiosyncratically large return that appears to compensate for the friction, then an unsophisticated agent who observes this will want to switch back to the conventional asset. The probability of this occurring is constant for a given r, ζ, P, and P', and is given by considered

$$\int_{-\infty}^{P'-P-\zeta} f(e)\, de. \tag{2}$$

where $f(e)$ is the probability distribution function of the estimation error.

Unsophisticated agents who observe a relatively favorable return to the conventional asset will not always be able to return to it immediately. If an innovation is less liquid, then an agent will have a lower probability of being able to swap back to the conventional asset in a given period. If, for example, the asset is perfectly illiquid—that is, there is no secondary market and the sophisticated agent will not repurchase it—then the unsophisticated agent must hold it in all remaining periods.

The assumptions we have made about unsophisticated agents allow us to summarize their behavior using a flexible class of models from epidemiology: the susceptible–infected–susceptible (SIS) model.[7] As in the SIS model, we partition agents into two categories: conventional asset holders, C_t, and innovation holders, I_t. The net mass of agents flowing through the innovation-holder group is given by

$$I_{t+1} - I_t = \psi I_t C_t - \xi_t I_t. \tag{3}$$

The product of the masses of I_t and C_t yields the mass of conventional asset holders who identify an innovation holder when they search the portfolios of unsophisticated agents. The parameter ψ is the probability of believing that the return to the innovation exceeds the return to the conventional asset, conditional on identifying an innovation holder. We assume that this conditional probability is fixed. The exogenous variable, ξ_t, is the probability of exiting, which contains a fixed component and an exogenous shock: $\xi_t = \xi + \nu_t$.

The parameter ξ is the product of the probability of being able to exit at a given point in time, which depends on the innovation's liquidity, and the probability of believing that the conventional asset is generating a higher return. The shock, ν_t, captures news events, which may change beliefs for all agents about the returns of the assets, and liquidity. We will assume that $\nu_t \sim \text{iid}(0, \sigma^2)$.

The corresponding flow into the group holding the conventional asset is given as

$$C_{t+1} - C_t = \xi_t I_t - \psi I_t C_t. \tag{4}$$

Notice that the flows are reversed. Here, we have a mass of unsophisticated agents, $\xi_t I_t$, who wanted to sell the financial innovation and were able to do so. We also have a mass of unsophisticated agents, $\psi I_t C_t$, replacing the conventional asset with the innovation.

Finally, it is important to observe that conversions from the innovation back to the conventional asset do not depend on the frequency of contact between conventional asset holders and innovation holders. Rather, they are generated by the aforementioned exogenous factors exclusively. Furthermore, these events do not permanently influence future choices. If an agent returns to the conventional asset, she will disregard the news events that lead her to that choice, and will instead rely on contemporaneous news events, as well as industry-level memory, summarized by the share of assets in the innovation.

[7]In epidemiological models, agents start out susceptible to an infectious disease. Contact with other infected agents transmits the disease. Over time, agents recover and return to the susceptibles category. In the susceptible–infected–resistant (SIR) version of the model, agents permanently recover from the disease and exit the pool of susceptibles.

2.2 *Sophisticated agents*

We will illustrate the role of sophisticated agents by considering the case where there is only one sophisticated agent ($N = 1$) who faces no competition. The only choice the sophisticated agent must make is whether or not to develop the financial innovation and introduce it to her clients.[8] This choice will pin down the initial conditions for the system generated by the choices of unsophisticated agents.

If the sophisticated agent does not create the innovation, she will instead sell one unit of the conventional asset for a fee, P. The agent will discount the future returns at rate β. If the agent chooses to offer clients the conventional asset at $t = 0$, her discounted flows of profit are

$$\pi_C = \sum_{t=0}^{\infty} \beta^t P C_0 = \frac{P C_0}{1 - \beta}. \tag{5}$$

In contrast, if the sophisticated agent develops a financial innovation at cost c and then offers it to clients for a fee, P', she will receive a mass-weighted return from the two assets. The weights will depend explicitly on the evolution of the mass of the financial innovation holders. This yields the following expected return from developing an innovation:

$$\pi_I = -c + \sum_{t=0}^{\infty} \beta^t P' I_t + \sum_{t=0}^{\infty} \beta^t P \underbrace{(C_0 - I_t)}_{C_t} \tag{6}$$

$$\rightarrow \quad \pi_I = -c + \frac{P C_0}{1 - \beta} + \sum_{t=0}^{\infty} \beta^t I_t (P' - P). \tag{7}$$

We may write the difference in profits, $\hat{\pi}$, as

$$\hat{\pi} = \pi_I - \pi_C = -c + \sum_{t=0}^{\infty} \beta^t I_t (P' - P). \tag{8}$$

As equation (8) indicates, discounted profits from constructing an innovation depend explicitly on the growth of the innovation-holder group. Since the sophisticated agent is a profit-maximizer, she will introduce the innovation to her clients if $\hat{\pi} > 0$, setting $C_0 = \mu - I_0$, where μ is her mass of clients. These initial conditions, combined with the system that describes the behavior of unsophisticated agents, determines whether the innovation will persist or disappear in equilibrium.

A stochastic steady-state equilibrium satisfies the following three conditions: (i) zero net movement between the conventional asset and the innovation; (ii) no aggregate shocks or shocks to parameter values; and (iii) subgame perfect Nash equilibrium among the sophisticated agents. The first condition is satisfied when $C_{t+1} - C_t = 0$ and $I_{t+1} - I_t = 0$. The second condition is satisfied when $\xi_t = \xi$. And the third condition is

---

[8]Once introduced, the innovation may spread to other agents who are not the sophisticated agent's clients. This is described in more detail in Section 2.1.

satisfied trivially when $N = 1$ and the sophisticated agent maximizes profits. We will explore the $N = 2$ case later in the paper.

Let us first consider the set of equilibria where at least one sophisticated agent issues the innovation. Here, equilibrium condition (i) implies, for the net flow equation for innovation holders,

$$0 = I_{t+1} - I_t = \psi C_t I_t - \xi_t I_t. \tag{9}$$

We may rearrange this equation as

$$\rightarrow \quad \frac{1}{C_t} = \frac{\psi}{\xi_t}. \tag{10}$$

The right hand side of this threshold condition is referred to as $\mathbb{R}_0$ or the basic reproductive number (BRN) in the epidemiology literature. We will adopt the same terminology for this model. Here, $\mathbb{R}_0$ can be interpreted as the average number of conventional asset holders who switch to the innovation after one unsophisticated agent adopts it. This is a measure of the strength of the contagiousness of the innovation.

Among the equilibria where the innovation is introduced, there are two possible outcomes for unsophisticated agents. In the first, the innovation disappears entirely: $(C^*, I^*) = (C_0, 0)$. In the second, the innovation persists: $(C^*, I^*) = (1/\mathbb{R}_0, C_0 - 1/\mathbb{R}_0)$. Additionally, in the latter equilibrium, the size of the group that holds the innovation depends critically on the value of $\mathbb{R}_0$, the measure of contagion within the unsophisticated agent group. These derivations are shown in the Appendix and are standard for the SIS class of epidemiological models.

For our purposes, the most important property of $\mathbb{R}_0$ is that it determines which equilibrium is stable. In particular, if $\mathbb{R}_0 > 1/C_0$, then the second equilibrium—where the financial innovation persists—is stable.[9] Otherwise, the first equilibrium is stable and the financial innovation disappears.[10] Since the sophisticated agent's choice to enter affects C_0, producing the innovation may also change the equilibrium.

In the next section, we will derive several theoretical results for the $N = 1$ and $N = 2$ cases. We will also show how the model is compatible with stylized facts from the financial innovation literature, and what mechanisms the model uses to explain them.

2.3 Theoretical results

Lerner's (2006) review of financial innovations finds that large financial firms are more likely to innovate, but not more likely to obtain a patent on those financial innovations. In our first theorem, we use the model to explore a possible mechanism for the size effect. We will accomplish this by constructing an equilibrium where size predicts the choice to innovate, and will then identify the forces that generate the equilibrium.

---

[9]If an equilibrium is stable, then the system will move toward it once the innovation is introduced. If an equilibrium is unstable, then the system will move away from it—and to the stable equilibrium—unless it starts at the unstable equilibrium point and receives no perturbations.

[10]See Keeling and Rohani (2008) for a proof of the equivalent condition in epidemiology.

THEOREM 1. *Let $N = 1$, $C_0 = \mu$, and $P' > P$. There exists a client base size threshold, $\bar{\mu}$, below which a sophisticated agent will not develop a financial innovation.*

PROOF. Consider a firm that has a client base of size $\tilde{\mu} = \bar{\mu} - \varepsilon$, where $\mathbb{R}_0 < 1/\tilde{\mu}$ and $I^* = 0$. If the sophisticated agent develops and introduces the innovation, then the upper bound on per period profit differences will be $(P' - P)I_0$, since I will decline on the path toward I^*. To the contrary, if the firm has a client base of size $\tilde{\tilde{\mu}} = \hat{\mu} + \varepsilon$, such that $\mathbb{R}_0 > 1/\tilde{\tilde{\mu}}$, then $(P' - P)I_0$ will form the lower bound on the path toward an equilibrium, where $I^* > I_0$.

We have established that $(P' - P)I_0$ is the upper boundary for profits when $\mu = \tilde{\mu}$ and the lower boundary when $\mu = \tilde{\tilde{\mu}}$. This implies that $\hat{\pi} = -c + \sum_{t=0}^{\infty} \beta^t (P' - P)I_0 = -c + \frac{(P'-P)I_0}{1-\beta}$ is the upper boundary for the discounted lifetime profit difference when $\mu = \tilde{\mu}$ and the lower boundary when $\mu = \tilde{\tilde{\mu}}$. Since $I_0 = \chi\mu$, we may rewrite this as $\hat{\pi} = -c + \frac{(P'-P)\chi\mu}{1-\beta}$. If $\bar{\mu} = \frac{c(1-\beta)}{(P'-P)\chi}$, then $\hat{\pi} = 0$. If $\mu = \tilde{\tilde{\mu}}$, then $\hat{\pi} > 0$ and the agent will produce the innovation. If, instead, $\mu = \tilde{\mu}$, then $\hat{\pi} < 0$ and the agent will not produce the innovation. Thus, there exists a client base size threshold, $\bar{\mu}$, below which a sophisticated agent will not develop a financial innovation. □

The intuition behind this proof is that size can be a critical determinant of profitability, even when a financial firm has the sole right or ability to produce a financial innovation. In the absence of competition, financial firms that are sufficiently large may select their preferred equilibrium by choosing to enter. In contrast, smaller firms may desire an equilibrium where the innovation persists indefinitely, but will not be able to generate enough visibility to achieve that effect. Instead, they will collect a sequence of declining fees from their dwindling client base until the financial innovation disappears.

This result depends crucially on the SIS core of the model. The interaction of innovation-holding unsophisticated agents with non-innovation-holding unsophisticated agents is what allows a new financial product to gain visibility and achieve permanence. Without permanence, it will not be profitable for the firm to pay the fixed cost to enter. This channel for the size effect is distinct from what might be derived in a model with no interactions between unsophisticated agents, since it does not rely on any assumptions about financing constraints, advertising budgets, or internal expertise. It also does not require a firm to have the capacity to expend more resources, but instead emphasizes the importance of client interactions across firms in a dynamic setting. Even in the absence of the aforementioned causes, the contagion effect will ensure the profitability of innovation if a firm is sufficiently large, and, thus, innovations will tend to be produced by larger firms, generating the size effect.

Next, we will consider why firm size has no impact on patenting, even though it does on innovation (Lerner (2006), Herrera and Schroth (2011)). We will assume that there are now two sophisticated agents of identical size making the entry decision.

THEOREM 2. *Let $N = 2$, $\mu_1 = \mu_2 = \mu$, and $P' > P$. For a given c, P, and P', there exists a subgame perfect Nash equilibrium (SPNE) where both sophisticated agents develop the innovation and the first mover does not patent.*

PROOF. There are four possible payoffs for the first mover: $\pi_{I,I}$, $\pi_{I,C}$, $\pi_{C,I}$, and $\pi_{C,C}$, where the first subscript is sophisticated agent 1's move and the second subscript is sophisticated agent 2's move. For this exercise, we consider only the set of equilibria where producing the innovation dominates producing the conventional asset. This rules out the equilibria that generate payoffs of $\pi_{C,I}$ and $\pi_{C,C}$. Since the first mover knows that producing the conventional asset is a dominated choice, she knows that the sophisticated agent who moves second will always choose to innovate. Thus, she must choose between patenting ($\pi_{I,C}$) and not patenting ($\pi_{I,I}$).

Now, assume $\mathbb{R}_0 < 1/\mu$, $\mathbb{R}_0 > 1/(2\mu)$, $C_0 = 2\mu$, and $\tilde{C}_0 = \mu$. If agent 1 excludes agent 2 with a patent, then agent 1 receives $\pi_{I,I} = -c + \sum_{t=0}^{\infty} \beta^t P(\mu - \tilde{I}_t) + \sum_{t=0}^{\infty} \beta^t P' \tilde{I}_t$. If agent 1 does not patent, but instead allows agent 2 to produce the innovation also, then agent 1 receives $\pi_{I,C} = -c + \sum_{t=0}^{\infty} \beta^t P(\mu - I_t/2) + \sum_{t=0}^{\infty} \beta^t P' I_t/2$. The discounted profit difference can be simplified to the expression $\pi_{I,I} - \pi_{I,C} = \sum_{t=0}^{\infty} \beta^t (P' - P)(I_t/2 - \tilde{I}_t)$. At $t = 0$, we know that $\tilde{I}_t = \mu$ and $\tilde{I}_t/2 = \mu$. Thus, there is no difference in profit flows in the first period. However, in subsequent periods, $\tilde{I}_t \to \tilde{I}^* = 0$ and $I_t/2 \to I^*/2 > 0$. This implies that not patenting will yield a weakly positive profit flow in each period on the equilibrium path: $\pi_{I,I} - \pi_{I,C} > 0$. Thus, there exists an equilibrium where both sophisticated agents develop the innovation and the first mover does not patent.   □

The intuition for this proof relies on the gains that come from moving to an equilibrium in which a positive amount of the innovation is held. If a sophisticated agent develops the innovation alone, then it may only capture a declining sequence of fees on the path to the no-innovation equilibrium. Issuing simultaneously allows them to sell more of the innovation, even though they cannot capture the entire market. It is important to note that we were able to derive this result without assuming that the costs of development or issuance were split; however, incorporating a cost split would provide a further rationale for simultaneous issuance or explicit syndication.

This argument can be extended to the case where patenting does not change the equilibrium category. Below, we consider two sophisticated agents, each of whom is sufficiently large to select an equilibrium independently. Even if these agents face no cost reduction for simultaneous issuance, they will prefer a joint release if convergence to equilibrium is sufficiently fast when both client bases are introduced to the innovation.

THEOREM 3. *Let $N = 2$, $\mu_1 = \mu_2 = \mu$, and $P' > P$. For a sufficiently short time to convergence, τ, there exists an equilibrium where both sophisticated agents develop the innovation and the first mover does not patent.*

PROOF. Assume that each sophisticated agent has a mass of clients of size μ, where $\mathbb{R}_0 > 1/\mu$, and also that equilibrium is reached within τ periods. For notational simplicity, let $\tilde{c} = -c + \sum_{t=0}^{\infty} \beta^t P\mu$. If the first agent enters alone, issues the innovation, and excludes the second agent with a patent, then he/she will receive no more than $\bar{\pi}_{I,C} = \tilde{c} + (P' - P)(\mu - \frac{1}{\mathbb{R}_0})/(1 - \beta) > \pi_{I,C}$. This assumes that the agent skips the transition path and receives the new equilibrium payoff in all periods. If, instead, the first mover does not

patent and both agents develop the innovation, then each agent's discounted payoff will be bounded from below by

$$\underline{\pi}_{I,I} = \tilde{c} + (P' - P) \sum_{t=0}^{\tau-1} \beta^t \frac{\chi\mu}{2} + (P' - P)\beta^\tau \sum_{t=0}^{\infty} \beta^t \left(\mu - \frac{1}{2\mathbb{R}_0} \right) \qquad (11)$$

$$\rightarrow \quad \underline{\pi}_{I,I} = \tilde{c} + (P' - P)\frac{\chi\mu}{2}\frac{1 - \beta^\tau}{1 - \beta} + (P' - P)\beta^\tau \frac{\mu - \dfrac{1}{2\mathbb{R}_0}}{1 - \beta}. \qquad (12)$$

The first mover will choose not to patent if the following condition holds:

$$\tilde{c} + (P' - P)\frac{\chi\mu}{2}\frac{1 - \beta^\tau}{1 - \beta} + (P' - P)\beta^\tau \frac{\mu - \dfrac{1}{2\mathbb{R}_0}}{1 - \beta} > \tilde{c} + \frac{(P' - P)\left(\mu - \dfrac{1}{\mathbb{R}_0}\right)}{1 - \beta} \qquad (13)$$

$$\rightarrow \quad \tau < \ln \left\{ \frac{\mu(\chi - 1) + \dfrac{1}{\mathbb{R}_0}}{\mu\left(\dfrac{\chi}{2} - 1\right) + \dfrac{1}{2\mathbb{R}_0}} \right\} \Big/ \ln \beta. \qquad (14)$$

Thus, if the amount of time spent on the convergence path, τ, is sufficiently small, then there exists an equilibrium where both sophisticated agents develop the innovation and the first mover does not patent. □

There are two parts to the intuition behind the proof. First, issuing the innovation jointly may improve profits, even when those profits are divided across two firms, and even when both of those firms are capable of achieving an equilibrium with a positive amount of the financial innovation independently. The second part of the intuition is that rapid convergence to the new equilibrium will ensure that these profits are not too distant in the future, and, therefore, not too heavily discounted. Sufficiently rapid convergence—within τ periods—will ensure that the choice to enter jointly dominates the choice to enter and then patent, regardless of what happens on the transition path.

Theorems 2 and 3 demonstrate different channels through which the model is compatible with the same stylized fact: even though financial innovations tend to improve the profitability of their originators, financial firms typically do not patent or otherwise attempt to protect innovations they develop. To the contrary, they often coordinate syndicated releases of financial innovations. Theorem 2 considers the importance of coordination for increasing visibility: under certain circumstances, a single financial firm will not be able to introduce an innovation in a way that causes it to persist; however, two or more financial firms may be able to generate enough visibility to cause it to persist in equilibrium. In Theorem 3, patenting is eschewed to pull the distant equilibrium outcome of persistence into the present. Moving these otherwise heavily discounted profit flows into the near future makes it more profitable for firms to issue the innovation jointly, even if it means that profits must be split.

Finally, we consider Lerner's (2006) finding that innovators tend to be less profitable ex ante than firms that do not innovate. This holds, even though innovations themselves

tend to make the firm more profitable ex post. Herrera and Schroth (2011) also demonstrate that innovation producers—as opposed to imitators—tend to collect higher fees and issue larger quantities, improving ex post profitability. Why, then, do innovators tend to be less profitable ex ante, even though the act of innovating makes them more profitable ex post? This seems counterintuitive, since a serial innovator would experience constant improvements in profitability over competitors, but would have below average profits in expectation.

One seemingly plausible answer is that less profitable firms innovate because they have worse outside options. That is, P differs across firms, and the firms with the lowest values of P will adopt the innovation, which offers $P' > P$. Thus, if we look at a cross section of firms at a point in time, we will find that those who choose to innovate were less profitable in the past, but improved their profitability by adopting the innovation. As we will show below, this turns out to be a stronger assumption than we need to generate the result within this framework. We will demonstrate how this difference can arise and persist, even when fees and innovations are identical.

THEOREM 4. *Let $N = 2$ and let $\mu_1 = \mu_2 = \mu$. If innovation-inducing frictions are sufficiently abundant and if innovations always disappear in equilibrium, then firms who innovate will be less profitable than competitors in the period prior to the innovation's development, but will experience an improvement in profitability immediately afterward.*

PROOF. Assume that $N = 2$ in a market segment where all financial innovations disappear in equilibrium.[11] Let each innovation's fee be P'. Assume that sophisticated agent 1 developed an innovation $\hat{t}$ periods prior. Assume also that sophisticated agent 2 developed the same innovation $\hat{t} - 1$ periods prior. Agent 1's discounted profits from continuing to issue the "old" innovation are given as

$$\pi_O = \frac{P\mu}{1 - \beta} + \sum_{t=\hat{t}}^{\infty} \beta^{\hat{t}-t} I_t (P' - P). \tag{15}$$

If agent 1 instead develops the "new" innovation, she will receive

$$\pi_N = -c + \frac{P\mu}{1 - \beta} + \sum_{t=0}^{\infty} \beta^t I_t (P' - P). \tag{16}$$

We may write the difference in profits as

$$\hat{\pi} = \pi_N - \pi_O = -c + \sum_{t=0}^{\infty} \beta^t (I_t - I_{t+\hat{t}})(P' - P). \tag{17}$$

---

[11]Some classes of innovations, such as those that exploit a technological opportunity or those that address a particular problem related to a business cycle stage, may be inherently transitory. This contrasts with innovations that result from permanent changes in taxes or regulations.

Now, pick $\hat{t}$, such that

$$c = \sum_{t=0}^{\infty} \beta^t (I_t - I_{t+\hat{t}})(P' - P). \tag{18}$$

Agent 1 will be indifferent between adopting the innovation and not. We will assume that agents on the margin adopt the innovation. For agent 2, the difference in profits is

$$\hat{\hat{\pi}} = -c + \sum_{t=0}^{\infty} \beta^t (I_t - I_{t+\hat{t}-1})(P' - P). \tag{19}$$

For each period, t, prior to reaching equilibrium, we know that $I_{t+\hat{t}-1} > I_{t+\hat{t}}$. Thus, we know that agent 2 will not adopt the innovation in period t:

$$c > \sum_{t=0}^{\infty} \beta^t (I_t - I_{t+\hat{t}-1})(P' - P). \tag{20}$$

In the period after the innovation is adopted, per period profits for agent 1 will be $(P' - P)I_0 + P\mu$, which exceeds $(P' - P)I_{\hat{t}-1} + P\mu$—agent 1's profits in the period prior to adoption—since $P'I_0 \geq P'I_1 \geq \cdots \geq P'I^*$. Furthermore, $(P' - P)I_{\hat{t}} + P\mu$ is lower than $(P' - P)I_{\hat{t}-1} + P\mu$—agent 2's profits in the period prior to adoption. Thus, agent 1 innovates first and improves her profitability after innovating, but generates a lower per period profit in the periods immediately prior to the innovation's development. □

This proof demonstrates how innovators can be less profitable than noninnovators in the immediate past, but more profitable in the future than they were in previous periods without making any strong assumptions. As long as innovation-inducing frictions are abundant, innovation development is not simultaneous, and innovations disappear in equilibrium, both stylized facts will hold, even if firms are otherwise identical and financial innovations are identical. These results arise from the decay in profitability of nonpermanent innovations.

The SIS model's role is only implicit in Theorem 4: we have assumed that innovations disappear in equilibrium, but we have not explicitly stated why this is the case. This is equivalent to imposing restrictions on either $\mathbb{R}_0$ or the initial mass of unsophisticated agents who adopt the innovation. This could be an accurate description of innovations that are constructed in response to frictions generated by the business cycle or to specific technological opportunities. These frictions subside over time, which keeps $\mathbb{R}_0$ low, preventing the spread of the innovation by setting it on a monotonic path toward the no-innovation equilibrium.

3. Empirical applications

In this section, we will discuss two applications of the model that involve estimating $\mathbb{R}_0$. In both cases, we will consider innovations that are successful in the sense that they are still actively being issued; however, one innovation grows in share throughout the

entire sample period while the other declines after the financial crisis. For each case, we will (i) estimate a structural equation from the model, (ii) use the point estimates to compute $\mathbb{R}_0$, (iii) use the delta method to compute the standard error of $\mathbb{R}_0$, and then (iv) perform a hypothesis test to determine whether or not the financial innovation will spread and persist in equilibrium. We will also consider recursive estimates of $\mathbb{R}_0$ that allow for time variation in the parameter values, and will use those estimates to make predictions about the permanence of innovations. We will conclude by discussing the quality of the structural model's fit and the robustness of threshold condition rejections.

3.1 *Asset-backed securities*

In the first application, we consider the introduction of private, asset-backed securities (ABS). The term "asset-backed securities" typically refers to bonds with payment flows derived from an underlying pool of assets, such as consumer loans or mortgage debt. ABS, however, can refer to debt that is securitized by either private financial firms or government-sponsored agencies (GSEs). Here, we avoid debt that is securitized by GSEs and instead focus on privately securitized debt.

So as to estimate the model's structural parameters, we must first identify a mapping from the data to the model's variables, C and I. I assume that the strategic decisions described in the theoretical section were made in periods prior to the first available observation, since the innovation has already been introduced. Furthermore, for simplicity, I assume that all debt can potentially be securitized and I normalize the total stock of debt to 1. This simplifies the model by allowing us to compute the mass of securitized debt, I, as

$$I_t = 1 - C_t. \tag{21}$$

I will use quarterly Flow of Funds (FoF) data from the Federal Reserve Board of Governors to construct C. To accomplish this, three series were used: (i) the total level of household liabilities; (ii) the total value of financial assets at GSEs; and (iii) the total value of financial assets for issuers of asset-backed securities. All three series provide values for the aggregate economy and span the period from 1983:Q4 to 2014:Q4. I selected 1983 as the start date because that is when private ABS are first introduced in the FoF data.[12]

From these three series, I constructed the values for C, the proportion of households liabilities that were not being held at government-sponsored enterprises (securitized or not) and were not already securitized, as shown by

$$C_t = \frac{\text{total liabilities}_t - \text{GSE financial assets}_t - \text{securitized debt}_t}{\text{total liabilities}_t - \text{GSE financial assets}_t}. \tag{22}$$

---

[12]All data for the ABS exercise are taken from the Federal Reserve Board's Flow of Funds tables. All series fall under Z.1. Statistical Release: Financial Accounts of the United States. The unique identifiers for the series used are (i) the total level of household liabilities (Z1/Z1/FL154190005.Q), (ii) the total value of financial assets at GSEs (Z1/Z1/FL404090005.Q), and (iii) the total value of financial assets for issuers of asset-backed securities (Z1/Z1/FL674090005.Q).
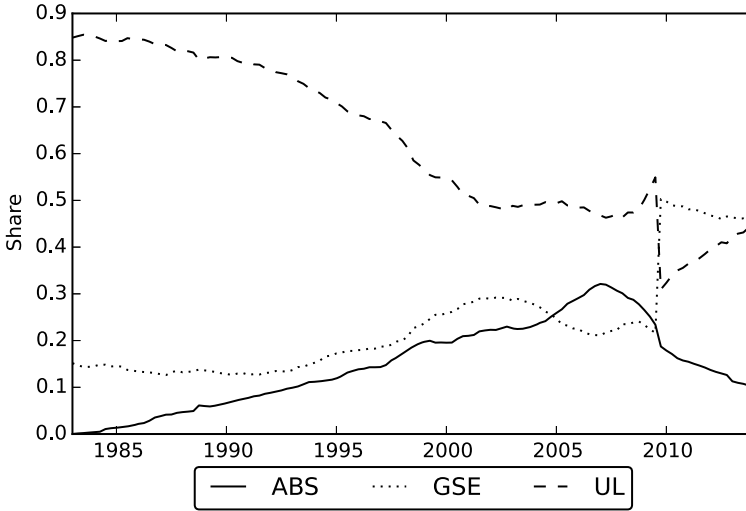
FIGURE 1. The component shares of asset-backed securities (ABS), government-sponsored enterprise assets (GSE), and unsecuritized liabilities (UL). The shares are constructed using total dollar values of each fund type, taken from the Federal Reserve Board's Flow of Funds data. UL are defined as total liabilities (TL) minus ABS minus GSE.

These liabilities could be considered the "conventional asset" for the institutions that held them. Now that we have constructed C—and also I—we estimate a transformation of equation (3):

$$\frac{I_{t+1} - I_t}{I_t} = \psi \frac{I_t C_t}{I_t} - (\xi + \nu_t)\frac{I_t}{I_t} \tag{23}$$

$$\rightarrow \quad \frac{I_{t+1} - I_t}{I_t} = \psi C_t - \xi + \nu_t. \tag{24}$$

Figure 1 shows a plot of the normalized series used to construct C and I. Note that unsecuritized liabilities (UL) are C and I is $1 - C$. Also, all component series sum to 1 in each period, and ψ and ξ retain their structural interpretations, even though we identify them using equation (24), which does not readily provide a clear interpretation. The jump in GSE share in 2010 marks the start of the government's conservatorship of Freddie Mac and Fannie Mae and the conversion of mortgage guarantees to on-balance-sheet items. In contrast, the private ABS share grows until the financial crisis, but then declines thereafter to the end of the sample.

Furthermore, note that imposing (21) transforms (24) into a univariate equation. I estimate (24) on quarterly data for the 1983:Q4–2014:Q4 period. The results are given in column 1 of Table 1. The parameter estimates imply that $\mathbb{R}_0 > \frac{1}{C_0}$, which predicts that private ABS will grow in share and persist in a stable equilibrium. Table 1 also provides $\mathbb{R}_0$ and its standard error. A separate set of estimates is given in column 2 of Table 1, which excludes the financial crisis and the changes that immediately preceded it, using data only during the 1983:Q4–2006:Q1 period.

Table 1. ABS and ETF regression results.

	ABS (1)	ABS (2)	ETF
ψ	0.236***	0.202***	1.478***
	(0.037)	(0.039)	(0.282)
ξ	0.164***	0.129***	1.343***
	(0.030)	(0.033)	(0.272)
R_0	1.437***	1.565***	1.100***
	(0.243)	(0.376)	(0.024)
N	116	85	76
R^2	0.271	0.233	0.258
F-stat.	27.54***	26.56***	39.61***

Note: This table displays the results from three OLS regressions and three parameter transformations. In all cases, the regression takes the form of equation (24). That is, the percentage change in the share of the financial innovation is regressed on a constant and the lagged share of conventional assets. The coefficients, ψ and ξ, as well as R^2 and the F test results are associated with the two original regressions. In addition to this, R_0 is reported, which is computed as ψ/ξ. ABS (1) and ETF are estimated using all available data. ABS (2) uses a subset of the data that ends prior to 2007, excluding the financial crisis and period that follows. Newey–West standard errors are reported for ξ and ψ. Standard errors for R_0 are constructed using the delta method. Section 3.3. provides a more detailed analysis of the model fit using the Akaike information criterion (AIC). Significance levels are given as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

The variable R_0 was computed as ψ/ξ, yielding 1.437 for the full sample and 1.565 when the financial crisis is excluded. The standard error of R_0 was then computed as follows: let Σ represent the covariance matrix and let $D = [1/\xi, -\psi/\xi^2]$ be the row vector of partial derivatives of R_0. Then, by the delta method, the variance of R_0 is $\mathrm{var}(R_0) = D\Sigma D'$. In this case, $\mathrm{SE}(R_0) = 0.243$ for the full sample and 0.376 for the subsample that excludes the financial crisis. Now recall that the threshold condition is violated if

$$R_0 > \frac{1}{C_0}. \tag{25}$$

If we perform a hypothesis test using 1 as the initial share of the conventional asset (i.e., $C_0 = 1$), then we may reject the threshold condition in both cases.[13] That is, even if we include the financial crisis in the sample—during which the ABS share dropped substantially—the model still predicts that private ABS will persist in equilibrium.

In addition to providing ordinary least squares (OLS) estimates for the entire sample period, we also plot recursive estimates for R_0 for each quarter from 1988:Q4 to 2014:Q4 in Figure 2 (i.e., starting 5 years after they were introduced). We construct each quarter's estimate of R_0 using Kalman-filtered, smoothed estimates of ξ_t and ψ.[14] We then compute the standard error of R_0 using the delta method. This approach allows for variation in ξ_t, which could be caused by changes in liquidity or returns.

---

[13]We might instead use the final sample period value of C_t as C_0 and reinterpret C_0 as the initial observation in the forecast period. However, for the purposes of this paper, we will focus on the initial observation in the sample period.

[14]The filtering and smoothing is done using R's dynamic linear models package. See Petris (2009) for a complete description of the implementation.
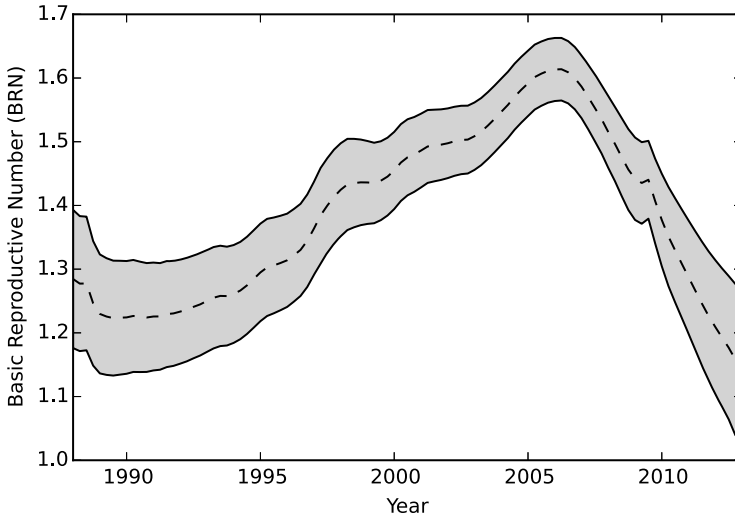
FIGURE 2. The basic reproductive number $\mathbb{R}_0$ for ABS share and its associated, 95% confidence band, updated quarterly from 1988:Q4 to 2014:Q4. We construct each quarter's estimate of $\mathbb{R}_0$ using Kalman-filtered, smoothed estimates of ξ_t and ψ. We then compute the standard error of $\mathbb{R}_0$ using the delta method.

The recursive estimates suggest that the rejection of the threshold condition was statistically significant in 1988, five years after the introduction of private ABS. This predicts the expansion and equilibrium persistence of private ABS. The significance of the rejection started to decline prior to the financial crisis and has continued to decline since the crisis ended. This effect is substantially larger than what we saw in the OLS estimates, since the recursive estimates allow for reduced dependence on older observations. Importantly, however, $\mathbb{R}_0$ remains statistically significantly different from 1, even in the final period, 2014:Q4.

3.2 *Exchange-traded funds*

Exchange-traded funds (ETFs) are a type of security that tracks the value of some underlying basket of stocks, bonds, or commodities. Unlike mutual funds, ETFs experience price changes during trading hours, are traded throughout the day, and are not subject to a minimum holding period requirement.

ETFs were introduced in the United States in 1993, and steadily grew in popularity thereafter. In contrast to ABS, information was publicly available about ETFs at their release. However, the steady-state popularity and relative fee size of ETFs was unknown initially, and the returns history was unknown to investors, even though it could be constructed in some cases.

I used four quarterly Flow of Funds (FoF) series from the Federal Reserve Board of Governors (BoG) to estimate ψ, ξ, and $\mathbb{R}_0$: (i) closed-end funds, (ii) exchange-traded
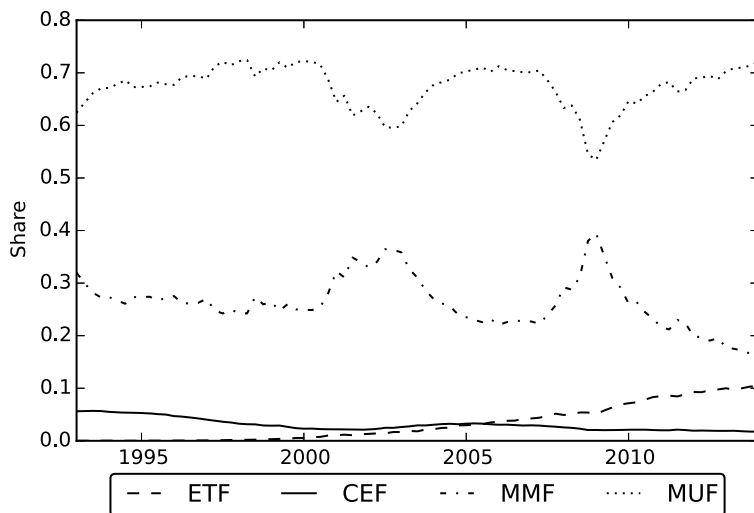
FIGURE 3. The component shares of ETFs and three substitute products: (i) closed-end funds (CEFs), (ii) money-market mutual funds (MMFs), and (iii) mutual funds (MUFs). The shares were constructed using the total dollar values of each fund type, taken from the Federal Reserve Board's Flow of Funds data.

funds, (iii) money market mutual funds, and (iv) mutual funds.[15] Each series provides the total financial assets for all funds of that type at a quarterly frequency. All series were obtained for the 1993:Q4–2014:Q4 period, since 1993 was the first year in which ETFs were available in the United States. I computed the total amount of financial assets susceptible to ETF conversion by adding up the series values for each of the substitute fund types: (i) money market mutual funds, (ii) mutual funds, and (iii) closed-end funds. All three fund types were predecessors to ETFs and also likely substitute investment vehicles. Figure 3 shows a plot of the share associated with each fund type. Note that the share for ETFs is I and $C = 1 - I$. The ETF share has been growing steadily since they were introduced and was not affected by the financial crisis. In contrast, the mutual fund share tended to drop and the money market mutual fund share tended to rise during recessions in the sample period.

I then used the total value of financial assets in the funds to compute C and I: C was defined as the fraction of asset value that was concentrated in non-ETFs; I was defined as the fraction of total asset value concentrated in ETFs. With these definitions in place, I recovered an estimate of ψ and ξ, and then constructed the basic reproductive ratio and its standard error using the method outlined in Section 3.1. I found that $\mathbb{R}_0 = 1.1$ with a standard error of 0.024. If we use 1 as the initial share of C_0, then the threshold

---

[15]All data for the ETF exercise are taken from the Federal Reserve Board's Flow of Funds tables. All series fall under the Z.1. Statistical Release: Financial Accounts of the United States. The unique identifiers for the series used are (i) closed-end funds (Z1/Z1/FL554090005.Q), (ii) exchange-traded funds (Z1/Z1/FL564090005.Q), (iii) money market mutual funds (Z1/Z1/FL634090005.Q), and (iv) mutual funds (Z1/Z1/FL654090005.Q).
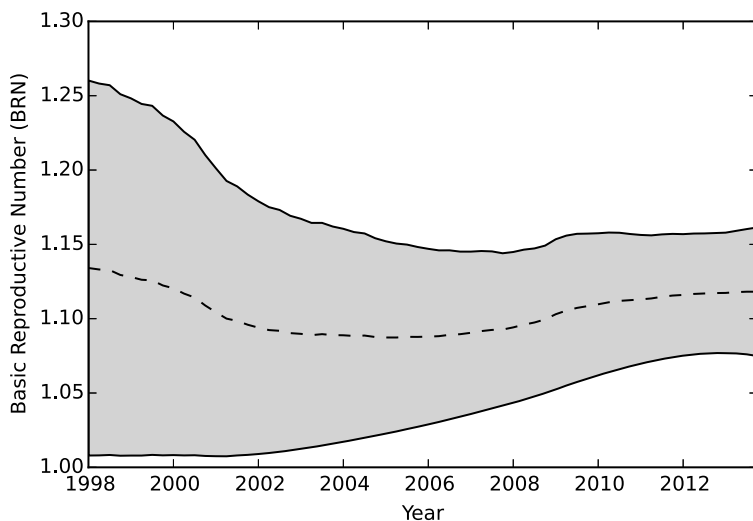
FIGURE 4. The recursive estimates of the basic reproductive number $\mathbb{R}_0$ for ETF share and its associated standard error. Estimates begin five years after ETFs were introduced and span the period from 1998:Q4 to 2014:Q4. We construct each quarter's estimate of $\mathbb{R}_0$ using Kalman-filtered, smoothed estimates of ξ_t and ψ using the approach described in the previous subsection.

condition is rejected, predicting that ETFs will persist in equilibrium. Column 3 of Table 1 provides the results for this exercise.

The model suggests that some of the parameters, such as ξ_t, may evolve over time, but we assume that these changes are independent and identically distributed (iid) for simplicity. In practice, a change in the innovation's liquidity or return could be persistent and might shift ξ_t and, subsequently, $\mathbb{R}_0$. The underlying friction size, ζ, may also change if we consider a sufficiently long horizon. For these reasons, we repeat the exercise from the previous subsection and plot recursive estimates for $\mathbb{R}_0$ for each year from 1998:Q1 to 2014:Q4 in Figure 4. We construct each quarter's estimate of $\mathbb{R}_0$ using Kalman-filtered, smoothed estimates of ξ_t and ψ using the approach described in the previous subsection.

The recursive estimate of $\mathbb{R}_0$ is statistically significantly greater than 1 as early as 1998, but declines until 2004 and then rises slightly thereafter. The precision of the estimate improves throughout the period, which makes the rejection of the threshold condition more statistically significant, even as the estimates initially drop. Unlike ABS, the value of $\mathbb{R}_0$ for ETFs does not drop substantially during or after the financial crisis.

3.3 *Model fit*

To assess the structural model's fit, we compare it to augmented versions that include one, two, and three lags of the independent variable. We conduct the comparison using Akaike's (1974) information criterion and Burnham and Anderson's (2002) method for measuring the relative probability that a model minimizes information loss. For ABS, the AIC suggests that the specification with three lags is the best fit; however, the difference across models is negligible. Applying Burnham and Anderson (2002) to the AIC, the

original specification is 0.307 times as probable to minimize the information loss as the version with three lags. The models with one and two additional lags are 0.769 times and 0.327 times as probable to minimize the information loss as the model with three lags.

For ETFs, the model with one additional lag minimizes the AIC; however, the original, structural model performs well: it is 0.961 times as probable to minimize the information loss. Adding two and three additional lags lowers the relative probability to 0.688 and 0.478, respectively.

Overall, the structural models for both ETFs and ABS perform well relative to augmented versions that include additional lags. Although, neither minimizes the information loss, the gains from adding lags are small. For this reason, we retain the original, structural form of the model, which provides a clearer interpretation of the model parameters.

3.4 *Robustness of rejections*

In many applications, it will be unclear whether it is appropriate to place certain assets in the C category. For example, there may be legal obstacles that prevent some conventional asset holders from selling these assets and adopting the innovation. If the data are not sufficiently disaggregated, then it will not be possible to define C in such a way that excludes funds that cannot be converted to the innovation. Thus, if we define the conventional asset group as $1 - I_t$, then we may be overstating the amount of funds that may be invested in the innovation. Recall that we use the following regression to recover ψ and ξ, so that $\mathbb{R}_0$ can be constructed:

$$\frac{I_{t+1} - I_t}{I_t} = -\xi + \psi \tilde{C}_t + \nu_t. \tag{26}$$

The mass of funds that can be converted to the innovation may be mismeasured. In particular, there may be a third group, R_t, that cannot hold the innovation,[16] but also cannot be separately identified from C_t:

$$\tilde{C}_t = C_t + R_t. \tag{27}$$

This will lead to attenuation bias in ψ, which will also attenuate $\mathbb{R}_0 = \psi/\xi$. Thus, omitting classes of assets that are not truly substitutes of the financial innovation will tend to understate the contagiousness of a security. This will decrease the frequency with which the threshold condition is rejected correctly, but will not cause more frequent, incorrect rejections of the threshold condition. This suggests that a high estimate of $\mathbb{R}_0$ may provide more meaningful predictive content than a low estimate of $\mathbb{R}_0$.

4. Conclusion

I build a general model of financial innovation that is used to explain the process through which financial product innovations are developed and spread. The approach

---

[16]Alternatively, we might think that some institutions in the C_t group have already held the innovation, but switched back to the conventional asset permanently after a bad experience.

focuses on the inherently social behavior that emerges when little information is available about a new financial innovation. Instead of analyzing historical returns, potential early adopters examine the behavior and outcomes of the current cohort of competitors to infer the properties of an innovation.

In the theoretical model, I combine the social behavior described above with strategic interactions between innovation issuers. The resulting model has two categories of equilibria, only one of which can be stable: in one, the innovation persists; in the other, it disappears. The choice of equilibria depends on two things: (i) $\mathbb{R}_0$, the measure of the innovation's contagiousness among unsophisticated agents, and (ii) the strategic choices of potential innovation issuers. The model is used to explain why innovators tend to be larger and less profitable prior to issuance, and why issuers not only fail to protect innovations from imitators, but often invite imitation through syndicated releases. I show that this holds true under general conditions, even when there is no cost reduction from cooperation.

In addition to the theoretical findings, this paper develops $\mathbb{R}_0$ as a measure of contagion for financial innovations. I show how $\mathbb{R}_0$, and the related threshold condition, can be used to predict whether a financial innovation will grow in market share and persist in equilibrium. This can be done using only a short time series of aggregate data, which is important, given the scarcity of data available on new financial innovations. I also evaluate the econometric robustness of the techniques, as well as the robustness of threshold condition checks to model selection.

Finally, I evaluate the approach in two separate empirical applications. I show how the violations of the threshold condition could be used to make predictions about the take-off of asset-backed securities (ABS) and exchange traded funds (ETFs).

<div align="center">Appendix</div>

<div align="center">A.1 Equilibrium derivations</div>

Below, we characterize the existence of equilibria for the unsophisticated agent component of the model. In the first equilibrium, the innovation disappears; in the second, it persists. These derivations are identical to those for the SIS model from epidemiology (Keeling and Rohani (2008)). We start by noting that an equilibrium is defined by the absence of movement across categories:

$$0 = g + \xi I^* - \psi I^* C^*, \tag{28}$$

$$0 = \psi I^* C^* - \xi I^*. \tag{29}$$

In the preceding equations, the superscript asterisk (*) indicates a steady-state value. Note that the second equation can be factored as

$$0 = I^* (\psi C^* - \xi). \tag{30}$$

When $g = 0$, this has two solutions: $I^* = 0$ and $C^* = \xi/\psi = 1/\mathbb{R}_0$. The first solution is the no-innovation equilibrium, since no agents hold the financial innovation in the steady

state. When we plug $I^* = 0$ into the first equation, we find that $C^* = \mu$, since $C + I = \mu$. Thus, the no-innovation equilibrium may be characterized as

$$\left(C^*, I^*\right) = (\mu, 0). \tag{31}$$

Next, we consider the second solution: $C^* = 1/\mathbb{R}_0$. From this, we get $I^* = \xi(\mathbb{R}_0 - 1)/\psi$. This gives us the following equilibrium with a positive volume of the financial innovation:

$$\left(C^*, I^*\right) = \left(\frac{1}{\mathbb{R}_0}, \frac{\xi}{\psi}(\mathbb{R}_0 - 1)\right). \tag{32}$$

A well known result in the epidemiology literature is that the second equilibrium is stable if $\mathbb{R}_0 > \frac{1}{C_0}$ and the first equilibrium (no innovation) is stable otherwise.[17] In our case, this suggests that if $\mathbb{R}_0 > \frac{1}{C_0}$ and a small amount of the financial innovation is introduced, then we will move to the equilibrium with a positive amount of the innovation, where the financial innovation grows in importance and remains in use in the steady state.

In general, the threshold condition is said to hold if the system returns to the no-innovation equilibrium after a sophisticated agent introduces some mass of the financial innovation. If, instead, the system moves toward an equilibrium with the innovation, then threshold condition is violated.

References

Akaike, H. (1974), "A new look at the statistical model identification." *IEEE Transactions on Automatic Control*, 19 (6), 716–723. [631]

Antweiler, W. and M. Frank (2004), "Is all that talk just noise? The information content of Internet message boards." *Journal of Finance*, 59 (3), 1259–1294. [614]

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992), "A theory of fads, fashion, custom, and cultural change as informational cascades." *Journal of Political Economy*, 100 (5), 992–1026. [615]

Burnham, K. and D. R. Anderson (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, Berlin. [631]

Campbell, T. (1988), *Money and Capital Markets*. Scott Foresman, Glenview, IL. [613]

Carroll, C. (2003), "Macroeconomic expectations of households and professional forecasters." *Quarterly Journal of Economics*, 118, 269–298. [615]

Cockburn, I. and R. Henderson (1998), "Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery." *Journal of Industrial Economics*, 46 (2), 157–182. [614]

Cohen, W. M. and D. A. Levinthal (1989), "Innovation and learning: The two faces of R&D." *The Economic Journal*, 99, 569–596. [614]

---

[17]This can be demonstrated by constructing the Jacobian and checking the eigenvalues. See Keeling and Rohani (2008) for a short proof.

Frame, S. and L. White (2004), "Empirical studies of financial innovation: Lots of talk, little action?" *Journal of Economic Literature*, 42 (1), 116–144. [613, 614]

Graham, B. and D. Dodd (1996), *Security Analysis: The Classic 1934 Edition*. Mcgraw-Hill, New York. [614]

Herrera, H. and E. Schroth (2011), "Advantageous innovation and imitation in the underwriting market for corporate securities." *Journal of Banking & Finance*, 35 (5), 1097–1113. [615, 621, 624]

Hong, D., H. Hong, and A. Ungureanu (2011), "An epidemiological approach to opinion and price-volume dynamics." AFA 2012 Chicago Meetings Paper. [615]

Hong, H., J. D. Kubik, and J. C. Stein (2004), "Social interactions and stock market participation." *Journal of Finance*, 59, 137–163. [614]

Hong, H., J. D. Kubik, and J. C. Stein (2005), "Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers." *Journal of Finance*, 60, 2801–2824. [614]

Ivkovic, Z. and S. Weisbenner (2007), "Information diffusion effects in individual investors common stock purchase: Cover the neighbors investment choices." *Review of Financial Studies*, 20, 1327–1357. [614]

Jaffe, A. B., M. Trajtenberg, and R. M. Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations." *Quarterly Journal of Economics*, 108 (3), 578–598. [614]

Kaustia, M. and S. Knupfer (2012), "Peer performance and stock market entry." *Journal of Financial Economics*, 104 (2), 321–338. [614]

Keeling, M. and P. Rohani (2008), *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, Princeton, NJ. [614, 620, 633, 634]

Lerner, J. (2006), "The new financial thing: The origins of financial innovations." *Journal of Financial Economics*, 79, 223–255. [614, 615, 620, 621, 623]

Miller, M. H. (1986), "Financial innovation: The last twenty years and the next." *Journal of Financial and Quantitative Analysis*, 21 (4), 459–471. [613]

Modigliani, F. and M. Miller (1958), "The cost of capital, corporation finance and the theory of investment." *American Economic Review*, 48 (3), 261–297. [616]

Mukandavire, Z., D. Smith, and G. Morris (2013), "Cholera in Haiti: Reproductive numbers and vaccination coverage estimates." *Scientific Reports*, 3, 1–8. [615]

Petris, G. (2009), "An R package for dynamic linear models." *Journal of Statistical Software*, 36, 12. [628]

Pfajfar, D. and E. Santoro (2012), "News on inflation and the epidemiology of inflation expectations." Discussion Paper 2012-048, Tilburg University, Center for Economic Research. [615]

Schroth, E. (2006), "Innovation, differentiation and the choice of an underwriter: Evidence from equity-linked securities." *Review of Financial Studies*, 19, 3. [615]

Shive, S. (2010), "An epidemic model of investor behavior." *Journal of Financial and Quantitative Analysis*, 45, 169–198. [615]

Sommer, M. and C. Carroll (2004), "Epidemiological expectations and consumption dynamics." In *Money Macro and Finance (MMF) Research Group Conference 2003*, 92. Money Macro and Finance Research Group. [615]

Tufano, P. (1989), "Financial innovation and first-mover advantages." *Journal of Financial Economics*, 25, 213–240. [615]

Tufano, P. (2003), "Financial innovation." In *Handbook of the Economics of Finance*, Vol. 1A (G. Constantinides, M. Harris, and R. Stulz, eds.), 307–336, Elsevier, Amsterdam. [614, 616]