Cascades in Automation and Technology Adoption

Julian Klix

Master Thesis, MRes in Economics,
Universitat Pompeu Fabra, May 15, 2026

Supervisor: Alexander Frug
Additional Reader: Massimo Motta

Abstract

Motivated by deautomation patterns and evidence for cascades in technology adoption, I develop a model of herding behavior in innovation. Firms compete strategically, observe private signals about the uncertain general prospect of innovating and try to infer additional information from other firms' observable actions. This can result in cascades in which private information is forgone in favor of following the prevalent action. Introducing strategic interaction between firms prompted by direct competition alters the incentives to innovate, making adoptive behavior and innovation cascades less likely in the case of substitutes and more likely in the case of complements.
Furthermore, the implications for economic policy involving public information disclosure and its role in breaking up inefficient and welfare-detrimental innovation cascades are discussed, as well as further avenues of research.

1 Introduction

"Imitation is not just the sincerest form of flattery - it's the sincerest form of learning."
George B. Shaw

Few topics have been as pervasive in the public economic discourse over the last years as automation and its potential effects on employment, wages and productivity. This ubiquity of particular technologies and innovations in the general focus appears to be a common phenomenon, emerging both on a large scale as in the case of automation, production downsizing or the adoption of e-commerce technology as well as in specific sectors such as the acquisition of CT scanners by hospitals or the adoption of the QWERTY-standard by typewriter manufacturers.
This raises the question of which factors might be able to explain this simultaneous focus on the adoption of particular technologies. It is clear that not all instances of innovation trends can be attributable to simply being superior in profitability.[1] Quite on the contrary, innovations concerning automation appear to be governed by substantial uncertainty, while production downsizing

[1]This clear superiority appears to be the driving force, for instance, behind the widespread adoption of labor division during the industrial revolution.

(Cameron, 1998) and the acquisition of CT scanners (Banta, 1980) were shown to be manifestations of a bandwagon effect. Given such an uncertain environment, the regard for others' adoption behavior and the sequential nature of adoption decisions, the theory of cascades and herding behavior appears to be one plausible approach to explain trends and fad-like behavior in technology adoption. In fact, several studies find evidence for occurrences of herding in innovation over time, from plank roads (Bikhchandani et al., 1998) to organic farming (Chatzimichael et al., 2014) to the aforementioned electronic commerce technologies (Walden and Browne, 2003).

Based on this, I develop a model in which firms either act as monopolists or compete in complements or substitutes and face a sequential decision whether to adopt an available technology in order to innovate their production process. Innovation will affect profits depending on the success of the adoption. The global success of adoption is unknown ex-ante, with each firm receiving a private signal about the innovation prospect and, in the main models, being able to observe the adoption choices of preceding firms. Since private information is only partially informative, firms also try to infer the aggregate probability of successfully innovating from observable choices. As in the standard herding models, if previous actions sufficiently coincide, firms disregard their own signal in favor of imitating the adoption behavior of their predecessors, resulting in an adoption or omission cascade.

I find that the introduction of additional payoff externalities through competition and strategic interaction adds a forward-looking component to the purely backward-looking nature of standard herding models. This results in a more complex characterization of adoption behavior with increased adoption thresholds and a decrease in innovation probability when additional adopters decrease profits (i.e., in the case of negative externalities) and vice versa, as the firms' incentives to adopt the technology change by internalizing the signaling capacity of their own actions to subsequent firms. In aggregate, that means that adoption cascades are less likely to occur, and thus, the expected number of adopters is lower the stronger the competition (in substitutes). Similarly, strategic interaction between complements increases the cascade probability of innovation and the expected number of adopters.

Examining the phenomenon of adoption patterns and their relation to other dimensions of economic activity can provide useful insights for designing economic policies and evaluating technological trends and their effects on the economy. This paper develops a basis on which to assess the prospects of governmental intervention in their ability to influence trends in technology adoption. I discuss how the possibility of a persistent innovation trend (manifesting in an adoption cascade) in particular could open an avenue for governmental intervention. To this end, providing additional or improved information about the technology could provide the means to interrupt an ongoing (potentially wasteful) adoption cascade.

The rest of the paper is structured as follows: The remainder of Section 1 gives an overview of the related literature. Sections 2 and 3 introduce the model of cascading behavior in technology adoption in its baseline form and with adoption externalities through competition respectively and sketch their solution. Section 4 analyzes the result of the models on an aggregate level and examines implications for assessing technology abandonment and the design of economic policies, while Section 5 concludes the paper by discussing potential shortcomings of the modeling approach used and proposing potential avenues of further research.

Related Literature

Imitative behavior has been a well-known phenomenon in economics, spurring a variety of explanatory approaches, from sanctions for divergent behavior (Akerlof, 1980) to conformity preferences (Bernheim, 1994). In this avenue, models of herding behavior were developed simultaneously in two seminal papers by Banerjee (1992) and Bikhchandani et al. (1992). In both models, herding behavior arises when individuals infer information from the observable actions of preceding agents that prompts them to disregard their own private signals. The result is a so-called *information cascade*, which arises as a rational consequence of a standard preference framework without the need for conformity-inducing punishments or rewards. The main difference between the two models lies within the informational environment: while the former is embedded into a continuous, potentially uninformative signal space, the latter (which will be closer to the cascades present in this model) takes place in a discrete signal space and can lead to wrong or suboptimal herding behavior.

Subsequently, a variety of avenues of research have formed that apply the initial idea of cascades as the cause of imitation to new domains. As one of the first papers to study herding in a strategic market environment, Scharfstein and Stein (1990) develops a model with investment decisions after the observation of private signals and previous actions. The strategic component increases the probability of herding, as individuals are concerned not only about direct payoffs but also their reputation in a secondary labor market.
A related investment model is proposed by Chamley and Gale (1994), as well as Chamley (2004), in which agents can additionally choose the decision time in order to optimize the obtained information, leading to herds of decision delay as long as the discount factor is small enough. More recently, Kasahara (2015) analyzes a model of technology adoption in continuous time with endogenous decision times that generates a variety of different behavior patterns depending on the specific environment, including cascades of early adoption.
Following a slightly different approach, Guarino et al. (2011) develops a model of information cascades with asymmetrically observable actions. Decisions then depend on the number of observed actions and lead to the formation of cascades of observable actions only, while Melissas (2005) develops a model in which firms develop a public technology by privately exerting effort, which can lead to cascades in effort levels (and thus adoption cascades). Lim et al. (2016) builds a model in which herding is prompted solely by network externalities from previous adopters rather than

inference from observable actions.

The two papers that arguably are closest two this modeling approach, however, are Dasgupta (2000) and Choi (1997), since both consider models of cascading that include a notion of payoff externalities in the form of network effects. The former develops a model with a continuous signal space (similar to Banerjee, 1992) where agents only experience positive payoffs if all agents' actions coincide. This relatively extreme form of payoff externalities results in the emergence of cascades without any intrinsic value for the actions.
On the other hand, Choi (1997) also develops a model of innovation featuring adoption cascades. It takes, however, a few fundamentally different approaches in modeling the adoption process that distances it from the original cascading framework; The model features uncertainty only partially as the first adoption reveals the value of the technology. In the absence of private signals, the herding behavior arising in the model is solely due to risk aversion and the payoff externalities rather than through inference from observable actions. Vergari (2004) extends this framework by adding a general uncertainty in valuations caused by a stochastic economic environment.

Cascading behavior, however, is not merely a theoretical idea. Several studies, including Çelen and Kariv (2004) and the seminal work by Anderson and Holt (1997), have replicated the cascading behavior predicted by Bikhchandnai et al. (1992) and Banerjee (1992) in a laboratory setting. While Kübler and Weizsäcker (2004) also replicated cascading behavior in the lab, they showed that the depth of reasoning (and inference) of subjects is often limited.
The evidence for cascades, however, is not limited to an experimental context but extends into the adoption of technology by firms. In a follow-up paper, Bikhchandani et al. (1998) backs-up the original framework of informational cascades by arguing for cascades as a potential explanation for innovation behavior and presenting empirical evidence (among others Kennedy, 1997 and Chaudhuri et al., 1997) for herding in technology adoption and product introduction. Later, Walden and Browne (2003) provide evidence for herding in the adoption of e-commerce technologies, Xu et al. (2012) in the adoption of Open Standard Networks and Chatzimichael et al. (2014) shows the existence of adoption cascades in the German and Greek agricultural sector.

The aspect of cascades and herding is also present in the technology diffusion literature. For instance, Rogers (1974) attributes the phenomenon of technology abandonment to, among others, occurrences of cascading behavior, while Geroski (2000) offers cascades as a conceptual explanation for, particularly, increasing diffusion rates, and Oorschot et al. (2018) identifies herding models as part of the theory explaining innovation.

Lastly, modeling technology adoption as a decision problem governed by uncertainty is motivated by Camilo et al. (2020), which uses Spanish firm-level data and the firm-level automation framework of Humlum (2019) to analyze the patterns of robot abandonment largely overlooked by the automation literature. It finds that approximately every second firm that adopts robots will abandon them subsequently, which is well explained by a model of firms uncertain about the success of automation that only learn about the firm-specific automated production costs after having robotized.[2]

While the idea of cascades and herding behavior in technology adoption is well established in the literature, the proposed model, to the best of my knowledge, is the first to model cascades in technology adoption in a market environment with observable actions and competition externalities on profits as well as to discuss the connection between cascading behavior and technology abandonment. Furthermore, it is closely integrated with the original framework developed by Bikhchandnani et al. (1992), which makes it possible to apply results and extensions based on the original model relatively directly.

2 Baseline Model

A set of homogeneous firms $i \in \{1, \ldots, n\}$ compete in (locally) separated markets. Each firm chooses sequentially whether to adopt (denoted by $a_i = 1$) or omit ($a_i = 0$) an available technology with i denoting each firm's position in the queue.[3] Adopting the technology can have either high or low success, determined globally by the state of the world $s \in \{H, L\}$ which is unknown to the firms. Each firm, however, receives a private signal θ_i of precision $\lambda \geq 0.5$ about the state s, i.e., $\mathbb{P}[\theta_i = \mathbf{1}_s(H)] = \lambda$.[4]
The profits of firm i depend on the adoption decision and the state of the world, i.e., $\pi_i = \pi(a_i, s)$. If deciding not to adopt the technology, firm i receives profits $\pi(0, s)$ independent of the state s. If it decides to adopt, on the other hand, its profits depend on the success of the technology, with $\pi(1, H) > \pi(0, s) > \pi(1, L)$.[5]
Lastly, in order to make an adoption decision, the firms need to form beliefs over the state of the world. To that end, denote by p^s the common prior belief coinciding with the actual probability of high success, that is $p^s := \mathbb{P}[s = 1]$. On the other hand, let $\beta(\mathcal{I}_i)$ denote the posterior belief of firm i with information set $\mathcal{I}_i$, i.e., $\beta(\mathcal{I}_i) = \mathbb{P}[s = 1 | \mathcal{I}_i]$.

Since $\pi(1, H) > \pi(0, s) > \pi(1, L)$, a firm i will adopt the technology if and only if it is certain enough about the success of innovating. Put differently, firm i will automate whenever $\beta(\mathcal{I}_i) > \bar{\beta}$ where $\bar{\beta}$ is the threshold at which the firm is indifferent between adopting and not adopting. It is given by

[2]In the model, the time-invariant automation cost reflects the firm's intrinsic suitability for automation that is only revealed after initially automating.

[3]As is standard in the herding literature, this model follows an *exogenous queue assumption*, ensuring that firms cannot choose their position in the decision-making queue.

[4]This is an informal way of saying $\mathbb{P}[\theta_i = 1 | s = H] = \mathbb{P}[\theta_i = 0 | s = L] = \lambda$. Note that because of the structure of the signal and state space the precision λ has to at least 0.5. For any $\lambda < 0.5$ the firms could simply use $\tilde{\theta}_i := 1 - \theta_i$ with precision $1 - \lambda > 0.5$ instead.

[5]This is again an assumption; however, every other case would make the adoption decision trivial.

$$\mathbb{E}_\beta[\pi(1,s)] = \beta\pi(1,H) + (1-\beta)\pi(1,L) = \pi(1,L) + \beta\big(\pi(1,H) - \pi(1,L)\big) = \pi(0,s)$$

$$\implies \bar{\beta} \coloneqq \frac{\pi(0,s) - \pi(1,L)}{\pi(1,H) - \pi(1,L)} \in (0,1)$$

I refer to beliefs $\beta(\mathcal{I}_i) > \bar{\beta}$ which prompt adoption as *affirmative beliefs* in contrast to *dissuasive beliefs* when $\beta(\mathcal{I}_i) < \bar{\beta}$. Finally, whenever a firm is indifferent ($\beta(\mathcal{I}_i) = \bar{\beta}$) it will break ties in favor of its own signal, that is $a_i = \theta_i$.

Independence and Observable Information

Before turning towards the main model with *observable actions*, two simpler cases are worth discussing briefly. In the first case, firms cannot observe anything and are thus making the adoption decision independently, relying only on their private information θ_i. The posterior beliefs of a firm i are, thus, given by

$$\beta(\theta_i = 1) = \mathbb{P}[s = H|\theta_i = 1] = \frac{\mathbb{P}[s = H]\mathbb{P}[\theta_i = 1|s = H]}{P[\theta_i = 1]} = \frac{p^s\lambda}{p^s\lambda + (1-p^s)(1-\lambda)} =: \beta^1$$

$$\beta(\theta_i = 0) = \mathbb{P}[s = H|\theta_i = 0] = \frac{\mathbb{P}[s = H]\mathbb{P}[\theta_i = 0|s = H]}{P[\theta_i = 0]} = \frac{p^s(1-\lambda)}{p^s(1-\lambda) + (1-p^s)\lambda} =: \beta^0$$

Note that $\lambda > 0.5$ ensures that $\beta^1 > p^s > \beta^0$ (see Appendix A.1). Under reasonable (or non-trivial) parametrizations, meaning $\bar{\beta} \in (\beta^0, \beta^1)$, it follows that if firms observe only their private signal, they will behave accordingly. Put differently, under independence, firms strictly follow their own signal and innovate if and only if they receive a positive signal.
In contrast, if each firm's signal is (publicly) observable rather than private information, it is straightforward to show that each firm will follow the signal that prevails in aggregate.[6] As a consequence, each firm's own signal becomes less important and pivotal the later it decides.

Observable Action

Under *observable actions*, rather than being able to observe competitors' signals directly, each firm is able to observe previous actions. These actions can then be used for making inference about the corresponding signals. Solving the decision problem sequentially, the first firm is faced with the same situation as any independent firm, having to rely solely on its own signal θ_1, resulting in $a_1 = \theta_1$. Therefore, as long as firms follow their signal, firms indirectly observe the history of signals.[7] As a consequence, any following firm can face two kind of histories $\mathcal{I}_{i-1}$:

[6] See Appendix A.3 for more details.

[7] In this case a history refers to a sequence of actions that allow for a direct recuperation of signals. As a result we obtain $(a_1, \dots, a_{i-1}) = (\theta_1, \dots, \theta_{i-1}) = \mathcal{I}_{i-1}$ and can thus write $\mathcal{I}_i = (\mathcal{I}_{i-1}, \theta_i)$

1. *Continuation history.* The inferred previous signals $\mathcal{I}_{i-1}$ are ambiguous enough that the adoption decision depends on the own signal θ_i, that is $\beta(\mathcal{I}_{i-1}, 1) > \bar{\beta} > \beta(\mathcal{I}_{i-1}, 0)$. Consequently, the firms continue to follow their individual signal.

2. *Cascade history.* The inferred previous signals $\mathcal{I}_{i-1}$ are conclusive enough that the adoption decision is independent of the own signal θ_i, that is $\beta(\mathcal{I}_{i-1}, 0) > \bar{\beta}$ or $\beta(\mathcal{I}_{i-1}, 1) < \bar{\beta}$. As a direct consequence, firm i disregards its individual signal (instead $a_i = \text{mod}(\mathcal{I}_{i-1})$). Thus, any subsequent firms have no means to infer θ_i. Since firms are homogeneous, every firm $k > i$ can only infer $\mathcal{I}_{i-1}$ from all $k-1$ observable actions, leading to the same adoption decision irrespective of their individual signal θ_k. This behavior of firms to disregard their own information and adopt the predominant observable action is referred to as a *cascade*.

One characteristic of cascades is that they are *absorbing*; once emerged, the aggregation of additional information via the inference process is stopped. That leads to the predominant action prevailing and any subsequent history prompting the same action.
The partition of possible histories into these two classes depends on the environment, particular on the value of $\bar{\beta}$. To that end, each possible history is sufficiently represented (in the sense of *sufficient statistics*) by the excess actions $\Delta_i^a = \Sigma \mathbf{1}_{a_k}(1) - \Sigma \mathbf{1}_{a_k}(0) = 2\Sigma a_k - i$. Assuming $\bar{\beta} \in \{\beta^0, \beta^1\}$ there are three relevant cases to consider, illustrated in Figure 1

1. $\bar{\beta} = p^s$: This is the original setting used in Bikhchandani et al. (1992). The resulting partition is symmetric with firms following their own signal as long as $|\Delta_{i-1}^a| \leq 1$ and cascades emerging whenever $\Delta_{i-1}^a \geq 2$ (adoption) or $\Delta_{i-1}^a \leq -2$ (omission).[8]

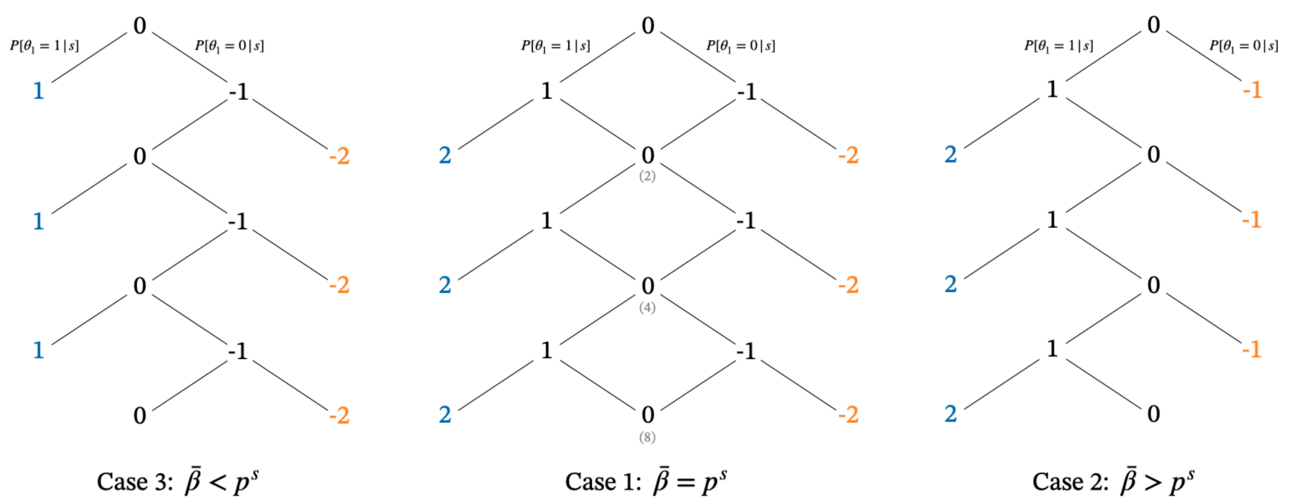


Figure 1: Decision tree of the baseline model for each case. Each node, featuring the excess actions Δ_{i-1}^a, represents the decision made by a firm, with the left (right) branches representing the action when $\theta_i = 1$ ($\theta_i = 0$). Blue (orange) nodes denote the start of an adoption (rejection) cascade and the numbers in brackets the amount of paths leading to a specific node.

[8]Note that even though the tie-braking rule in case of indifference is slightly different than in the original paper (symmetric mix of actions), the result is the same.

2. $\bar{\beta} > p^s$: Compared to the first case, the higher $\bar{\beta}$ implies that firms need to be more optimistic in order to induce adoption. As a consequence, firms are faster to decline, and omission cascades start already when $\Delta^a_{i-1} \leq -1$, while adoption cascades continue to emerge for $\Delta^a_{i-1} \geq 2$.

3. $\bar{\beta} < p^s$: Analogous to the previous case, firms are more easily convinced to adopt. Thus, adoption cascades emerge whenever $\Delta^a_{i-1} \geq 1$ while $\Delta^a_{i-1} \leq -2$ induces omission cascades.

With this we can therefore express the posterior belief for any firm depending directly on the observable actions and their individual signal, $(\Delta^a_{i-1}, \theta_i)$. In the first case, this would yield

$$\beta(\Delta^a_{i-1}, \theta_i) = \frac{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i)-\mathfrak{d}_i^-}}{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i)-\mathfrak{d}_i^-} + (1-p^s)\lambda^{(1-\theta_i)-\mathfrak{d}_i^-}(1-\lambda)^{\mathfrak{d}_i^+ + \theta_i}}$$

where $\mathfrak{d}_i^+ := \text{median}(0, \Delta^a_{i-1}, 2)$ and $\mathfrak{d}_i^- := \text{median}(-2, \Delta^a_{i-1}, 0)$ denote the maximal inferable adoption and omission signals respectively for the first case.[9] The corresponding expression for the cases 2. and 3. as well as the derivations for the respective partitioning can be found in Appendix A.2.

Consider an exemplary model with $n = 4$ firms for case 1, depicted in Figure 2. Each node features the adoption threshold of the corresponding history, which in this case is constantly $\bar{\beta} = p^s = 0.6$. Each branch contains the corresponding firm's belief depending on the observed history and private signal.[10] Whenever enough information is aggregated (in this case, $a_1 = a_2$), the private signal of firm $i = 3$ is not sufficiently strong to induce the corresponding action. After observing two

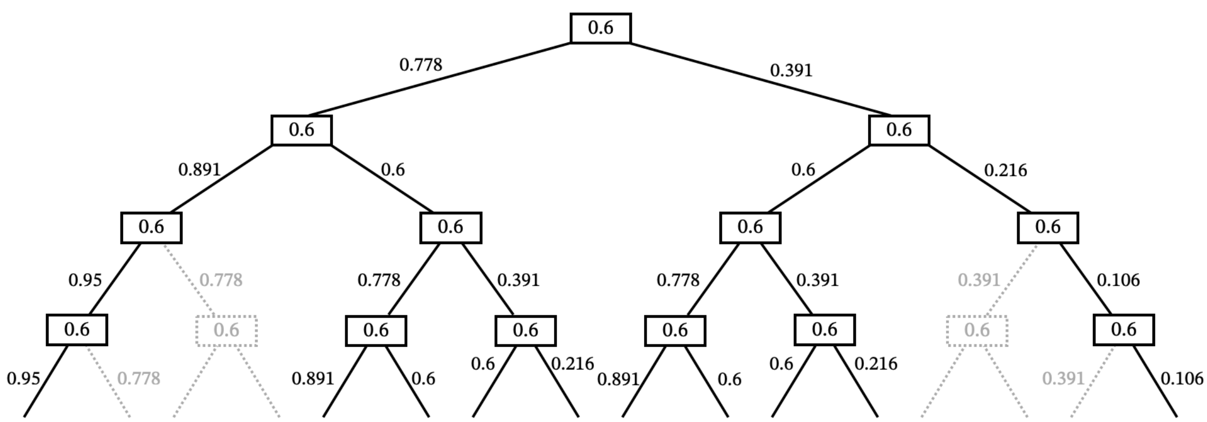


Figure 2: Structure of a baseline model with $n = 4$ firms. The bordered numbers at each node are the adoption thresholds $\bar{\beta} = p^s = 0.6$, the numbers at the left (right) branches the belief held by the corresponding firm after observing $\theta_i = 1$ ($\theta_i = 0$). Movement along left (right) branches corresponds to $a_i = 1$ ($a = 0$) with grayed out portions representing never-played actions due to cascading behavior.

[9]Put differently, whenever a firm observes a difference of more than two actions it is part of a cascade. Thus, it can make no inference about the signals of the additional actions.

[10]For instance, after observing $a_1 = 1$ firm 2 can either obtain $\theta_2 = 1$, resulting in a belief $\beta(2) = 0.891$ or $\theta_2 = 0$ with $\beta(0) = p^s = 0.6$.

adoptions, for instance, firm 3 can only hold beliefs $\beta(3) = 0.95$ or $\beta(1) = \beta^1 = 0.778$, which both surpass the threshold $\bar{\beta} = 0.6$, resulting both times in an adoption: the firm disregards its private information and initiates a cascade. At this point, non-reachable branches of the tree are grayed out. Since no additional information is provided through its adoption decision, firm $i = 4$ faces the identical situation, making the same inference and choosing the same action and, again, ignoring its private signal in the process.

3 Competitive Model

Now, I extend the baseline model to capture additional characteristics of technology adoption in a market setting by allowing for heterogeneity across firms and market interactions. The information and signal structure remain unchanged.

Denote by $A_{-i} := \Sigma_{j \neq i} a_j$ the number of competing adopters for firm i. Then the profits of firm i are given by $\Pi_i(a_i, A_{-i}, s) = \pi_i(a_i, s) - \gamma V_i(A_{-i}, s)$, where π_i captures the direct, individual profits of the adoption decision. On the other hand, V_i captures the interaction effect of other firms successfully adopting and is dependent on the state (with $V_i(\cdot, 1) > V_i(\cdot, 0)$) and increasing in the number of adopters ($\partial V_i / \partial A_{-i} > 0$). Lastly, the competition parameter γ captures the strength of competition in the market. A positive γ indicates standard competition between substitutable goods, whereas $\gamma < 0$ represents complementary interaction effects, with higher $|\gamma|$ denoting stronger competition. It is easy to see that this extension nests the baseline model presented in Section 2 by setting $\gamma = 0$ and $\pi_i(\cdot) = \pi(\cdot)$ for all i.

Diminishing Adoption Gains

One first extension of he baseline model is to look at firms that still act independently ($\gamma = 0$) but where firms that adopt earlier obtain higher profits, in the sense that $\pi_i(1, H) \geq \pi_j(1, H)$ for $i < j$, with all other profits constant across firms. Since there are no interaction effects, the adoption threshold is given, analogous to before, by

$$\bar{\beta}_i := \frac{\pi(0, s) - \pi(1, L)}{\pi_i(1, H) - \pi(1, L)} \in (0, 1)$$

It follows immediately, that $\bar{\beta}_i \leq \bar{\beta}_j$ for $i < j$. Since the threshold is increasing, later firms become gradually less optimistic about adopting, which can lead to firms not adopting even if faced with a cascade. This is the case whenever $\bar{\beta}_i > \beta^{\times}$ where

$$\beta^{\times} := \frac{p^s \lambda^3}{p^s \lambda^3 + (1 - p^s)(1 - \lambda)^3}$$

denotes the participation threshold (derivation in Appendix A.4).

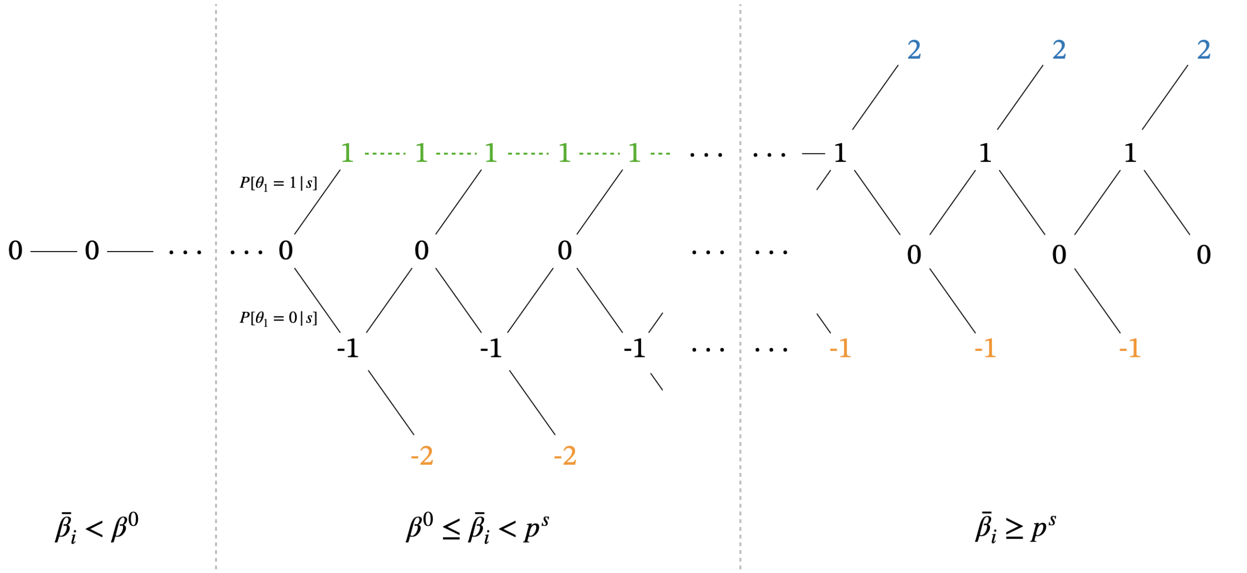


Figure 3: Decision tree of the model with heterogeneous adoption thresholds $\bar{\beta}_i$. Each node, featuring the excess actions $\mathfrak{d}_i^+ + \mathfrak{d}_i^-$, represents the decision made by a firm, with the upper (lower) branches representing the action when $\theta_i = 1$ ($\theta_i = 0$). Blue (orange) nodes denote the start of a permanent adoption (omission) cascade, while green nodes mark the start of temporary adoption cascades.

Interesting and distinct behavior arises when $\bar{\beta}_1 < p^s$.[11] If there exist firms for which $\bar{\beta}_i < \beta^0$, those firms will always hold affirmative beliefs, making their adoption non-informative for all following firms. The first firm i with $\bar{\beta}_i \in (\beta^0, p^s)$ faces a history of observable but non-informative actions, making it equivalent to the first mover in the first baseline case. This first case behavior (cascades at $\Delta^a_{i-1} \geq 1$ or $\Delta^a_{i-1} \leq -2$) continues until the threshold has increased to the point that $\bar{\beta}_i \geq p^s$. At this point, the amount of observed informative actions required to start a cascade pivots. In particular, the condition for an adoption cascade increases from $\Delta^a_{i-1} \geq 1$ to $\Delta^a_{i-1} \geq 2$. Consequentially, the limited information aggregated in an early adoption cascade is insufficient to continue to sustain the cascade in the high-threshold environment. Instead, the cascade is disrupted, and firms start following their own signal and aggregating information again until eventually a new cascade emerges. By contrast, a non-adoption cascade is sustained more easily (from $\Delta^a_{i-1} \leq -2$ to $\Delta^a_{i-1} \leq -1$) and thus an early non-adoption cascade will continue when thresholds increase. The structure of the model with increasing $\bar{\beta}_i$ is illustrated in Figure 3.

Market Interaction

In most environments, firms do not act as monopolists in an isolated market. Their behavior is, rather, a result of strategic interactions between competing firms. In the context of technology adoption, these interactions manifest in the externalities of other firms' innovations, V_i, by relaxing the constraint $\gamma = 0$.

In order to analyze how these interaction effects influence behavior, it is necessary to study their

[11] If instead $\bar{\beta}_1 > p^s$ it is easy to see that this is equivalent to the third case of the baseline model with n' firms where $n' = \max\{i : \bar{\beta}_i < \beta^\times\}$.

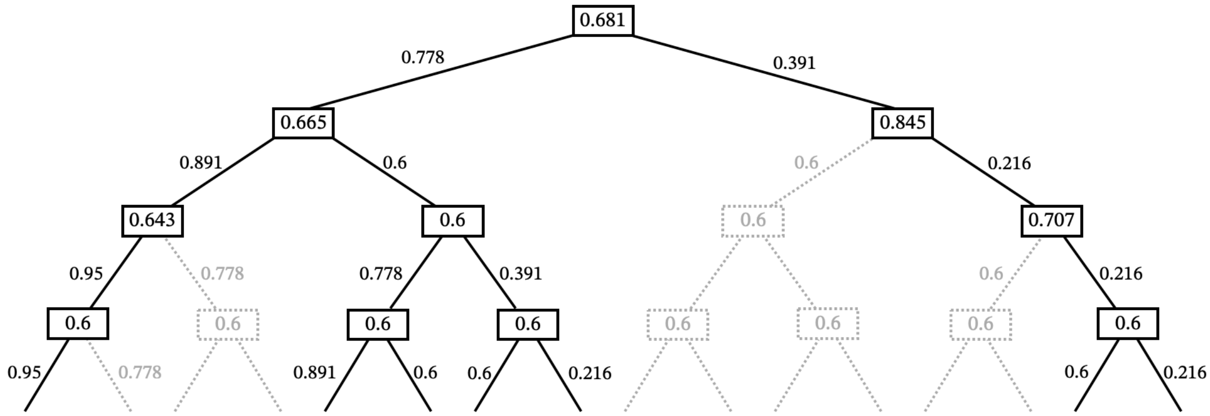


Figure 4: Structure of a competitive model between $n = 4$ firms with substitutes ($\gamma > 0$). The bordered numbers at each node are the adoption thresholds $\bar{\beta}_i^M$, obtained via simulation (see Appendix A.8), the numbers at the left (right) branches the belief held by the corresponding firm after observing $\theta_i = 1$ ($\theta_i = 0$). Movement along left (right) branches corresponds to $a_i = 1$ ($a = 0$) with grayed out portions representing never-played actions due to cascading behavior.

influence on the adoption thresholds $\bar{\beta}_i^M$. By including interaction effects, the model becomes not only *backward-looking* (with respect to the inference about others' private signals) but also *forward-looking* as firms internalize the effect of their action on subsequent adoption decisions. Figure 4 depicts an example of a competitive model between four firms (in contrast to the baseline model in Figure 2). The dualism of forward- and backward-looking elements poses some limitations on solving the model with competition effects in general, as it results in the model not being solvable by pure forward or backward induction. Instead, equilibrium behavior has to be found in an iteratively convergent manner. The following, simplified reasoning consists of the initial step of the iteration and captures the fundamental characteristics of adoption behavior under competition. The results are formalized in Proposition 1 and provide an intuition for the way in which the competitive model differs from the baseline.

In order to analyze this class of models, I will follow an ensemble approach with respect to the sequence of individual signals θ_i. This approach (also referred to as *coupling*) interprets the sequence of signals $(\theta_1, \ldots, \theta_n)$ as one realization of a joint stochastic process rather than a sequence of individual realizations of independent stochastic processes.[12]

Now, consider firm i's decision and fix, for simplicity, the strategy (i.e., the threshold $\bar{\beta}_j$) of every firm $j \neq i$. The profit of firm i from a_i for any $(\theta_1, \ldots, \theta_n)$ and s can be expressed as $\pi_i(a_i, s) - \gamma V_i(a_i, (\theta_1, \ldots, \theta_n), s)$.[13] The firm cannot influence a_j for $j < i$, so differences in V_i are caused by the influence of a_i on a_j for $j > i$. Coupling the signal sequence allows for directly comparing actions a_i with respect to V_i for any possible state of the signal space. Because the start of a cascade is directly tied to the observed difference of actions, it has to be that

[12]This is a common technique used frequently in time series econometrics to establish certain stochastic properties. It is, however, also used in decision theory to derive stochastic dominance, as for instance in Gossner et al. (2018).

[13]This is not implying that $A_{-i} = \Sigma_{j \neq i} a_j$ is equal independent of a_i. Instead, given a_i, the sequence of actions and thus also A_{-i} is fully determined by $(\theta_1, \ldots, \theta_n)$.

$A_{-i}(a_i = 1) \geq A_{-i}(a_i = 0)$ (see Appendix A.5,). Since $\partial V_i(\cdot)/\partial A_{-i} \geq 0$ it follows that for any signal sequence $V_i(a_i = 1, s) \geq V_i(a_i = 0, s)$.[14] With that, the threshold in the model with market interaction can be expressed as

$$\bar{\beta}_i^M(\gamma) := \frac{\pi(0, s) - \pi(1, L) + \gamma \Delta V_i^L}{\pi_i(1, H) - \pi(1, L) + \gamma(\Delta V_i^L - \Delta V_i^H)}$$

where $\Delta V_i^s := \mathbb{E}_A[V_i|1, s] - \mathbb{E}_A[V_i|0, s] \geq 0$.[15] Define the change in threshold through competitive interaction effects, $\zeta_i(\gamma) := \bar{\beta}_i^M(\gamma) - \bar{\beta}_i$. The following properties of $\zeta_i(\gamma_i)$ can then be shown (see Appendix A.7).

Proposition 1. *Under fixed strategies, competition between substitutes ($\gamma > 0$) raises the adoption threshold ($\zeta_i(\gamma) > 0$) and competition between complements ($\gamma < 0$) decreases it ($\zeta_i(\gamma) < 0$). Further, the threshold change ζ_i is increasing in competition intensity, i.e., $\partial\zeta_i/\partial\gamma > 0$.*

The proposition is in line with the results observable in the example in Figure 4. Compared to the baseline (Figure 2), the thresholds for adoption increase and make innovation less likely. Appendix A.9 presents additional examples for differences in magnitude of γ and competition with substitutes that are equally in line with the remaining claims in Proposition 1. Taking into account the effect of competitors adopting on profits decreases the incentives to adopt as it sends a positive signal to subsequent competitors whose adoption decision decreases profits. With stronger competition (and thus stronger externalities), the strategic aspect of avoiding signals increases until it outweighs the direct benefits of adopting. This is the case, for instance, for firm $i = 2$ after observing $a_1 = 0$ in Figure 4. Thus, under strong enough competition, the effect of γ on $\bar{\beta}_i^M$ could be strong enough to inhibit the emergence of adoption cascades altogether.

The simplifying assumption that strategies are fixed, made to illustrate the key characteristics of adoption cascades in the presence of competitive effects is inaccurate in the sense that competing players adapt their strategies depending on each firms' strategic behavior. Because, as outlined before, the strategic behavior is both backward- and forward-looking, the equilibrium behavior of each model is found in a converging manner: Imposing strategic thresholds for one firm changes the thresholds of other firms as well, to which the strategic thresholds adapt again. This means, however, that even though the strategies of competitors change in reaction, the overall tendency of the thresholds to increase, described in Proposition 1, is unaffected.

4 Results

After having characterized the behavior of individual firms in the baseline and competitive models in Sections 2 and 3 respectively, this section will shift focus to a more aggregate point of view. Knowing the adoption strategy of each firm and the probability structure of the signal and state

[14] Note that this implies that without fixing a signal sequence $\mathbb{E}_A[V_i|1, s]$ FOSD $\mathbb{E}_A[V_i|0, s]$, see Appendix A.6.

[15] For more details, again, see Appendix A.6.

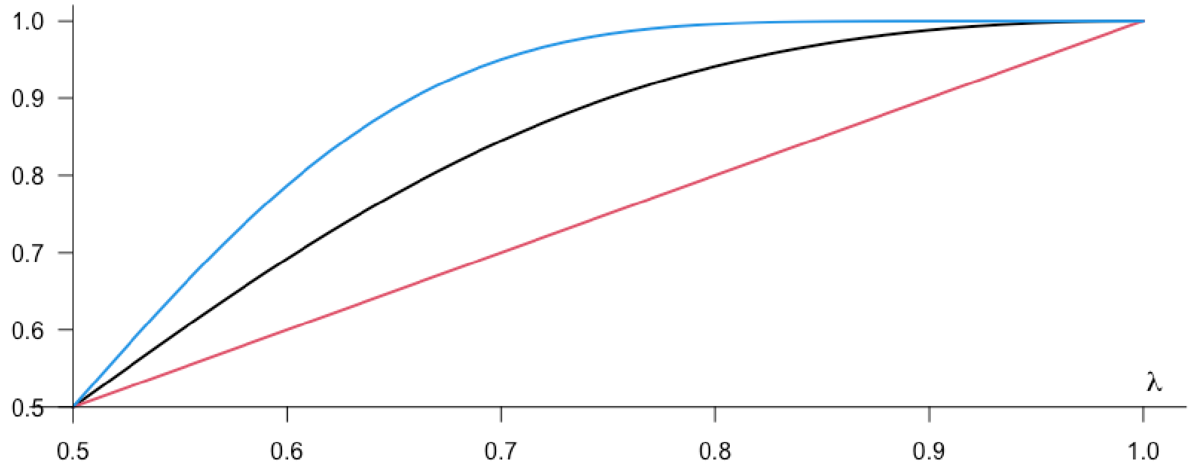


Figure 5: Probability of ending up in the correct cascade (black) after $n = 16$ firms in comparison to the probability of the last firm playing the correct action $a_n = \mathbf{1}_s(H)$ for independent firms (red) and observable signals (blue).

space, the probabilities of a cascade emerging can be calculated.

For that purpose, there are two different points from which to analyze the occurrence of cascades. In the *ex-post* view, the state s is assumed to having already realized. The assessment of actions and cascades can, thus, be conditioned on the state, giving rise to the notion of a *correct action*, $a_i = \mathbf{1}_s(H)$, as corresponding to the profitability of the state. Consequently, a *correct cascade* is one that prompts the correct action. Figure 5 depicts the probabilities of a correct cascade emerging at one point for the baseline case, together with the equivalent for the alternative baseline specifications with independent firms and publicly observable signals.[16] The probabilities depicted clearly reflect the accuracy of the obtained information. In the case of independent firms (red line), the last firm -as any other firm before- has to rely solely on its own signal. The accuracy of the information set of the last firm is, thus, equivalent to the signal precision λ. In the standard case of observable actions, in contrast, the inference firms are able to make leads to the aggregation of information, which increases accuracy (and thus the probability of deciding correctly) compared to independence. The emergence of cascades, however, stops the process of aggregating information, which leads to lower probabilities compared to the case with observable signals and continuous aggregation.

Figure 6 depicts the probabilities of ending up in a correct cascade for the different cases of $\bar{\beta}$. The black line is identical to the standard model (black line as well) in Figure 5. Cases 2 ($\bar{\beta} > p^s$) and 3 ($\bar{\beta} < p^s$) produce different probabilities depending on the realization of the state s due to the asymmetrical cascade cut-offs. Instead, it is more useful to summarize them depending on whether they facilitate the correct (omission in case 2 and adoption in case 3) or incorrect (conversely) cascade. Then, facilitated cascades (red) are more likely to occur than in the standard case. On the other hand, even if the emergence condition is the same for regular and non-facilitated cascades, recall that, as can be seen in Figure 1, the facilitated cascade reduces the possibilities for a non-facilitated cascade to occur, which results in a lower probability compared to the baseline.

[16]For a more extensive overview, including the limiting case $n \to \infty$, see Appendix A.12.

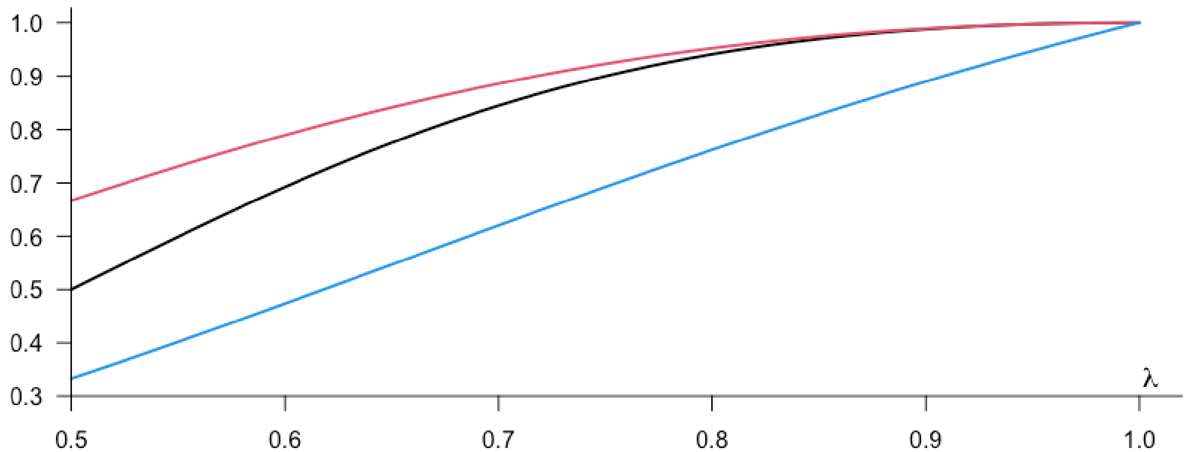


Figure 6: Probability of ending up in the correct cascade in the limit ($n \to \infty$). The black line corresponds to the first case ($\bar{\beta} = p^s$) while the red (blue) line corresponds to the case facilitating a correct (incorrect) cascade.

In contrast to the state-dependent ex-post probabilities, the *ex-ante* view is concerned with each action emerging unconditionally, that is absent any notions of optimality. Instead it comprises the ex-post probabilities of each realization of the state s and is, thus, dependent on the success probability and prior p^s. Figure 7 summarizes the probability of an adoption cascade emerging in the limit ($n \to \infty$) for the different cases of $\bar{\beta}$. Note that because cascades are absorbing, in the limit, a cascade has to emerge almost surely. Thus the probabilities for an omission cascade emerging are the complementary probabilities to the ones depicted in Figure 7.[17]

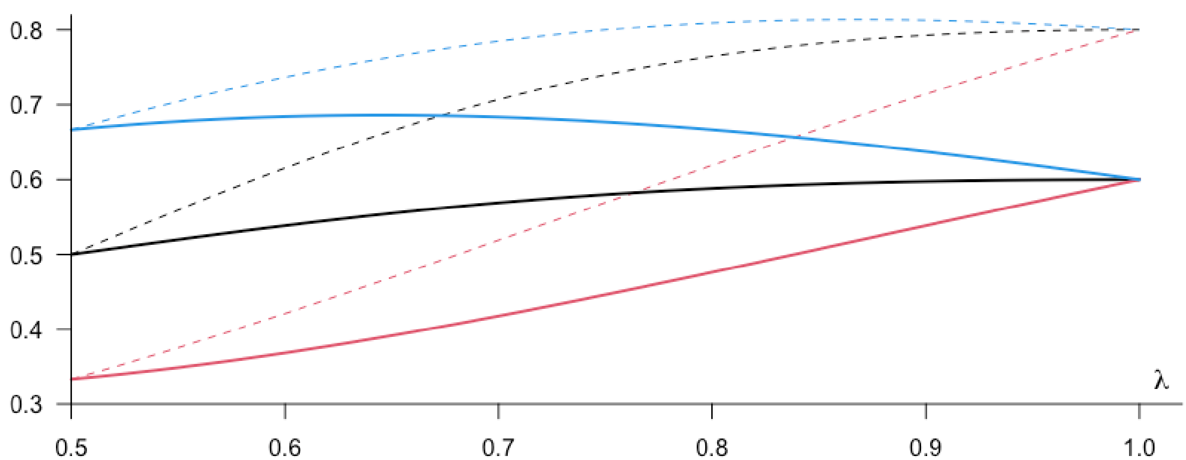


Figure 7: Probability of ending up in an adoption cascade in the limit ($n \to \infty$). The black line corresponds to the first case ($\bar{\beta} = p^s$) while the red (blue) line corresponds to the second (third) case with $\bar{\beta} > p^s$ ($\bar{\beta} < p^s$). Dashed lines represent the cascade probabilities with an increase in p^s.

[17] Also, note that there is nothing inherently favoring adoption cascades. The fact that the probabilities for an adoption cascade are higher is just a direct result of the fact that $p^s > 0.5$.

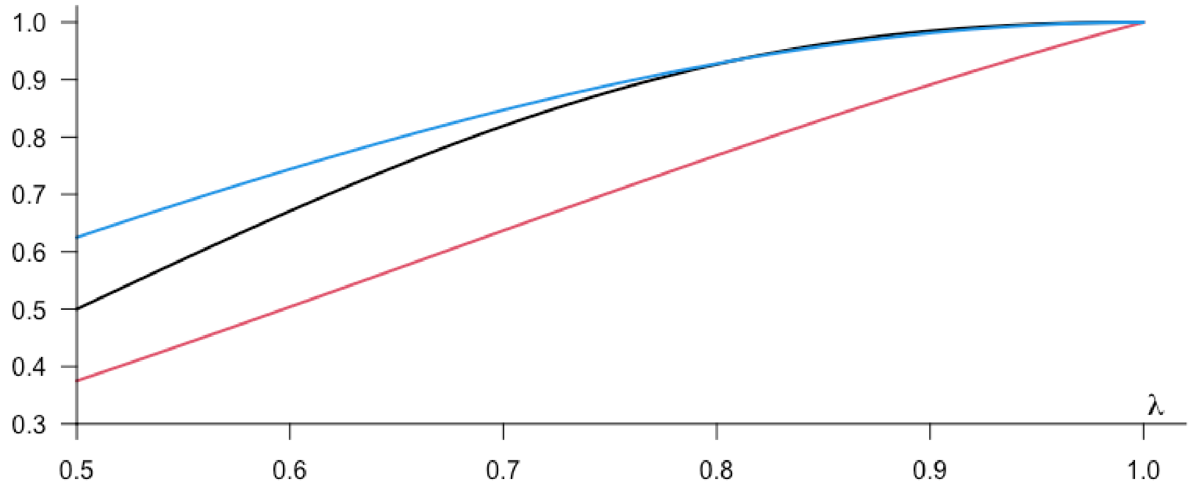


Figure 8: Probability of ending up in the correct cascade in a model with $n = 4$ firms. The black line represents the baseline model (Figure 2), while the red (blue) line corresponds to the exemplary competitive model with complements (substitutes), depicted in Figure 4 (10b).

As mentioned before, in comparison to the purely backward-looking baseline model, the competitive model encompasses both forward- and backward-looking elements, limiting the degree to which a general characterization of firm behavior is possible. As a result, although there are general trends in aggregate behavior of the model, the particular behavior depends on the specific model environment. I will, thus, illustrate the general behavior using the previously explored exemplary environment (Appendix A.9 and Figures 2, 4 and 10).

Similar to before, the models produce asymmetric behavior for adoption and omission. Therefore, the likelihood of correct cascades depends on the realized state s as well. Figure 8 contains the probability of ending up in a cascade in the two competitive models (with complements and substitutes) compared to the black baseline in the case of $s = H$.[18] Since the negative externalities from competition in complements ($\gamma > 0$) reduce incentives to innovate by increasing the adoption threshold ($\zeta_i > 0$), the corresponding probability (red line) is lower than in the baseline ($\gamma = 0$).[19] Conversely, the positive externalities from competition in substitutes cause the probabilities of an adoption cascade to increase.

In the case of externalities from adoption there is one additional aspect of interest. Figure 9 depicts the expected number of adopters in the baseline and both competition cases for the case of $s = H$, that is $\mathbb{E}[A|s = H]$. By the same arguments as before, competition between complements (red line) decreases the individual incentives to adopt and thereby the expected quantity of adoption while substitutes introduce positive externalities, raising the number of firms innovating in expectation.

[18] While this need not hold in general, in this case the optimal actions for $\gamma = 1$ (Figure 4) and $\gamma = -1$ (Figure 10b) are mirror-images of other. Thus, the probability of a correct cascade under $\gamma = 1$ and $s = H$ equals the probability of a correct cascade under $\gamma = -1$ and $s = L$.

[19] Note that for calculating the cascade probability, $a_1 = 0$ was treated as the start of a cascade, since it induces a cascade for firms 2 and 3, even if $i = 4$ follows its own signal again.

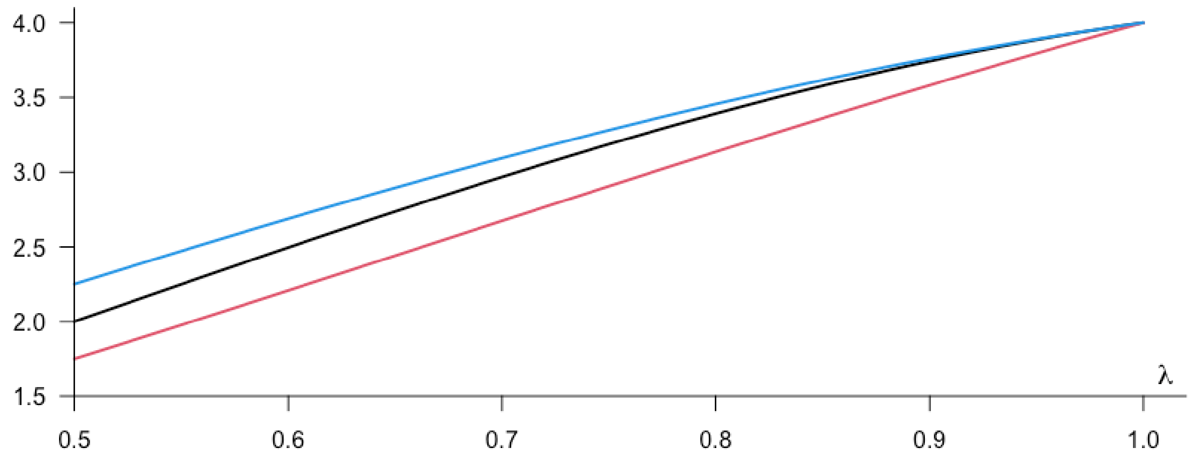


Figure 9: Expected number of adopters $\mathbb{E}[A|s = H]$ in a model with $n = 4$ firms. The black line represents the baseline model (Figure 2), while the red (blue) line corresponds to the exemplary competitive model with complements (substitutes), depicted in Figure 4 (10b).

Adoption Cascades and Technology Abandonment

One of the aspects that motivated this model was the observation of a substantial portion of firms that adopted industrial robots and abandoned them promptly afterward (Camilo et al., 2020). This phenomenon appears counterintuitive and is usually overlooked or ignored by the automation literature that assumes robotization to be irreversible (see, for instance, the seminal work by Acemoglu and Restrepo, 2018).

However, it fits into a larger body of evidence for *disenchantment discontinuances* (Rogers, 1974), that is, rejecting a technology (promptly) after having adopted it. Studies found not only a substantial fraction of firms discontinuing innovations across a variety of sectors, one of the key characteristics of discontinuances is that they are considerably more common among later adopters. One of the explanatory factors introduced by Rogers (1974) is that (particularly later) adopters base their decision on imitating previous adopters rather than on objective evaluations based on scientific studies of the innovation's consequences.

As this model considers a single, global success probability s, there is no sense in which late adopters differ from early adopters. To that end, it seems reasonable to consider a potential extension proposed by Bikhchandani et al. (1992). Instead of assuming perfectly correlated valuations (or equivalently perfectly correlated success probabilities $s_i = s$), it could be more realistic to allow for individual states s_i to be positively correlated, generating firm-level heterogeneity in innovation success. This heterogeneous but correlated success of innovation can be thought of as the result of a combination of sector-wide global and firm-specific individual characteristics. Following the conjecture of the original model, introducing such imperfectly correlated states should affect adoption thresholds and general and aggregate results quantitatively but not qualitatively.

In the absence of aggregated information from previous actions, adoption is purely (or at least predominantly) based on each firm's individual characteristic. With increasing levels of information aggregation, the innovation decisions should shift to be increasingly determined by global characteristics (captured by the inference from observable actions). Thus, firms adopting later would be more susceptible to following cascades and disregarding private information about their individual characteristics in the process. This circumstance is exacerbated if, as Rogers (1974) suggests, early adopters act like *fashion-leaders* in the sense that better-informed firms (that is $\lambda_1 \geq \lambda_2 \geq \ldots$) decide first.

The Role of Economic Policy

As is commonly the case in economics, economic policies can be designed to achieve a broad variety of goals. In the case of innovation and the adoption of technology, economic policy could be concerned with promoting the adoption of specific technologies due to domestic interests (for instance, different rivaling technologies might affect domestic sectors differently) or ethic, social, national security or environmental concerns. Alternatively, economic policy could be driven by consumer welfare (for instance, by disincentivizing innovations that would lead to a consolidation of market power). While these aspects are not explicitly present in the model, they can be mostly subsumed under a class of policies that try to elicit or dissuade one specific action.[20]

However, economic policy should additionally also aim to maximize the direct welfare from the adoption decisions. Since $\pi_i(1, H) > \pi(0, s) > \pi(1, L)$, maximizing $W = \Sigma\pi_i(a_i, s)$ is equivalent to eliciting the correct action $a_i = \mathbf{1}_s(H)$. Assuming the government faces a similar uncertainty about the state s as every firm, it cannot target a specific action directly. Rather, as illustrated by Figure 5, the probability of correct adoption decisions can be increased by increasing the level of information aggregation (and thus moving it closer to the case of publicly observable signals). Since the sole responsible for hindering information aggregation is the emergence of cascading behavior, this can be achieved by breaking up formed cascades.

It appears reasonable to assume that the government, rather than observing the true state or being able to affect fundamentals such as λ, is in a similar position as any firm and receives a private signal θ_G of precision λ_G. Rather than deciding on whether to adopt, it instead can decide when and if to (truthfully) reveal its signal publicly.[21]

While information cascades generate highly stable behavior, they are themselves relatively fragile and prone to the emergence of additional information. Consequentially, the cascades' fragility can be leveraged when designing policies that try to elicit specific actions or increase information aggregation. Policies that target one specific action should, therefore, try to break converse cascades, while policies trying to incentivize the correct action should disrupt cascades in general. Thus, the government should publically disclose information after the corresponding cascade emerged.

[20] Note that in contrast to the class of policy discussed next, the targeted policy need not be the correct action given the state s.

[21] This is similar to the assumptions made by Bikhchandani et al. (1992) when examining the role of public information disclosure. In particular, this upholds the fundamental assumption that governmental intervention is unexpected in the sense that firms make no inference from the absence of public disclosure.

While the previously explored avenue of intervention to direct adoption decisions is, in the most abstract terms, concerned with influencing the belief formation, another potential avenue of government action could target the transformation of beliefs into actions. More precisely, since the adoption process is governed by comparing the actual belief to the adoption threshold, the government can not only influence the belief itself (as in the previous cases) but also the belief threshold. As exemplified by Figure 7, increasing the threshold decreases the probability of the technology being adopted in general and of an adoption cascade emerging in particular. While this approach can not affect the general likelihood of a cascade emerging, it is a useful approach to prompt one specific action by influencing the threshold accordingly.

Given that the adoption threshold is a direct result of the firm's expected profit maximization, the most straightforward implementation is to directly target the expected profits of adoption and omission. Subsidizing the desired action (or taxing the reverse) achieves this goal by automatically moving the threshold in the corresponding direction. In case the policy needs to be budget balanced (which the first avenue always is as long as public disclosure is costless), an action-dependent subsidy can be combined with an action-independent tax paid by every firm.[22]

Lastly, under certain conditions, the government can exploit the connection between cascading behavior and market interaction explored in Sections 3 and 4. If the goals of eliciting a certain action and the effect of an increase in market competition coincide, the government can achieve both goals simultaneously by any market regulation policy that increases or facilitates competition.

5 Conclusion

This paper contributes to the literature on cascading behavior and technology adoption and diffusion by developing a model of cascading behavior in technology adoption with payoff externalities capturing the fundamental dynamics of competitive interaction between firms. These externalities differ from commonly explored payoff externalities, such as network effects, in that they asymmetrically depend on one specific action; the adoption of the technology.[23]

Following the framework of Bikhchandani et al. (1992) closely, the model features uncertainty in adoption values, sequential decision making, private information and public actions, as well as externalities from the number of adopters. Similar to the standard herding models, I find that enough coinciding actions prompt the emergence of a cascade in which private information is forgone in

[22] Since the number of firms choosing the desired action is (weakly) smaller than the total number of firms, the universal fee is necessarily smaller than the subsidy. This results in net-positive transfers for choosing the desired action and net-negative otherwise, with the special case of net-zero transfers if every firm acts as desired. This mechanism is similar in spirit to the budget-balanced two-stage mechanism proposed by Shichijo and Nakayama (2009).

[23] While standard network externalities assume that the effect is increasing with the number of conforming players. In this model, however, the competitive advantage of firms that adopt a technology is independent of a firm's own adoption decision, and thus, the externality increases with the number of adopters, even for omitting firms.

favor of inference from the observable actions. The addition of payoff externalities changes the profitability of each action, affecting the adoption threshold and thereby adoption behavior and cascading probabilities.

The cascading mechanism and framework employed by this paper to explain and study technology adoption behavior is undoubtedly not the only possible approach to model technology adoption in general nor imitative adoption behavior in particular. The intention of this paper, therefore, is to elaborate and illustrate one potential explanation for a particular pattern in innovation behavior. To which frameworks and patterns the model is applicable is linked to some of the fundamental properties of the model.

Firstly, it operates in an environment of uncertainty about the value of adoption. Thus, it lends itself particularly to the adoption of newly emerged technologies (as opposed to those that are already relatively established), as they can be expected to be accompanied by a higher degree of uncertainty. Notably, the role of uncertainty and learning appears to play a fundamental role in the area of automation technology (Frey and Osborne, 2017; Camilo et al., 2020).

Secondly, the information available should be (predominantly) private. Since firms should be expected to use all sources of information available, the relevant question becomes whether firms' information originates mainly from public sources. As argued by Bikhchandani et al. (1998) and Kennedy (1997), joint research is unlikely to occur between competitors and data from prior experiences as well as individual market research are usually private. However, even if outcomes of innovation are not immediately or directly observable to competitors, they become inferable from the general economic success in the long run. Thus, assuming private signals is, again, more sensible for newly emerged rather than established technologies.

Thirdly, actions have to be observable. While it is a rather reasonable assumption that the decision to adopt a technology is visible, the same is not necessarily true for the opposite. Albeit there are approaches to model cascading in an environment with asymmetry in the observability of actions (one of which is discussed later), this framework focuses on the case where both actions are equally observable, which in reality is particularly reasonable whenever the consideration to innovate is likely to be observable in itself, such as for big-scale or controversial decisions. Automation decisions are especially probable to fall into the latter category as the threat of labor displacement might prompt the vocal involvement of labor unions, thus, increasing publicity.

Lastly, the adoption decisions have to be exogenously sequential. The sequential nature of adoption decisions itself should be less controversial since firms deciding simultaneously to innovate is considerably less likely than decisions being made sequentially.[24] The assumption that this sequential order is given exogenously is, however, more contestable. Arguably, one of the biggest areas of the herding literature studies the relaxation of the exogenous queue assumption by including the decision point in the firms' action set. In this light, there are two cases to be made for upholding

[24]In fact, nearly the entire technology diffusion literature is built on the assumption that adoption is sequential in nature.

sequential exogeneity. On the one hand, firms might be limited in their ability to freely choose a decision point by factors such as the availability of capital and financing or properties of the production process.[25] On the other hand, more importantly, endogeneizing the decision time does not change the actual adoption problem after choosing a time but rather adds a second layer in front. As such, analyzing the model with the decision order given would be a necessary step before allowing firms to also choose when to adopt, and in some circumstances might be a sufficiently reasonable approximation of decision behavior in itself.

While the model is motivated by patterns of innovation behavior observable particularly in the realm of automation and the adoption of industrial robots and rooted in this competitive market context, the setup itself is relatively general. While this makes it possible to adapt this framework to other environments with payoff externalities that match the original cascading framework without major modification, the generic setup of this model carries some shortcomings as well. The brevity is due to the specification of adoption and omission values directly as fundamentals of the model instead of micro-founding it in the market behavior of firms. Thus, the model can not directly analyze the effect of changes in the market fundamentals (for instance, a change in demand or the cost structure) on the cascading behavior.
A possible solution would be the extension into a two-layered model. The first layer solves the optimization problem of the firm given the state s and profile of actions A, yielding a functional mapping into a profit function as specified in the original competitive model. The second layer consists of the competitive model as described in Section 3 with the advantage that results based on the profit and externality functions Π_i and V_i can now be mapped directly to the fundamentals of the first layer market.

In addition, another major drawback compared to the original model by Bikhchandani et al. (1992) and the baseline model in Section 2 is the characterization of behavior in the competitive model. While in the former the adoption behavior is specifiable for the general case, the latter is only able to characterize the overall dynamics for the general case. The specific behavior, given by each firm's adoption threshold $\bar{\beta}_i^M$, has to be obtained on a case-by-case basis for every specific environment. As a consequence, the general effects of changes in the competitive environment (for instance, an increase in the competition parameter γ) can only be assessed qualitatively, whereas a quantitative assessment would require the specification of a concrete example. The differences in the calculation of cascade probabilities, depicted in Figures 5 and 8 illustrates exactly this fact.

This model is situated in a particular area of research that has not been explored intensively before, with models addressing similar environments following some fundamentally distinct modeling approaches. Thus, there are some natural avenues to be explored next. As discussed before, one

[25] For instance, a car manufacturer might only be able to automate the production of a particular car model when starting production and setting up new production lines.

potential extension concerns the structure of the signal space. Modifying the innovation success to be imperfectly correlated would allow to capture heterogeneity in technology adoption more realistically and analyze firms' innovation behavior and the effects of policy intervention in a more nuanced manner.
Furthermore, in its current state, the model features a minimal state, signal and action space. Extending the domain of states, signals and actions would also allow to model innovation more realistically. In particular, this extension could capture the decision between rivaling technologies as well as innovation intensity.

A more separate avenue of research concerns the fundamental modeling of sequential technology adoption. As mentioned before, in this model, as in the standard herding models, such as Bikhchandani et al. (1992), the order (or decision time) is exogenously given and both actions, adoption and omission, are equally observable. A different approach that captures a distinct but similar environment is to assume that only adoption decisions are observable and firms decide instead when to adopt a technology, similar to Guarino et al. (2011). Since adoption decisions depend on the posterior beliefs about adoption success and the decision order is no longer exogenous, the uniqueness of a solution should be directly related to a source of firm heterogeneity.[26] One natural source of heterogeneity would be to assume a heterogeneous signal precision λ_i. It could be investigated whether this approach results in a situation in which the best-informed firms move first, acting as so-called *fashion-leaders* (Bikhchandani, 1992) and whether, if enough firms innovate soon enough, later firms could be persuaded to disregard their own signal and follow the majority of observable signals. With the main difference to the present model being that this approach only makes adoption cascades possible as omission can not be observed, it could be further analyzed whether adding a market environment with payoff externalities similar to the one in Section 3 leads to comparable results in which the interaction effects create a trade-off to adoption that reduces the probability of a firm innovating compared to the baseline without externalities.

Understanding the driving forces behind firms' adoption behavior is a crucial step in analyzing the diffusion of technology and developing mechanisms to prevent the wasteful adoption of unsuccessful technologies. Imitation is a well-documented phenomenon in technology adoption which can arise from a variety of reasons, from publicly known superiority to network effects to cascades and following market-leaders (Rogers, 1974). Imitation is also of particular importance to welfare concerns as it leads to an amplification of the welfare effect of the imitated action. The results from this paper illustrate one mechanism, adoption cascades, through which such imitative behavior can arise, and effects a competitive environment and payoff externalities have on innovation herding and offer insights into how economic policy can be used to improve welfare by reducing the likelihood of undesired cascades.

[26] Sticking to the established notation, for there to be a unique firm willing to adopt at any point in time we require that there is a unique firm with $\beta_i(\mathcal{I}_i) > \bar{\beta}_i$ at this point in time.

References

Acemoglu, D. & P. Restrepo (2018). Modeling automation. *AEA Papers and Proceedings*, 108, 48-53.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The Quarterly Journal of Economics*, 94(4), 749-775.

Anderson, L. R., & C. A. Holt (1997). Information cascades in the laboratory. *The American Economic Review*, 87(5), 847-862.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3), 797-817.

Banta, H. D. (1980). The diffusion of the computed tomography (CT) scanner in the United States. *International Journal of Health Services*, 10(2), 251-269.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, 102(5), 841-877..

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

Bikhchandani, S., D. Hirshleifer & I. Welch (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives*, 12(3), 151-170.

Cameron, K. S. (1998). Strategic organizational downsizing: An extreme case. *Research in Organizational Behavior*, 20, 185-230.

Camilo, A., D. Gökalp, D. Iurchenko, J. Klix & J. Rubinoff (2020). Industrial robots and where to find them: evidence and theory on derobotization. *UPF Repository*, Master's Degree in Economics and Finance, 27.

Çelen, B., & S. Kariv (2004). Distinguishing informational cascades from herd behavior in the laboratory. *The American Economic Review*, 94(3), 484-498.

Chamley, C. (2004). Rational herds: Economic models of social learning. *Cambridge University Press*.

Chamley, C. & D. Gale (1994). Information revelation and strategic delay in a model of investment. *Econometrica*, 62(5), 1065-1085.

Chaudhuri, S., J. Jayartne & A. Chang (1997). Informational Externalities and the Branch Location Decisions of Banks: An Empirical Analysis. Columbia University, Working Paper.

Chatzimichael, K., M. Genius & V. Tzouvelekas (2014). Informational cascades and technology adoption: Evidence from Greek and German organic growers. *Food Policy*, 49, 186-195.

Choi, J. P. (1997). Herd behavior, the "penguin effect," and the suppression of informational diffusion: an analysis of informational externalities and payoff interdependency. *The Rand Journal of Economics*, 407-425.

Dasgupta, A. (2000). Social Learning with Payoff Complementarities. *Econometric Society World Congress Contributed Papers,* (322).

Frey, C. B. & M. A. Osborne (2017). The future of employment: How susceptible are jobs to computerisation?. *Technological forecasting and social change*, 114, 254-280.

Geroski, P. (2000). Models of technology diffusion. *Research policy*, 29(4-5), 603-625.

Gossner, O., J. Steiner & C. Stewart (2018). Attention please!. *University of Zurich, Department of Economics, Working Paper,* (308).

Guarino, A., H. Harmgart & S. Huck (2011). Aggregate information cascades. *Games and Economic Behavior*, 73(1), 167-185.

Humlum, A. (2019). Robot Adoption and Labor Market Dynamics. *Princeton University.*

Kasahara, T. (2015). Strategic technology adoption under dispersed information and information learning. *International Journal of Innovation and Technology Management*, 12(6).

Kennedy, R. (1997). Strategy Fads and Strategic Positioning: An Empirical Test for Herd Behavior in Prime-Time Television Programming. *Harvard Business School, Working Paper.*

Kübler, D. & G. Weizsäcker (2004). Limited depth of reasoning and failure of cascade formation in the laboratory. *The Review of Economic Studies*, 71(2), 425-441.

Lim, Y., A. Ozdaglar & A. Teytelboym (2016). A simple model of cascades in networks. *Working Paper.*

Melissas, N. (2005). Herd behaviour as an incentive scheme. *Economic Theory*, 26(3), 517-536.

Oorschot, J. v., E. Hofman & J. I. Halman (2018). A bibliometric review of the innovation adoption literature. *Technological Forecasting and Social Change*, 134, 1-21.

Rogers, E. M. (1974). Diffusion of Innovations. *Free Press*, Third Edition.

Scharfstein, D. & J. Stein (1990). Herd Behavior and Investment. *The American Economic Review*, 80(3), 465-479.

Shichijo, T. & Y. Nakayama (2009). A two-step subsidy scheme to overcome network externalities in a dynamic game. *The BE Journal of Theoretical Economics*, 9(1), 1-20.

Vergari, C. (2004). Herd behaviour, strategic complementarities and technology adoption. *SSRN Working Paper*, 2004063.

Walden, E. A. & G. J. Browne (2003). Running with the pack: An empirical examination of information cascades in the adoption of novel technologies. *Lubbock: Texas Tech. University.*

Xu, S. X., C. Zhu & K. X. Zhu (2012). Why do firms adopt innovations in bandwagons? Evidence of herd behaviour in open standards adoption. *International Journal of Technology Management*, 59(1/2), 63-91.

Appendix

A.1 Derivation of Belief-Order

For $\beta^1 > p^s$ we require

$$\begin{aligned} \frac{p^s\lambda}{p^s\lambda+(1-p^s)(1-\lambda)} > p^s &\iff \frac{\lambda}{p^s\lambda+(1-p^s)(1-\lambda)} > 1 \\ &\iff \lambda > p^s\lambda+(1-p^s)(1-\lambda) \\ &\iff (1-p^s)\lambda > (1-p^s)(1-\lambda) \\ &\iff \lambda > (1-\lambda) \iff \lambda > \tfrac{1}{2} \end{aligned}$$

Similarly, for $\beta^0 < p^s$ we require

$$\begin{aligned} \frac{p^s(1-\lambda)}{p^s(1-\lambda)+(1-p^s)\lambda} < p^s &\iff \frac{1-\lambda}{p^s(1-\lambda)+(1-p^s)\lambda} < 1 \\ &\iff 1-\lambda < p^s(1-\lambda)+(1-p^s)\lambda \\ &\iff (1-p^s)(1-\lambda) < (1-p^s)\lambda \\ &\iff \lambda > (1-\lambda) \iff \lambda > \tfrac{1}{2} \end{aligned}$$

A.2 Derivation of General Belief and Partitioning

Take an arbitrary sequence of k signals $\Theta_k = (\theta_1, \dots, \theta_k) \in \{0,1\}^k$. Then, the posterior belief $\beta(\Theta_k)$ is given by

$$\begin{aligned} \beta(\Theta_k) = \mathbb{P}[s=H|\Theta] = \frac{\mathbb{P}[s=H]\mathbb{P}[\Theta|s=H]}{P[\Theta]} &= \frac{\mathbb{P}[s=H]\mathbb{P}[\Theta|s=H]}{\mathbb{P}[s=H]\mathbb{P}[\Theta|s=H]+\mathbb{P}[s=L]\mathbb{P}[\Theta|s=L]} \\ &= \frac{p^s\lambda^{\Sigma\theta_i}(1-\lambda)^{\Sigma(1-\theta_i)}}{p^s\lambda^{\Sigma\theta_i}(1-\lambda)^{\Sigma(1-\theta_i)}+(1-p^s)\lambda^{\Sigma(1-\theta_i)}(1-\lambda)^{\Sigma\theta_i}} \end{aligned}$$

Note that the belief comprises two exponents, $\Sigma\theta_i$ and $\Sigma(1-\theta_i)$. Cancel with the smaller of both, i.e., divide both numerator and denominator by $\lambda^{\min(\Sigma\theta_i,\Sigma(1-\theta_i))}(1-\lambda)^{\min(\Sigma\theta_i,\Sigma(1-\theta_i))}$, obtaining

$$\begin{aligned} \beta(\Theta_k) &= \frac{p^s\lambda^{\max(0,2\Sigma\theta_i-k)}(1-\lambda)^{\max(0,k-2\Sigma\theta_i)}}{p^s\lambda^{\max(0,2\Sigma\theta_i-k)}(1-\lambda)^{\max(0,k-2\Sigma\theta_i)}+(1-p^s)\lambda^{\max(0,k-2\Sigma\theta_i)}(1-\lambda)^{\max(0,2\Sigma\theta_i-k)}} \\ &= \frac{p^s\lambda^{\max(0,\Delta_k^\theta)}(1-\lambda)^{-\min(0,\Delta_k^\theta)}}{p^s\lambda^{\max(0,\Delta_k^\theta)}(1-\lambda)^{-\min(0,\Delta_k^\theta)}+(1-p^s)\lambda^{-\min(0,\Delta_k^\theta)}(1-\lambda)^{\max(0,\Delta_k^\theta)}} \\ &= \frac{p^s\lambda^{\theta_k+\max(0,\Delta_{k-1}^\theta)}(1-\lambda)^{(1-\theta_k)-\min(0,\Delta_{k-1}^\theta)}}{p^s\lambda^{\theta_k+\max(0,\Delta_{k-1}^\theta)}(1-\lambda)^{(1-\theta_k)-\min(0,\Delta_{k-1}^\theta)}+(1-p^s)\lambda^{(1-\theta_k)-\min(0,\Delta_{k-1}^\theta)}(1-\lambda)^{\theta_k+\max(0,\Delta_{k-1}^\theta)}} \end{aligned}$$

where $\Delta_k^\theta \coloneqq 2\Sigma\theta_i - k = \Sigma\theta_i - \Sigma(1-\theta_i)$. With that we can categorize signal histories $\Theta_k = (\Theta_{k-1}, \theta_k)$ with respect to the beliefs $\beta(\Theta_{k-1}, \theta_k)$ they induce.
First, as already stated in Section 2, we obtain $\beta(0,1) = \beta^1$ and $\beta(0,0) = \beta^1$. It follows also directly that $\beta(1,0) = \beta(-1,1) = p^s$. With that in mind, consider the three cases presented in Section 2 in order to determine Δ_{k-1}^θ that form part of a cascade, namely such that $\beta(\Theta_{k-1}, 0) > \bar{\beta}$ and $\beta(\Theta_{k-1}, 1) < \bar{\beta}$:

Case 1: $\bar{\beta} = p^s$

Assume $\Delta_k^\theta > 0$, then we can write the posterior $\beta(\Theta_{k-1}, 0)$ as

$$\beta(\Theta_{k-1}, 0) = \frac{p^s \lambda^{\Delta_{k-1}^\theta}(1-\lambda)}{p^s \lambda^{\Delta_{k-1}^\theta}(1-\lambda) + (1-p^s)\lambda(1-\lambda)^{\Delta_{k-1}^\theta}} > p^s$$

$$\iff \lambda^{\Delta_{k-1}^\theta}(1-\lambda) > p^s \lambda^{\Delta_{k-1}^\theta}(1-\lambda) + (1-p^s)\lambda(1-\lambda)^{\Delta_{k-1}^\theta}$$

$$\iff 1 > p^s + (1-p^s)\lambda^{1-\Delta_{k-1}^\theta}(1-\lambda)^{\Delta_{k-1}^\theta - 1}$$

$$\iff 1 > \left(\frac{1-\lambda}{\lambda}\right)^{\Delta_{k-1}^\theta - 1} \implies \Delta_{k-1}^\theta - 1 > 0 \iff \Delta_{k-1}^\theta > 1$$

Thus, every $\Delta_{k-1}^\theta \geq 2$ is part of an adoption cascade. Similarly, for $\Delta_{k-1}^\theta < 0$ we can express $\beta(\Theta_{k-1}, 1)$ as

$$\beta(\Theta_{k-1}, 1) = \frac{p^s \lambda (1-\lambda)^{-\Delta_{k-1}^\theta}}{p^s \lambda(1-\lambda)^{-\Delta_{k-1}^\theta} + (1-p^s)\lambda^{-\Delta_{k-1}^\theta}(1-\lambda)} < p^s$$

$$\iff \lambda(1-\lambda)^{-\Delta_{k-1}^\theta} < p^s \lambda(1-\lambda)^{-\Delta_{k-1}^\theta} + (1-p^s)\lambda^{-\Delta_{k-1}^\theta}(1-\lambda)$$

$$\iff 1 < p^s + (1-p^s)\lambda^{-\Delta_{k-1}^\theta - 1}(1-\lambda)^{\Delta_{k-1}^\theta + 1}$$

$$\iff 1 < \left(\frac{1-\lambda}{\lambda}\right)^{\Delta_{k-1}^\theta + 1} \implies \Delta_{k-1}^\theta + 1 < 0 \iff \Delta_{k-1}^\theta < -1$$

Consequentially, an omission cascades takes place whenever $\Delta_{k-1}^\theta \leq -2$. The last step in obtaining the belief cascade is to account for the inference possible from any observed sequence of actions. Since a cascade starts at $|\Delta_{k-1}^\theta| = 2$, every action in the cascade afterwards is uninformative, which is saying $\Delta_{k-1}^\theta = \min(\Delta_{k-1}^a, 2)$ for $\Delta_{k-1}^a \geq 0$ and $\Delta_{k-1}^\theta = \max(\Delta_{k-1}^a, -2)$ for $\Delta_{k-1}^a \leq 0$. Combined with the general formula we finally obtain

$$\beta(\Delta_{i-1}^a, \theta_i) = \frac{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i) - \mathfrak{d}_i^-}}{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i) - \mathfrak{d}_i^-} + (1-p^s)\lambda^{(1-\theta_i) - \mathfrak{d}_i^-}(1-\lambda)^{\mathfrak{d}_i^+ + \theta_i}}$$

$$\text{where } \mathfrak{d}_i^+ \coloneqq \text{median}(0, \Delta_{i-1}^a, 2) = \max(0, \min(\Delta_{k-1}^a, 2))$$
$$\text{and } \mathfrak{d}_i^- \coloneqq \text{median}(-2, \Delta_{i-1}^a, 0) = \min(0, \max(\Delta_{k-1}^a, -2))$$

which concludes the general belief formula for this case.

Case 2: $\bar{\beta} > p^s$

Consider the amount of excess actions sequentially, starting with $\Delta_k^a \geq 0$.

- $\Delta_k^a = 0$: As stated before, $\beta(0,0) = \beta^0 < p^s < \bar{\beta}$ and $\beta(0,1) = \beta^1 > \bar{\beta}$ and therefore, firms follow their own signal and no cascade is started.
- $\Delta_k^a = 1$: In this case, $\beta(1,0) = p^s < \bar{\beta}$ and $\beta(1,1) > \beta^1 > \bar{\beta}$. Thus, firms continue to follow their signal without prompting a cascade
- $\Delta_k^a = 2$: Now, $\beta(2,0) = \beta(0,1) = \beta^1 > \bar{\beta}$ which prompts an adoption cascade.
- $\Delta_k^a > 2$: Since, as derived above, $\Delta_k^a = 2$ starts a cascade, this implies $\Delta_k^\theta = \min(2, \Delta_k^a) = 2$

On the other hand, for $\Delta_k^a \leq 0$ we obtain

- $\Delta_k^a = 0$: As stated already, $\beta(0,0) = \beta^0 < p^s < \bar{\beta}$ and $\beta(0,1) = \beta^1 > \bar{\beta}$ with firms following their own signal and no cascade.
- $\Delta_k^a = -1$: In this instance, $\beta(-1,0) < \beta^0 < \bar{\beta}$ as well as $\beta(-1,1) = p^s < \bar{\beta}$, which leads to $a_k = 0$ for all θ_k and starts an omission cascade.
- $\Delta_k^a < -1$: Since $\Delta_k^a = -1$ starts a rejection cascade, it has to be $\Delta_k^\theta = \max(-1, \Delta_k^a)$

Combining both we obtain, similar to case 1

$$\beta(\Delta_{i-1}^a, \theta_i) = \frac{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i) - \mathfrak{d}_i^-}}{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i) - \mathfrak{d}_i^-} + (1-p^s)\lambda^{(1-\theta_i) - \mathfrak{d}_i^-}(1-\lambda)^{\mathfrak{d}_i^+ + \theta_i}}$$

but with $\mathfrak{d}_i^+ := \text{median}(0, \Delta_{i-1}^a, 2)$ and $\mathfrak{d}_i^- := \text{median}(-1, \Delta_{i-1}^a, 0)$ instead.

Case 3: $\bar{\beta} < p^s$

Symmetric to case 2, in this case we obtain that when $\Delta_k^a = 1$, $\beta(1,0) = p^s > \bar{\beta}$, prompting an adoption cascade and that for $\Delta_k^a = -2$, $\beta(2,1) = \beta(0,0) = \beta^0 < \bar{\beta}$. Thus, we obtain the general belief as

$$\beta(\Delta_{i-1}^a, \theta_i) = \frac{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i) - \mathfrak{d}_i^-}}{p^s \lambda^{\mathfrak{d}_i^+ + \theta_i}(1-\lambda)^{(1-\theta_i) - \mathfrak{d}_i^-} + (1-p^s)\lambda^{(1-\theta_i) - \mathfrak{d}_i^-}(1-\lambda)^{\mathfrak{d}_i^+ + \theta_i}}$$

with $\mathfrak{d}_i^+ := \text{median}(0, \Delta_{i-1}^a, 1)$ and $\mathfrak{d}_i^- := \text{median}(-2, \Delta_{i-1}^a, 0)$.

A.3 Public Signals

To obtain the action of each firm k in the case of publicly observable signals, consider the general belief formula derived in A.2 for a sequence of observed signals Θ_k with $\Delta_k^\theta := 2\Sigma\theta_i - k$

$$\beta(\Theta_k) = \frac{p^s \lambda^{\max(0,\Delta_k^\theta)}(1-\lambda)^{-\min(0,\Delta_k^\theta)}}{p^s \lambda^{\max(0,\Delta_k^\theta)}(1-\lambda)^{-\min(0,\Delta_k^\theta)} + (1-p^s)\lambda^{-\min(0,\Delta_k^\theta)}(1-\lambda)^{\max(0,\Delta_k^\theta)}}$$

Now, whenever $\Delta_k^\theta \geq 1$ (i.e., $\mathrm{mod}(\Theta_k) = 1$) the corresponding belief is given by

$$\beta(\Theta_k) = \frac{p^s \lambda^{\Delta_k^\theta}}{p^s \lambda^{\Delta_k^\theta} + (1-p^s)(1-\lambda)^{\Delta_k^\theta}} > \frac{p^s \lambda}{p^s \lambda + (1-p^s)(1-\lambda)} = \beta^1 \implies a_k = 1$$

Conversely, for $\Delta_k^\theta \leq -1$ (i.e., $\mathrm{mod}(\Theta_k) = 0$) the corresponding belief is given by

$$\beta(\Theta_k) = \frac{p^s (1-\lambda)^{\Delta_k^\theta}}{p^s (1-\lambda)^{\Delta_k^\theta} + (1-p^s)\lambda^{\Delta_k^\theta}} < \frac{p^s (1-\lambda)}{p^s (1-\lambda) + (1-p^s)\lambda} = \beta^0 \implies a_k = 0$$

Lastly, if $\Delta_k^\theta = 0$ we get $\beta(\Theta_k) = p^s$ prompting $a_k = \theta_k$, $a_k = 0$ or $a_k = 1$ in case 1,2 or 3 respectively.

A.4 Derivation of Absence Threshold

For a firm i to stop adopting regardless of not only its private signal but any possible observable history, it has to be willing to do so holding the most affirmative belief possible, that is, in an adoption cascade. Assuming that the belief threshold of the directly preceding firm $\bar{\beta}_{i-1} > p^s$, an adoption cascades takes place for $\Delta_{i-1}^a \geq 2$. The highest belief is obtained, when additionally $\theta_i = 1$ yielding the belief

$$\beta(\Delta_{i-1}^a, 1) = \frac{p^s \lambda^{2+1}}{p^s \lambda^{2+1} + (1-p^s)(1-\lambda)^{2+1}} = \frac{p^s \lambda^3}{p^s \lambda^3 + (1-p^s)(1-\lambda)^3} =: \beta^\times$$

Thus, whenever $\bar{\beta}_i > \beta^\times$, firm i will not innovate under any circumstances.

A.5 Signal Coupling

This section explores the effect of the adoption decision on subsequent decisions under the assumption of coupled signals and fixed strategies of competitors. Start from an arbitrary configuration of the baseline model. Take any sequence of signals $(\theta_1, \ldots, \theta_n)$ and a firm i. The signal sequence completely determines all adoption decisions $(a_1, \ldots, a_{i-1})$, which together with θ_i would also determine a_i^b in the baseline model. Now consider i's decision between $a_i = 1$ and $a_i = 0$ and its effect on subsequent adoption decisions. As shown in A.2, each threshold $\bar{\beta}_j$ in the baseline model corresponds to an upper and lower limit on the aggregate number of actions, Δ_{j-1}^a starting an adoption or rejection cascade respectively. Assume that before i takes action there are an aggregate Δ_{i-1}^a actions. The baseline strategy of i can thus be of three forms, depending of the value of the aggregate actions: $a_i^b = 1$, $a_i^b = \theta_i$ or $a_i^b = 0$. Depending on the realization of the signal sequence, four cases are possible that are summarized below, with Δ_i^a denoting the aggregate actions in the baseline case and $\bar{\Delta}_i^a$ when playing a_i instead:

	$\theta_i = 1, a_i = 1$	$\theta_i = 0, a_i = 1$	$\theta_i = 1, a_i = 0$	$\theta_i = 1, a_i = 0$
$a_i^b = 1$	$\bar{\Delta}_i^a = \Delta_i^a$	$\bar{\Delta}_i^a = \Delta_i^a$	$\bar{\Delta}_i^a < \Delta_i^a$	$\bar{\Delta}_i^a < \Delta_i^a$
$a_i^b = 1$	$\bar{\Delta}_i^a = \Delta_i^a$	$\bar{\Delta}_i^a > \Delta_i^a$	$\bar{\Delta}_i^a < \Delta_i^a$	$\bar{\Delta}_i^a = \Delta_i^a$
$a_i^b = 1$	$\bar{\Delta}_i^a > \Delta_i^a$	$\bar{\Delta}_i^a > \Delta_i^a$	$\bar{\Delta}_i^a = \Delta_i^a$	$\bar{\Delta}_i^a = \Delta_i^a$
Total	$\bar{\Delta}_i^a \geq \Delta_i^a$	$\bar{\Delta}_i^a \geq \Delta_i^a$	$\bar{\Delta}_i^a \leq \Delta_i^a$	$\bar{\Delta}_i^a \leq \Delta_i^a$

In conclusion, we can observe that for any sequence of $(\theta_1, \ldots, \theta_{i-1})$ and any θ_i the number of aggregate actions including a_i follows $\bar{\Delta}_i^a(a_i = 1) \geq \Delta_i^a \geq \bar{\Delta}_i^a(a_i = 0)$ which is equivalent to saying $A_{-i}(a_i = 1) \geq A_{-i}(a_i = 0)$.

A.6 Derivation of Market Interaction Threshold

Denote the expected interaction effect V_i given a history of previous actions $(a_1, \ldots, a_{i-1})$ by $\mathbb{E}_A[V_i|a_i, s] = \mathbb{E}[V_i(A_{-i}, s)|a_1, \ldots, a_i, s]$. Then the expected profits are given by

$$\begin{aligned}\mathbb{E}_A[\Pi_i|a_i = 0] &= \mathbb{E}[\Pi_i|a_1, \ldots, a_{i-1}, a_i = 0] = \pi(0, s) - \beta\gamma\mathbb{E}_A[V_i|0, H] - (1-\beta)\gamma\mathbb{E}_A[V_i|0, L] \\ &= \pi(0, s) - \gamma\mathbb{E}_A[V_i|0, L] - \beta\gamma(\mathbb{E}_A[V_i|0, H] - \mathbb{E}_A[V_i|0, L])\end{aligned}$$

$$\begin{aligned}\mathbb{E}_A[\Pi_i|a_i = 1] &= \beta(\pi_i(1, H) - \gamma\mathbb{E}_A[V_i|1, H]) + (1-\beta)(\pi(1, L) - \gamma\mathbb{E}_A[V_i|1, L]) \\ &= \pi(1, L) + \beta(\pi_i(1, H) - \pi(1, L)) - \gamma\mathbb{E}_A[V_i|1, L] - \beta\gamma(\mathbb{E}_A[V_i|1, H] - \mathbb{E}_A[V_i|1, L])\end{aligned}$$

Define $\Delta V_i^s \coloneqq \mathbb{E}_A[V_i|1, s] - \mathbb{E}_A[V_i|0, s]$, the threshold $\bar{\beta}_i^M$ satisfies $\mathbb{E}_A[\Pi_i|a_i = 0] = \mathbb{E}_A[\Pi_i|a_i = 1]$, yielding

$$\begin{aligned}&\pi(1, L) + \beta(\pi_i(1, H) - \pi(1, L)) - \gamma\mathbb{E}_A[V_i|1, L] - \beta\gamma(\mathbb{E}_A[V_i|1, H] - \mathbb{E}_A[V_i|1, L]) \\ &\qquad = \pi(0, s) - \gamma\mathbb{E}_A[V_i|0, L] - \beta\gamma(\mathbb{E}_A[V_i|0, H] - \mathbb{E}_A[V_i|0, L])\end{aligned}$$

$$\begin{aligned}&\beta(\pi_i(1, H) - \pi(1, L)) - \beta\gamma(\mathbb{E}_A[V_i|1, H] - \mathbb{E}_A[V_i|1, L]) + \beta\gamma(\mathbb{E}_A[V_i|0, H] - \mathbb{E}_A[V_i|0, L]) \\ &\qquad = \pi(0, s) - \pi(1, L) - \gamma\mathbb{E}_A[V_i|0, L] + \gamma\mathbb{E}_A[V_i|1, L]\end{aligned}$$

$$\begin{aligned}&\beta(\pi_i(1, H) - \pi(1, L)) - \beta\gamma(\mathbb{E}_A[V_i|1, H] - \mathbb{E}_A[V_i|0, H]) + \beta\gamma(\mathbb{E}_A[V_i|1, L] - \mathbb{E}_A[V_i|0, L]) \\ &\qquad = \pi(0, s) - \pi(1, L) + \gamma\Delta V_i^L\end{aligned}$$

$$\pi(0,s) - \pi(1,L) + \gamma\Delta V_i^L = \beta(\pi_i(1,H) - \pi(1,L)) - \beta\gamma\Delta V_i^H + \beta\gamma\Delta V_i^L$$
$$= \beta(\pi_i(1,H) - \pi(1,L) - \gamma\Delta V_i^H + \gamma\Delta V_i^L)$$

$$\implies \bar{\beta}_i^M(\gamma) \coloneqq \frac{\pi(0,s) - \pi(1,L) + \gamma\Delta V_i^L}{\pi_i(1,H) - \pi(1,L) - \gamma\Delta V_i^H + \gamma\Delta V_i^L} = \frac{\pi(0,s) - \pi(1,L) + \gamma\Delta V_i^L}{\pi_i(1,H) - \pi(1,L) + \gamma(\Delta V_i^L - \Delta V_i^H)}$$

A.7 Proof of Proposition 1

Before proving Proposition 1, the following lemmata are established:

Lemma 1. Take an arbitrary fraction $\frac{a}{b} < 1$ and $c > 0$. Then $\frac{a+c}{b+c} > \frac{a}{b}$.

Proof. $\frac{a}{b} = \frac{a(1+c/b)}{b(1+c/b)} = \frac{a+a/b\cdot c}{b+c} < \frac{a+c}{b+c}$, since $a + \frac{a}{b}c < a + c$. □

Lemma 2. $\Delta V_i^s > 0$.

Proof. As stated in A.5 it has to be that $A_{-i}(a_i = 1) \geq A_{-i}(a_i = 0)$ for any signal sequence $\Theta = (\theta_1, \ldots, \theta_n)$. Define an event space $\Omega = \{0,1\}^n \times \{H, L\}$ and an identical event space $\mathcal{F} = \Omega$. Lastly, define the probability function P as

$$P(\omega) = P(\theta_1, \ldots, \theta_n, s) = \mathbf{1}_s(H)p^s\lambda^{\Sigma\theta_i}(1-\lambda)^{\Sigma(1-\theta_i)} + \mathbf{1}_s(L)(1-p^s)\lambda^{\Sigma(1-\theta_i)}(1-\lambda)^{\Sigma\theta_i}$$

Then we can define the random variables $X_i^1 : \Omega \to \mathbb{R}^+$ with $X_i^1(\omega) = V_i(a_i = 1, \theta_1, \ldots, \theta_n, s)$ and $X_i^0 : \Omega \to \mathbb{R}^+$ with $X_i^0(\omega) = V_i(a_i = 0, \Theta, s)$. Since $A_{-i}(a_i = 1) \geq A_{-i}(a_i = 0)$ for any and $\partial V_i(\cdot)/\partial A_{-i} \geq 0$, it follows directly that for any $\omega \in \Omega$, $X_i^1(\omega) \geq X_i^0(\omega)$. Thus, X_i^1 first-order stochastically dominates X_i^0.
Lastly, for any $A \subset \Omega$ it is true that $\mathbb{E}[X_i^1|A] > \mathbb{E}[X_i^0|A]$, since

$$\mathbb{E}[X_i^1|A] = \sum_{\Omega} X_i^1(\omega)\frac{P(\omega \cap A)}{P(A)} > \sum_{\Omega} X_i^0(\omega)\frac{P(\omega \cap A)}{P(A)} = \mathbb{E}[X_i^0|A]$$

Therefore, for any $(a_1, \ldots, a_{i-1}, s)$ define

$$C = \{\theta_1, \ldots, \theta_{i-1} : (a_1(\theta_1), \ldots, a_{i-1}(\theta_1, \ldots, \theta_{i-1}) = (a_1, \ldots, a_{i-1}, s)\} \cup \{s\}$$

Then $\mathbb{E}_A[V_i|a_i, s] = \mathbb{E}[X_i^{a_i}|C]$ and thus $\Delta V_i^s \coloneqq \mathbb{E}_A[V_i|1,s] - \mathbb{E}_A[V_i|0,s] = \mathbb{E}[X_i^1|C] - \mathbb{E}[X_i^0|C] > 0$. □

Now consider the threshold $\bar{\beta}_i^M$ derived in A.6 and assume $\gamma > 0$. By Lemma 2, $\Delta V_i^H > 0$, thus

$$\bar{\beta}_i^M(\gamma) = \frac{\pi(0,s) - \pi(1,L) + \gamma\Delta V_i^L}{\pi_i(1,H) - \pi(1,L) + \gamma(\Delta V_i^L - \Delta V_i^H)} > \frac{\pi(0,s) - \pi(1,L) + \gamma\Delta V_i^L}{\pi_i(1,H) - \pi(1,L) + \gamma\Delta V_i^L}$$

By Lemma 2, also $\Delta V_i^H > 0$ and thus by Lemma 1

$$\bar{\beta}_i^M(\gamma) > \frac{\pi(0,s) - \pi(1,L) + \gamma \Delta V_i^L}{\pi_i(1,H) - \pi(1,L) + \gamma \Delta V_i^L} > \frac{\pi(0,s) - \pi(1,L)}{\pi_i(1,H) - \pi(1,L)} = \bar{\beta}_i$$

concluding that $\zeta_i(\gamma) = \bar{\beta}_i^M(\gamma) - \bar{\beta}_i > 0$ for $\gamma > 0$. Conversely, for $\gamma < 0$, by the same arguments

$$\bar{\beta}_i^M(\gamma) < \frac{\pi(0,s) - \pi(1,L) + \gamma \Delta V_i^L}{\pi_i(1,H) - \pi(1,L) + \gamma \Delta V_i^L} < \frac{\pi(0,s) - \pi(1,L)}{\pi_i(1,H) - \pi(1,L)} = \bar{\beta}_i$$

which implies that $\zeta_i(\gamma) = \bar{\beta}_i^M(\gamma) - \bar{\beta}_i < 0$ for $\gamma < 0$. For the last part, I will show that $\zeta_i(\gamma) > \zeta_i(\gamma')$ for $\gamma > \gamma'$, which is equivalent to showing $\bar{\beta}_i^M(\gamma) > \bar{\beta}_i^M(\gamma')$

$$\begin{aligned}
\bar{\beta}_i^M(\gamma) &= \frac{\pi(0,s) - \pi(1,L) + \gamma \Delta V_i^L}{\pi_i(1,H) - \pi(1,L) - \gamma \Delta V_i^H + \gamma \Delta V_i^L} > \frac{\pi(0,s) - \pi(1,L) + \gamma \Delta V_i^L}{\pi_i(1,H) - \pi(1,L) - \gamma' \Delta V_i^H + \gamma \Delta V_i^L} \\
&= \frac{\pi(0,s) - \pi(1,L) + \gamma' \Delta V_i^L + \alpha \Delta V_i^L}{\pi_i(1,H) - \pi(1,L) - \gamma' \Delta V_i^H + \gamma' \Delta V_i^L + \alpha \Delta V_i^L} \quad \text{where } \gamma = \gamma' + \alpha \text{ with } \alpha > 0 \\
&> \frac{\pi(0,s) - \pi(1,L) + \gamma' \Delta V_i^L}{\pi_i(1,H) - \pi(1,L) - \gamma' \Delta V_i^H + \gamma' \Delta V_i^L} = \bar{\beta}_i^M(\gamma')
\end{aligned}$$

concluding the proof that $\bar{\beta}_i^M(\gamma)$ increases in γ.

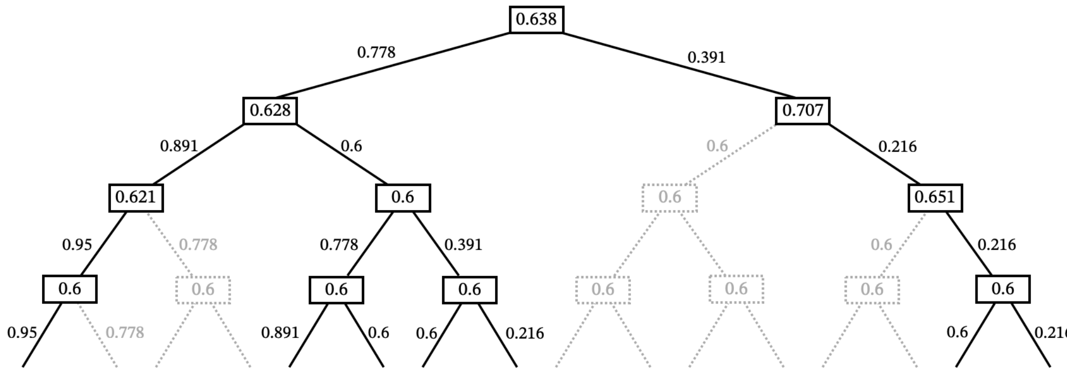
A.8 Simulation Methodology

Using R, I simulate $N = 10^7$ states s from a binomial distribution with $p^s = 0.6$ together with signal sequences $(\theta_1, \ldots, \theta_4)$ as independent binomial variables with $\lambda = 0.7$, and randomly assigned actions a_i.

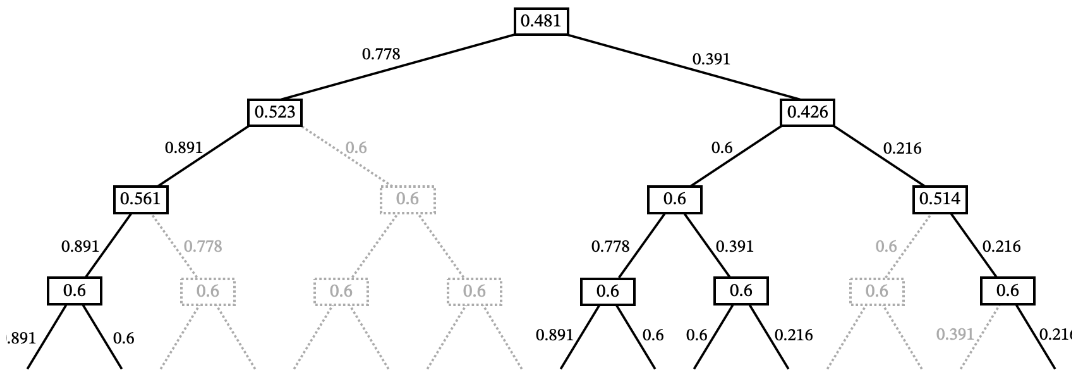
The profit function $\Pi_i(a_i, A_{-i}, s) = \pi(a_i, s) - \gamma V_i(A_{-i}, s)$ is parameterized with

$$\pi(a_i, s) = \begin{cases} 60 & \text{for } (1, H) \\ 10 & \text{for } (1, L) \\ 40 & \text{for } (0, s) \end{cases} \qquad \text{and} \quad V_i(A_{-i}, s) = \begin{cases} 10A_{-i} & \text{if } s = H \\ 2A_{-i} & \text{if } s = L \end{cases}$$

The beliefs held at each branch can then be estimated by the mean state $\bar{s}(\Theta, A)$ after conditioning on the corresponding sequence of signals and actions (Θ, A). Similarly, expected profits at each decision node can be estimated using the beliefs obtained prior and the profit function described above. Finally, adoption thresholds are obtained by finding the beliefs that would make each firm indifferent between adopting and omitting for each decision using a grid of beliefs β with precision $\Delta\beta = 10^{-4}$.



(a) Competitive model with weaker substitutes ($\gamma = 0.5$).



(b) Competitive model with complements ($\gamma = -1$).

Figure 10: Structure of two competitive models between $n = 4$ firms. The bordered numbers at each node are the adoption thresholds $\bar{\beta}_i^M$, obtained via simulation (see Appendix A.8), the numbers at the left (right) branches the belief held by the corresponding firm after observing $\theta_i = 1$ ($\theta_i = 0$). Movement along left (right) branches corresponds to $a_i = 1$ ($a = 0$) with grayed out portions representing never-played actions due to cascading behavior.

A.9 Additional Examples

This section provides two additional examples. Together with Figure 2 ($\gamma = 0$) and Figure 4 ($\gamma = 1$), Figure 10 ($\gamma = 0.5$ and $\gamma = -1$) illustrates the different characteristics of $\zeta_i(\gamma)$ outlined by Proposition 1 and proven in Appendix A.7.

A.10 Ex-Post Cascade Probabilities

This section derives the probabilities of ending up in the *correct* (adoption when $s = H$ and vice versa) and *incorrect* (conversely) cascade case by case. Figure 11 replicates Figure 1, illustrating the probability structure of all possible cases.

Case 1: $\bar{\beta} = p^s$

From any node with $\Delta_{i-1}^a = 0$ a correct cascade is started with probability λ^2 and an incorrect with probability $(1-\lambda)^2$. Assuming n to be even, the probability of reaching a cascade after exactly n firms made a decision is given by $\lambda^2[2\lambda(1-\lambda)]^{n/2-1}$ or $(1-\lambda)^2[2\lambda(1-\lambda)]^{n/2-1}$ respectively. The

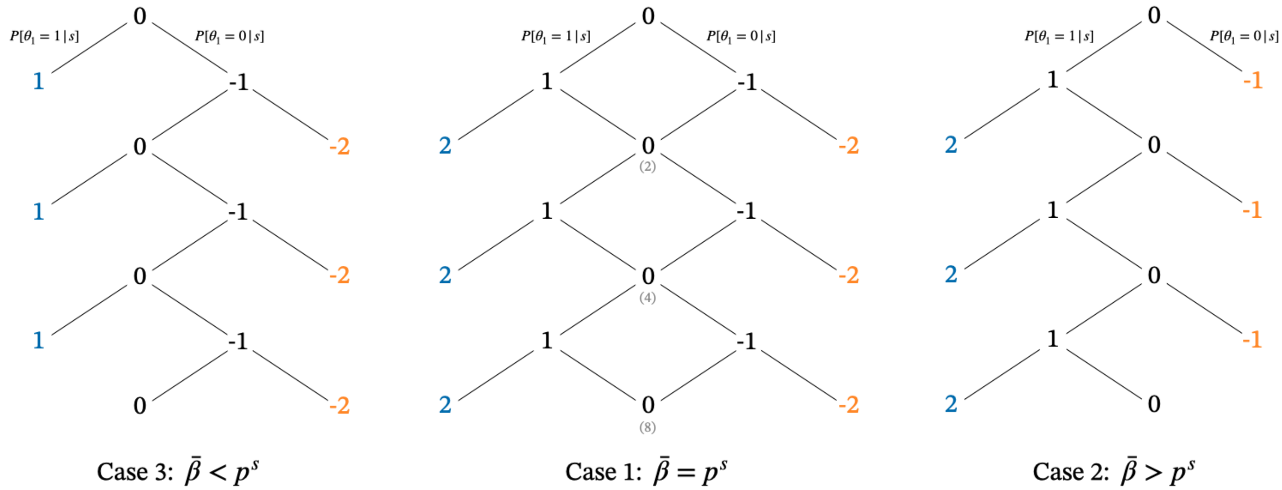


Figure 11: Decision tree of the baseline model for each case. Each node, featuring the excess actions Δ^a_{i-1}, represents the decision made by a firm, with the left (right) branches representing the action when $\theta_i = 1$ ($\theta_i = 0$). Blue (orange) nodes denote the start of an adoption (rejection) cascade and the numbers in brackets the amount of paths leading to a specific node.

probabilities of being in a cascade after n decisions or remaining in no cascade are thus given by

$$p^{\text{cor}} = \lambda^2 \sum_{i=1}^{n/2} [2\lambda(1-\lambda)]^{i-1} = \lambda^2 \frac{1 - [2\lambda(1-\lambda)]^{n/2}}{1 - 2\lambda(1-\lambda)}$$

$$p^{\text{inc}} = (1-\lambda)^2 \sum_{i=1}^{n/2} [2\lambda(1-\lambda)]^{i-1} = (1-\lambda)^2 \frac{1 - [2\lambda(1-\lambda)]^{n/2}}{1 - 2\lambda(1-\lambda)}$$

$$p^{\text{noc}} = [2\lambda(1-\lambda)]^{n/2}$$

since $\lambda(1-\lambda) \leq 1/4$. Thus, in the limit ($n \to \infty$) the probabilities of ending up in each cascade are

$$p^{\text{cor}} = \frac{\lambda^2}{1 - 2\lambda(1-\lambda)} \qquad p^{\text{inc}} = \frac{(1-\lambda)^2}{1 - 2\lambda(1-\lambda)} \qquad \text{and} \qquad p^{\text{noc}} = 0$$

Case 2/3: $\bar{\beta} \neq p^s$

In the second and third case, the probabilities need to account for the structural asymmetry in addition to the realized state s. The probability for the facilitated cascades (rejection in case 2 and adoption in case 3) starting from any node with $\Delta^a_{i-1} = 0$ is given by λ (correct) and $(1-\lambda)$ (incorrect). Starting a cascade after exactly n firms happens with $\lambda[\lambda(1-\lambda)]^{n/2-1}$ or $(1-\lambda)[\lambda(1-\lambda)]^{n/2-1}$, and thus the probabilities of being in a cascade after n decisions is given by

$$p^{\text{cor}} = \lambda \sum_{i=1}^{n/2} [\lambda(1-\lambda)]^{i-1} = \lambda \frac{1 - [\lambda(1-\lambda)]^{n/2}}{1 - \lambda(1-\lambda)}$$

$$p^{\text{inc}} = (1-\lambda) \sum_{i=1}^{n/2} [\lambda(1-\lambda)]^{i-1} = (1-\lambda) \frac{1 - [\lambda(1-\lambda)]^{n/2}}{1 - \lambda(1-\lambda)}$$

with the limiting probabilities

$$p^{\text{cor}} = \frac{\lambda}{1 - \lambda(1-\lambda)} \qquad \text{and} \qquad p^{\text{inc}} = \frac{1-\lambda}{1 - \lambda(1-\lambda)}$$

In contrast, the non-facilitated cascades (adoption in case 2 and rejection in case 3) start from any node with $\Delta^a_{i-1} = 0$ with probabilities λ^2 (correct) and $(1-\lambda)^2$ (incorrect), and after exactly n decisions with $\lambda^2[\lambda(1-\lambda)]^{n/2-1}$ or $(1-\lambda)^2[\lambda(1-\lambda)]^{n/2-1}$. The probabilities of being in a cascade or in no cascade after n decisions is then given by

$$p^{\text{cor}} = \lambda^2 \sum_{i=1}^{n/2} [\lambda(1-\lambda)]^{i-1} = \lambda^2 \frac{1 - [\lambda(1-\lambda)]^{n/2}}{1 - \lambda(1-\lambda)}$$

$$p^{\text{inc}} = (1-\lambda)^2 \sum_{i=1}^{n/2} [\lambda(1-\lambda)]^{i-1} = (1-\lambda)^2 \frac{1 - [\lambda(1-\lambda)]^{n/2}}{1 - \lambda(1-\lambda)}$$

$$p^{\text{noc}} = [\lambda(1-\lambda)]^{n/2}$$

with the limiting probabilities

$$p^{\text{cor}} = \frac{\lambda^2}{1 - \lambda(1-\lambda)} \qquad p^{\text{inc}} = \frac{(1-\lambda)^2}{1 - \lambda(1-\lambda)} \qquad \text{and} \qquad p^{\text{noc}} = 0$$

A.11 Ex-Ante Cascade Probabilities

Deriving the probabilities of ending up in either an adoption or rejection cascade before the realization of the state s differs between cases, combining ex-post probabilities derived in A.10 with probabilities of the state-space.

Case 1: $\bar{\beta} = p^s$

The probability of ending up in an adoption or rejection cascade after n firms are given by

$$p^{\text{ad}} = p^s p^{\text{cor}} + (1-p^s) p^{\text{inc}} = [p^s \lambda^2 + (1-p^s)(1-\lambda)^2] \frac{1 - [2\lambda(1-\lambda)]^{n/2}}{1 - 2\lambda(1-\lambda)}$$

$$p^{\text{re}} = p^s p^{\text{inc}} + (1-p^s) p^{\text{cor}} = [p^s (1-\lambda)^2 + (1-p^s)\lambda^2] \frac{1 - [2\lambda(1-\lambda)]^{n/2}}{1 - 2\lambda(1-\lambda)}$$

$$p^{\text{no}} = p^{\text{noc}} = [2\lambda(1-\lambda)]^{n/2}$$

which yields the limiting probabilities

$$p^{\text{ad}} = \frac{p^s\lambda^2 + (1-p^s)(1-\lambda)^2}{1-2\lambda(1-\lambda)} \qquad p^{\text{re}} = \frac{p^s(1-\lambda)^2 + (1-p^s)\lambda^2}{1-2\lambda(1-\lambda)} \qquad \text{and} \qquad p^{\text{no}} = 0$$

Case 2: $\bar{\beta} > p^s$

The probability of ending up in an adoption or rejection cascade after n firms are given by

$$p^{\text{ad}} = p^s p^{\text{cor}} + (1-p^s)p^{\text{inc}} = [p^s\lambda^2 + (1-p^s)(1-\lambda)^2]\frac{1-[\lambda(1-\lambda)]^{n/2}}{1-\lambda(1-\lambda)}$$

$$p^{\text{re}} = p^s p^{\text{inc}} + (1-p^s)p^{\text{cor}} = [p^s(1-\lambda) + (1-p^s)\lambda]\frac{1-[\lambda(1-\lambda)]^{n/2}}{1-\lambda(1-\lambda)}$$

$$p^{\text{no}} = p^{\text{noc}} = [\lambda(1-\lambda)]^{n/2}$$

which yields the limiting probabilities

$$p^{\text{ad}} = \frac{p^s\lambda^2 + (1-p^s)(1-\lambda)^2}{1-\lambda(1-\lambda)} \qquad p^{\text{re}} = \frac{p^s(1-\lambda) + (1-p^s)\lambda}{1-\lambda(1-\lambda)} \qquad \text{and} \qquad p^{\text{no}} = 0$$

Case 3: $\bar{\beta} < p^s$

The probability of ending up in an adoption or rejection cascade after n firms are given by

$$p^{\text{ad}} = p^s p^{\text{cor}} + (1-p^s)p^{\text{inc}} = [p^s\lambda + (1-p^s)(1-\lambda)]\frac{1-[\lambda(1-\lambda)]^{n/2}}{1-\lambda(1-\lambda)}$$

$$p^{\text{re}} = p^s p^{\text{inc}} + (1-p^s)p^{\text{cor}} = [p^s(1-\lambda)^2 + (1-p^s)\lambda^2]\frac{1-[\lambda(1-\lambda)]^{n/2}}{1-\lambda(1-\lambda)}$$

$$p^{\text{no}} = p^{\text{noc}} = [\lambda(1-\lambda)]^{n/2}$$

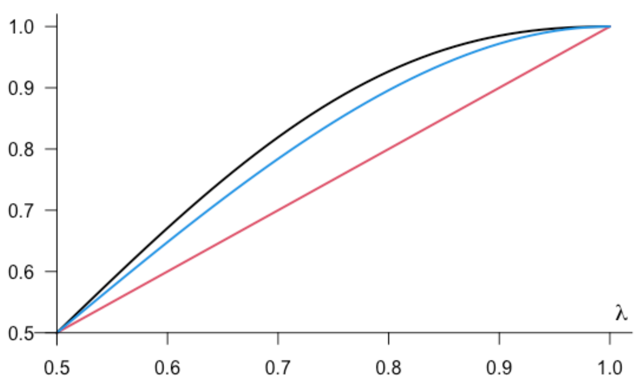
which yields the limiting probabilities

$$p^{\text{ad}} = \frac{p^s\lambda + (1-p^s)(1-\lambda)}{1-\lambda(1-\lambda)} \qquad p^{\text{re}} = \frac{p^s(1-\lambda)^2 + (1-p^s)\lambda^2}{1-\lambda(1-\lambda)} \qquad \text{and} \qquad p^{\text{no}} = 0$$

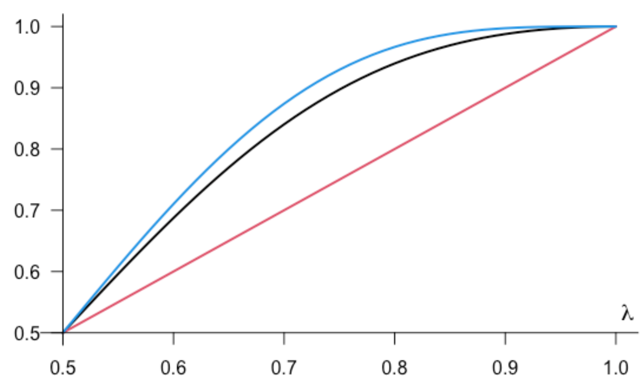
A.12 Additional Results

In addition to the results for the probability of an correct cascade emerging provided in Figure 5 in Section 4, Figure 12 contains an extensive overview of the cascade likelihoods for different market sizes. The only substantial differences is the convergence of the probability of correct actions in the independent model (blue line) with the increasing number of firms.

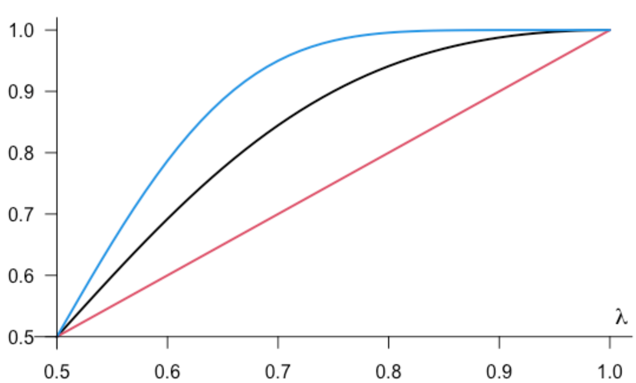
If the number of firms is small enough (Figure 12a, with $n = 4$ firms) it can be seen that the possibility of a correct cascade emerging is more powerful in prompting correct actions than the convergence of signals towards the true state s. As soon as the market size increases, the relationship reverses to the result expected in general.



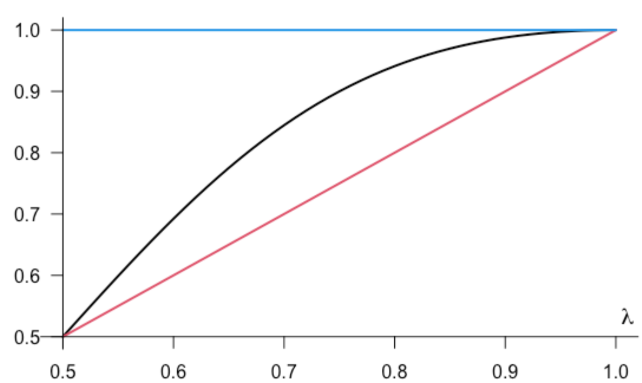
(a) Probabilities for $n = 4$.



(b) Probabilities for $n = 8$..



(c) Probabilities for $n = 16$..



(d) Probabilities for $n \to \infty$..

Figure 12: Probability of ending up in the correct cascade (black) in comparison to the probability of the last firm playing the correct action $a_n = \mathbf{1}_s(H)$ for independent firms (red) and observable signals (blue).