

Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Cowles Foundation Discussion Papers

Cowles Foundation

3-3-2015

Innovation Adoption by Forward-Looking Social Learners

Mira Frick
Yale University

Yuhta Ishii

Follow this and additional works at: <https://elischolar.library.yale.edu/cowles-discussion-paper-series>



Part of the [Economics Commons](#)

Recommended Citation

Frick, Mira and Ishii, Yuhta, "Innovation Adoption by Forward-Looking Social Learners" (2015). *Cowles Foundation Discussion Papers*. 2561.

<https://elischolar.library.yale.edu/cowles-discussion-paper-series/2561>

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

INNOVATION ADOPTION BY FORWARD-LOOKING SOCIAL LEARNERS

By

Mira Frick and Yuhta Ishii

February 2015

Revised August 2020

COWLES FOUNDATION DISCUSSION PAPER NO. 1988R



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

<http://cowles.yale.edu/>

Innovation Adoption by Forward-Looking Social Learners

Mira Frick Yuhta Ishii*

August 11, 2020

Abstract

We build a model studying the effect of an economy’s potential for social learning on the adoption of innovations of uncertain quality. Provided consumers are forward-looking (i.e., recognize the value of waiting for information), we show how quantitative and qualitative features of the learning environment affect observed adoption dynamics, welfare, and the speed of learning. Our analysis has two main implications. First, we identify environments that are subject to a *saturation effect*, whereby increased opportunities for social learning can slow down adoption and learning and do not increase consumer welfare, possibly even being harmful. Second, we show how differences in the learning environment translate into observable differences in adoption dynamics, suggesting a purely informational channel for two commonly documented adoption patterns—S-shaped and concave curves.

JEL codes: D81, D83, O33

Keywords: Innovation adoption; social learning; informational free-riding; strategic experimentation; exponential bandits.

1 Introduction

Suppose a new product of uncertain quality, such as a novel medical procedure or a new movie, is released into the market. In recent years, the rise of internet-based review sites, retail platforms, search engines, video-sharing websites, and social networking sites has greatly

*First version: October 2013. Frick: Yale University (mira.frick@yale.edu); Ishii: Pennsylvania State University (yxi5014@psu.edu). This paper is a revised version of chapters of our PhD dissertations at Harvard University. We are grateful to Drew Fudenberg, Attila Ambrus, Eric Maskin, and Tomasz Strzalecki for generous advice and encouragement. We also thank Nageeb Ali, Dirk Bergemann, Simon Board, Aislinn Bohren, Yeon-Koo Che, Thomas Covert, Martin Cripps, Ben Golub, Marina Halac, Johannes Hörner, Ryota Iijima, Boyan Jovanovic, Daniel Keniston, Daria Khromenkova, Scott Kominers, David Laibson, Greg Lewis, Chiara Margaria, César Martinelli, Stephen Morris, Giuseppe Moscarini, Pauli Murto, Aniko Öry, Luciano Pomatto, Sven Rady, Larry Samuelson, Heather Schofield, Jesse Shapiro, Ron Siegel, Andy Skrzypacz, Caroline Thomas, Jean Tirole, Chris Udry, Juuso Välimäki, Leeat Yariv, and numerous seminar audiences.

increased the *potential for social learning* in the economy: If a patient suffers a serious complication or a movie-goer has a positive viewing experience, this is more likely than ever to find its way into the public domain; and there are more and more people who have access to this common pool of consumer-generated information.

This paper builds a model studying the effect of an economy’s potential for social learning on the adoption of innovations of uncertain quality. We analyze how consumers’ informational incentives depend on quantitative and qualitative features of the learning environment, and how this affects observed adoption dynamics, welfare, and the speed of learning. Our analysis has two main implications. First, quantitatively, we suggest caution in evaluating the impact of increases in the potential for social learning: We identify environments that are subject to a *saturation effect*, whereby beyond a certain level, increased opportunities for social learning can slow down adoption and learning and do not improve consumer welfare, possibly even being harmful. Second, at a qualitative level, we show how differences in the learning environment translate into observable differences in adoption dynamics: This implies a new, purely informational channel for two of the most commonly documented adoption patterns—S-shaped and concave curves.

A central ingredient of our model is that consumers are *forward-looking* social learners. In choosing whether to adopt an innovation, forward-looking consumers recognize the option value of waiting for more information. With social learning, this information is created *endogenously*, based on the consumption experiences of other adopters. Equilibrium adoption dynamics must then resolve the following tension: If too many consumers adopt at any given time, then the expected amount of future information might be so great that all consumers would in fact strictly prefer to wait; conversely, if too few consumers adopt, it might not be worthwhile for anyone to wait. This tension depends non-trivially on the ease and nature of information transmission and is the source of the results of the preceding paragraph.

Forward-looking social learning is well-documented empirically, notably in the development economics literature studying the adoption of agricultural and health innovations.¹ However, its informational ramifications have largely remained unexplored theoretically: Existing learning-based models of innovation adoption typically assume either that learning is

¹Studies of social learning in this domain include [Besley and Case \(1993, 1994\)](#); [Foster and Rosenzweig \(1995\)](#); [Kremer and Miguel \(2007\)](#); [Conley and Udry \(2010\)](#); [Dupas \(2014\)](#). There is also evidence for *forward-looking* social learning: [Bandiera and Rasul \(2006\)](#) analyze the decision of farmers in Mozambique to adopt a new crop, sunflower. They find that if a farmer’s network of friends and family contains many adopters of the new crop, knowing one *more* adopter may make him *less* likely to initially adopt it himself. [Munshi \(2004\)](#) compares farmers’ willingness to experiment with new high-yield varieties (HYV) across rice and wheat growing areas in India. Farmers in rice growing regions, which compared with wheat growing regions display greater heterogeneity in growing conditions that make learning from others’ experiences *less* feasible, are found to be *more* likely to experiment with HYV than farmers in wheat growing areas.

social but consumers are myopic (e.g., Young, 2009), or that consumers are forward-looking but information arrives exogenously (e.g., Jensen, 1982). In either case, the dependence on the learning environment is trivial, both quantitatively (a greater ease of information transmission is always beneficial) and qualitatively (generating rich adoption patterns like S-shaped curves requires additional forces such as specific forms of consumer heterogeneity).²

In our model (Section 2), an innovation of fixed, but uncertain quality (better or worse than the status quo) is introduced to a large population of forward-looking consumers. In the baseline setting, consumers are (ex ante) identical, sharing the same prior about the quality of the innovation, the same discount rate, and the same tastes for good and bad quality. At each instant in continuous time, consumers receive stochastic opportunities to adopt the innovation. A consumer who receives an opportunity must choose whether to irreversibly adopt the innovation or to delay his decision until the next opportunity. In equilibrium, consumers optimally trade off the opportunity cost of delays against the benefit to learning more about the quality of the innovation.

Learning is summarized by a public signal process, representing news that is obtained endogenously—based on the experiences of previous adopters—and possibly also from exogenous sources, such as professional critics or government watchdog agencies. To study the importance of quantitative and qualitative features of the news environment, we build on the exponential-bandit framework widely used in the literature on strategic experimentation (see Section 1.1): Individual adopters’ experiences generate public signals at a fixed Poisson rate that we use to quantify the potential for social learning. Qualitatively, there is a natural distinction between *bad news* markets, where signal arrivals (*breakdowns*) indicate bad quality and the absence of signals makes consumers more optimistic about the innovation; and *good news* markets, where signals (*breakthroughs*) suggest good quality and the absence of signals makes consumers more pessimistic. As we discuss (see Section 2.2), this distinction may reflect whether bad or good quality innovations are more likely to generate “newsworthy” (e.g., extreme) payoff realizations, as well as limitations or reporting practices of the available social learning systems.

Section 3 analyzes and contrasts equilibrium adoption dynamics in bad and good news markets. As in many applications of Poisson learning, we focus on the stark but tractable case of *perfect* bad (respectively, good) news, where a single signal arrival conclusively indicates bad (respectively, good) quality. Thus, incentives are non-trivial only in the absence of signals. We first derive a useful structural result, whereby equilibrium incentives over time satisfy a quasi-single crossing property (Lemma 1): Absent signals, there can be at most one transition from strict preference for adoption to strict preference for waiting, or vice

²See the discussion in Section 1.1.

versa, with a possible period of indifference in between. Building on this, Theorems 1 and 2 establish equilibrium existence and uniqueness under bad and good news.³ Equilibrium adoption dynamics admit simple closed-form descriptions that are Markovian in current beliefs and in the mass of consumers who have not yet adopted.

Under bad news, the unique equilibrium is characterized by two times $0 \leq t_1^* \leq t_2^*$: Until t_1^* , no adoption takes place and consumers acquire information only from exogenous sources; from t_2^* on, all consumers adopt immediately when given a chance (absent breakdowns). If $t_1^* < t_2^*$, then throughout (t_1^*, t_2^*) there is *partial adoption*: Only some consumers adopt when given a chance, with others free-riding on the information generated by the adopters, where the flow of adopters on (t_1^*, t_2^*) ensures indifference between adopting and delaying throughout this interval. Ceteris paribus, a period of partial adoption arises in economies with a large enough potential for social learning and with sufficiently patient and not too optimistic consumers, while otherwise there is no partial adoption. By contrast, under good news, the equilibrium is *all-or-nothing*, featuring immediate adoption up to some time t^* and no adoption from t^* on (absent breakthroughs). Thus, regardless of the potential for social learning, consumers' discount rate or prior beliefs, there is never any partial adoption.

This highlights a new distinction (absent in existing strategic experimentation models) between the way in which bad and good news learning affects consumers' incentives. During a period where a consumer is prepared to adopt the innovation, he is willing to delay his decision only if he expects to acquire decision-relevant information in the meantime: Since originally he is prepared to adopt, such information must make him strictly prefer *not* to adopt. Under bad news, breakdowns have this effect, as they reveal the innovation to be bad. By contrast, under good news, breakthroughs conclusively reveal the innovation to be good and hence cannot be decision-relevant to a consumer who is already willing to adopt.

Turning to implications of the equilibrium analysis, Section 4.1 considers adoption curves, which plot the percentage of adopters in the population against time. Under bad news, the adoption curve exhibits an “S-shaped” growth pattern whenever $t_1^* < t_2^*$: Absent breakdowns, adoption levels increase convexly throughout the partial adoption region (t_1^*, t_2^*) , while growth is concave from time t_2^* on (reflecting the gradual depletion of the population during the immediate adoption phase). Convex growth throughout (t_1^*, t_2^*) is tied to consumer indifference during this region: Absent breakdowns, consumers grow increasingly optimistic, and their opportunity cost to delaying goes up. To maintain indifference, this is offset by an increase in the flow of new adopters, which raises the odds that waiting will produce information that allows consumers to avoid a bad innovation. By contrast, the all-or-nothing structure of the good news equilibrium implies that adoption proceeds in (possibly multiple) “concave

³Uniqueness is in terms of aggregate adoption behavior.

bursts;” concave curves also arise in bad news markets with very optimistic and impatient consumers or with little potential for social learning (in which case $t_1^* = t_2^*$). S-curves are the leading empirically documented adoption pattern, while concave curves are also commonly observed.⁴ As we discuss in Section 1.1, our model complements existing explanations by highlighting a purely informational channel that might contribute to these patterns.

Finally, Section 4.2 considers increases in the potential for social learning. Proposition 1 establishes the possibility of a saturation effect: If learning is via bad news and the equilibrium features partial adoption, then such increases are (ex ante) welfare-neutral, because they are fully balanced out by an expansion of the period (t_1^*, t_2^*) of informational free-riding. As a result, greater opportunities for social learning also strictly slow down the adoption of good products and do not translate into uniformly faster learning about the quality of the innovation. In Section 4.3, we further build on this finding to show that, when consumers are heterogeneous, increased opportunities for social learning not only need not improve welfare, but indeed can be strictly Pareto-harmful. By contrast, in environments where equilibrium is all-or-nothing, increasing the potential for social learning is (essentially) always strictly beneficial and speeds up learning at all times.

1.1 Related Literature

This paper links the large and interdisciplinary literature on innovation adoption, which seeks to explain why the diffusion of innovations is typically a drawn-out process and why S-shaped (and to a lesser extent concave) adoption patterns are prevalent, with the theoretical literature on strategic experimentation, on which we build to study the informational externalities that arise under forward-looking social learning.

Our contribution to the former literature is twofold. First, we identify a novel, purely informational channel for the aforementioned regularities: Under forward-looking social learning, S-shaped growth can arise due to some consumers delaying adoption in the hope of learning whether previous adopters suffered negative experiences, while concave adoption might arise when consumers are very optimistic, impatient, or when public signals are expected to reveal positive information about the innovation. Existing models instead rely on a combination of the following ingredients (for detailed surveys, see [Geroski, 2000](#); [Baptista, 1999](#)): (i) heterogeneity of potential adopters, where the distribution of characteristics is imposed exogenously to yield the desired adoption pattern—as in “probit” models (e.g., [David,](#)

⁴ Indeed, typical marketing textbooks devote a chapter to these two patterns; see [Hoyer et al. \(2012\)](#), Ch. 15, p. 425ff. and [Keillor \(2007\)](#) p. 46–61. The former type of curve is sometimes referred to as “logistic” and the latter as “exponential” or “fast-break”. In economics, S-curves are studied by [Griliches \(1957\)](#), [Mansfield \(1961, 1968\)](#), [Gort and Klepper \(1982\)](#), among many others; for (essentially) concave curves see some of the “group A” innovations in [Davies \(1979\)](#).

1969; Davies, 1979; Karshenas and Stoneman, 1993), or existing learning-based models with myopic consumers (Young, 2009) or exogenous signals (Jensen, 1982);⁵ (ii) non-informational “spillover” effects which, independently of the quality of the innovation, increase current adoption as a function of past adoption—e.g., through a process of contagion as in “epidemic” models (e.g., Mansfield, 1961; Bass, 1969; Mahajan and Peterson, 1985) or due to pure payoff externalities resulting from learning-by-doing (Jovanovic and Lach, 1989) or network effects (Farrell and Saloner, 1985, 1986); (iii) supply-side factors such as pricing (e.g., Bergemann and Välimäki, 1997; Cabral, 2012). To highlight the explanatory power of informational incentives alone, our model abstracts away from (i)–(iii), though in many settings these factors are, of course, likely to be at play as well. Second, however, Sections 4.2–4.3 investigate the effect of increased opportunities for social learning and obtain welfare implications and testable predictions that are outside the scope of existing models.

Our model builds on the framework of strategic experimentation with exponential/Poisson bandits, originating with Keller et al. (2005) and Keller and Rady (2010, 2015) (for a survey, see Hörner and Skrzypacz, 2016).⁶ We introduce two main departures: First, we assume that adoption (i.e., exit to the risky arm) is irreversible rather than allowing for continuous back-and-forth switching; this is natural for innovations such as medical procedures or movies, for which “consumption” is typically a one-time event, or for technologies with large switching costs. Second, we assume a continuum of agents, who each have a negligible influence on public information. These departures entail a new qualitative difference between bad and good news learning—the presence vs. absence of partial adoption regions—that has observable implications for adoption curves. In contrast, in the above strategic experimentation models, the unique symmetric Markov equilibrium features a region of partial experimentation/mixing under both bad and good news. Another implication of these departures is that, unlike strategic experimentation, our setting does not feature an “encouragement effect,” where agents have an incentive to increase current experimentation to drive up beliefs and induce more future experimentation by others. This yields new comparative statics that isolate the impact of informational free-riding: For example, while in the bad news environment of Keller and Rady (2015), an increase in the number of players or signal informativeness makes players more willing to experiment at pessimistic beliefs, the saturation effect in Proposition 1 relies on the opposite prediction.

Many recent papers study good or bad news Poisson learning in other economic settings

⁵Under myopic or exogenous learning, individual consumers follow a cutoff rule, adopting the innovation iff beliefs are above a certain threshold. Thus, regions of convex adoption growth cannot arise with identical consumers, and instead require specific forms of heterogeneity in prior beliefs or preferences.

⁶Keller et al. (2005) and Keller and Rady (2010) study learning via conclusive and inconclusive good news, while Keller and Rady (2015) consider (in)conclusive bad news.

(e.g., [Strulovici, 2010](#); [Bonatti and Hörner, 2011, 2017](#); [Halac et al., 2016, 2017](#); [Guo, 2016](#); [Cripps and Thomas, 2019](#); [Thomas, 2019a,b](#)). Like ours, most of these papers focus on conclusive signals for tractability. [Board and Meyer-ter Vehn \(2013\)](#), [Halac and Kremer \(2020\)](#), and [Khromenkova \(2015\)](#) contrast the implications of bad vs. good news learning for reputation-building, career concerns, and collective decision-making. [Wolitzky \(2018\)](#) considers steady-state adoption levels of an innovation when consumers learn from a random sample of others’ outcomes and compares efficiency for cost-saving vs. outcome-improving innovations. [Che and Hörner \(2018\)](#) take a mechanism design approach to incentivizing social learning about a new product. [Laiho and Salmi \(2018\)](#) build on our model by incorporating monopoly pricing and consumer heterogeneity.

Finally, informational externalities and strategic delay incentives are also studied by the observational learning literature with endogenous timing; see, e.g., [Chamley and Gale \(1994\)](#) and, more closely related, [Murto and Välimäki \(2011\)](#), where players privately obtain Poisson signals about the quality of a risky project at a fixed exogenous rate until they choose to irreversibly exit to a safe outside option.⁷ The key difference is that in this literature players hold private information about a payoff-relevant state and make inferences by observing others’ actions, whereas all news in our model is public and derived from previous adopters’ experiences. Information aggregates in random bursts in these models rather than smoothly as in our setting, and the aforementioned papers do not derive adoption curves or study the way in which they are shaped by qualitative and quantitative features of the learning environment.

2 Model

2.1 The Game

Time $t \in [0, +\infty)$ is continuous. At time $t = 0$, an innovation of unknown quality $\theta \in \{G = 1, B = -1\}$ and of unlimited supply is released to a continuum population of potential consumers of mass $\bar{N}_0 \in \mathbb{R}_{>0}$. Consumers are ex ante identical: They have a common prior $p_0 \in (0, 1)$ that $\theta = G$; they are forward-looking with common discount rate $r > 0$; and they have the same actions and payoffs, as specified below.

At each time t , consumers receive stochastic opportunities to adopt the innovation. Adoption opportunities are generated independently across consumers and across histories accord-

⁷In earlier papers on observational learning, the timing of players’ moves is exogenous (e.g., [Banerjee, 1992](#); [Bikhchandani et al., 1992](#); [Smith and Sørensen, 2000](#)). [Board and Meyer-ter Vehn \(2020\)](#) study innovation adoption on networks with exogenous timing and observational learning from others’ adoption decisions.

ing to a Poisson process with exogenous arrival rate $\rho > 0$.⁸ Upon an adoption opportunity, a consumer must choose whether to adopt the innovation ($a_t = 1$) or to wait ($a_t = 0$). If a consumer adopts, he receives an expected lump sum payoff of $\mathbb{E}_t[\theta]$, conditioned on information available up to time t , and drops out of the game. If the consumer chooses to wait or does not receive an adoption opportunity at t , he receives a flow payoff of 0 until his next adoption opportunity, where he faces the same decision again.

2.2 Learning

Over time, consumers observe public signals that convey information about the quality of the innovation. To highlight the importance of qualitative and quantitative features of the informational environment, we employ a variation of the Poisson learning models widely used in the literature on strategic experimentation. Let N_t denote the flow of consumers newly adopting the innovation at time t , which we define more precisely in Section 2.3. Then, conditional on quality θ , public signals arrive according to an inhomogeneous Poisson process with arrival rate $(\varepsilon_\theta + \lambda_\theta N_t)dt$, where $\lambda_\theta > 0$ and $\varepsilon_\theta \geq 0$ are exogenous parameters that depend on the quality θ of the innovation.

The signal process summarizes news events that are generated from two sources: First, the *social learning* term λN_t represents news generated endogenously, based on the experiences of other consumers: It captures the idea of a flow N_t of new adopters each generating signals at rate λdt .⁹ Thus, the greater the flow of consumers adopting the innovation at t , the more likely it is for a signal to arrive at t , and hence the absence of a signal at t is more informative the larger N_t . Second, we also allow for (but do not require) signals to arrive at a fixed *exogenous* rate εdt , representing information generated independently of consumers' behavior, for example, by professional critics or government watchdog agencies.

As in many applications of Poisson learning, we focus for tractability on *perfect* news processes, where a single signal provides *conclusive* evidence of the quality of the innovation. Qualitatively, there is then a natural distinction between two types of news environments: Learning is via **perfect bad news** (for short, **bad news**) if $\varepsilon_G = \lambda_G = 0$ and $\varepsilon_B = \varepsilon \geq 0$, $\lambda_B = \lambda > 0$; that is, the arrival of a signal (called a *breakdown*) is conclusive evidence that the innovation is bad. Learning is via **perfect good news** (for short, **good news**) if $\varepsilon_B = \lambda_B = 0$ and $\varepsilon_G = \varepsilon \geq 0$, $\lambda_G = \lambda > 0$; that is, a signal arrival (called a *breakthrough*) is conclusive evidence that the innovation is good. The nature of the news environment may

⁸Stochasticity of adoption opportunities might capture cognitive and time constraints that prevent consumers from considering whether or not to adopt the innovation at every instant in continuous time. We further discuss stochastic adoption opportunities in Section 5.

⁹We obtain qualitatively similar results when the social learning component at time t is taken to depend on the stock, $\int_0^t N_s ds$, rather than the flow of adopters at t . See the discussion in Section 5.

be influenced by limitations or usage practices of the available social learning systems (for example, if a review platform allows users to “Like” a product, but has no “Dislike” button); or by whether a bad or good quality innovation is more likely to generate newsworthy (e.g., extreme) payoff realizations.¹⁰

Quantitatively, we use $\Lambda_0 := \lambda \bar{N}_0$ as a simple measure of the **potential for social learning** in the economy, summarizing both the likelihood λ with which individual adopters’ experiences find their way into the public domain and the size \bar{N}_0 of the population which can contribute to and access the common pool of information.

Evolution of Beliefs: Under bad news learning, consumers’ posterior on $\theta = G$ permanently jumps to 0 at the first breakdown, while under good news, consumers’ posterior on $\theta = G$ permanently jumps to 1 at the first breakthrough. Let p_t denote consumers’ *no-news posterior*, i.e., the belief at t that $\theta = G$ conditional on no signals having arrived on $[0, t)$. Given a flow of adopters $N_{s \geq 0}$, Bayesian updating implies that

$$p_t = \frac{p_0 e^{-\int_0^t (\varepsilon_G + \lambda_G N_s) ds}}{p_0 e^{-\int_0^t (\varepsilon_G + \lambda_G N_s) ds} + (1 - p_0) e^{-\int_0^t (\varepsilon_B + \lambda_B N_s) ds}}. \quad (1)$$

In particular, if N_τ is continuous in an open interval $(s, s + \nu)$ for $\nu > 0$, then p_τ for $\tau \in (s, s + \nu)$ evolves according to the ODE

$$\dot{p}_\tau = ((\varepsilon_B + \lambda_B N_\tau) - (\varepsilon_G + \lambda_G N_\tau)) p_\tau (1 - p_\tau).$$

Note that the no-news posterior is continuous. Moreover, it is increasing under bad news and decreasing under good news.

2.3 Equilibrium

Since our interest is in the aggregate adoption dynamics of the population, we take as the primitive of our equilibrium concept the aggregate flow $N_{t \geq 0}$ of consumers newly adopting the innovation over time and do not explicitly model individual consumers’ behavior. Given our focus on perfect news processes, incentives are non-trivial only in the absence of signals:

¹⁰Slightly more formally, suppose payoffs of the good vs. bad quality innovation are distributed (i.i.d across consumers) according to F_G and F_B , where $\int_{-\infty}^{\infty} \xi dF_G(\xi) = 1$ and $\int_{-\infty}^{\infty} \xi dF_B(\xi) = -1$. Suppose payoff realizations ξ are newsworthy iff either $\xi \leq \underline{\xi}$ or $\xi \geq \bar{\xi}$, for some “extreme” low and high payoffs $\underline{\xi} < \bar{\xi}$, and that newsworthy payoffs generate public signals at some rate. Then bad news learning assumes $F_B(\underline{\xi}) > 0 = F_G(\underline{\xi})$ and $F_B(\bar{\xi}) = F_G(\bar{\xi}) = 1$, i.e., bad innovations sometimes generate extreme low realizations, but neither good nor bad innovations generate extreme high payoffs. Analogously, good news learning assumes $F_B(\underline{\xi}) = F_G(\underline{\xi}) = 0$ and $F_B(\bar{\xi}) = 1 > F_G(\bar{\xi})$.

¹¹Definition 1 imposes measurability on N , so this expression is well-defined.

Under bad news, no consumers adopt after a breakdown, while under good news all remaining consumers adopt at their first opportunity following a breakthrough. Thus, we henceforth let N_t denote the flow of new adopters at t conditional on no signals up to time t and define equilibrium in terms of this quantity. Capturing that aggregate adoption is predictable with respect to the news process of the economy, we require N_t to be a *deterministic* function of time. We consider all such functions that are feasible in the following sense:

Definition 1. A feasible flow of adopters is a right-continuous function $N: [0, +\infty) \rightarrow \mathbb{R}$ such that $N_t := N(t) \in [0, \rho \bar{N}_t]$ for all $t \in [0, +\infty)$, where $\bar{N}_t := \bar{N}_0 - \int_0^t N_s ds$.

Here \bar{N}_t denotes the mass of consumers remaining in the game at time t . We require that $N_t \leq \rho \bar{N}_t$ so that N_t is consistent with the remaining \bar{N}_t consumers independently receiving adoption opportunities at Poisson rate ρ . Any feasible adoption flow $N_{t \geq 0}$ defines an associated no-news posterior p_t^N as given by (1).

In equilibrium, we require that at each time t , N_t is consistent with optimal behavior by the remaining \bar{N}_t forward-looking consumers: A consumer who receives an adoption opportunity at t optimally trades off his expected payoff to adopting against his value to waiting, given that he assigns probability p_t^N to $\theta = G$ and expects the population's adoption behavior to evolve according to the process $N_{s \geq 0}$. For this we first define the value to waiting.

Let Σ_t denote the set of all right-continuous functions $\sigma: [t, +\infty) \rightarrow \{0, 1\}$. Under the Poisson process generating adoption opportunities, any $\sigma \in \Sigma_t$ defines a random time τ^σ at which, absent signals, the consumer will adopt the innovation and drop out of the game.¹² Let $W_t^N(\sigma)$ denote the expected payoff to waiting at t and following σ in the future, given the aggregate adoption flow $N_{s \geq 0}$. Specifically, under bad news learning, σ prescribes adoption at the random time τ^σ if and only if there have been no breakdowns prior to τ^σ , yielding

$$W_t^N(\sigma) := \mathbb{E} \left[e^{-r(\tau^\sigma - t)} \left(p_t^N - (1 - p_t^N) e^{-\int_t^{\tau^\sigma} (\varepsilon + \lambda N_s) ds} \right) \right],$$

where the expectation is with respect to the Poisson process generating adoption opportunities. Under good news learning, following σ means that at any adoption opportunity prior to τ^σ , adoption occurs only if there has been a breakthrough, and at τ^σ adoption occurs whether or not there has been a breakthrough. For any time $s \geq t$, denote by τ_s the random

¹²Formally, let $(X_s)_{s \geq t}$ denote the stochastic process representing the number of arrivals generated on $[t, s]$ by a Poisson process with arrival rate ρ , and let $(X_{s-})_{s > t}$ denote the number of arrivals on $[t, s)$. Then $\tau^\sigma := \inf\{s \geq t : \sigma_s \times (X_s - X_{s-}) > 0\}$, with the usual convention that $\inf \emptyset := +\infty$. It is well-known that the hitting time of a right-continuous process of an open set is an optional time. Therefore, the expectations in the definition of the value to waiting are well-defined.

time at which the first adoption opportunity after s arrives. Then $W_t^N(\sigma)$ is given by

$$\mathbb{E} \left[\left(p_t e^{-\int_t^{\tau^\sigma} (\varepsilon + \lambda N_s) ds} + (1 - p_t) \right) e^{-r(\tau^\sigma - t)} (2p_{\tau^\sigma} - 1) + p_t \int_t^{\tau^\sigma} (\varepsilon + \lambda N_s) e^{-\int_t^s (\varepsilon + \lambda N_k) dk} e^{-r(\tau_s - t)} ds \right],$$

where the expectation is again with respect to the Poisson process generating adoption opportunities.

The value to waiting at t is the payoff to waiting and behaving optimally in the future:

Definition 2. The value to waiting given a feasible adoption flow $N_{t \geq 0}$ is the function $W^N : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined by $W_t^N := \sup_{\sigma \in \Sigma_t} W_t^N(\sigma)$ for all t .

We are now ready to formally define equilibrium:

Definition 3. An equilibrium is a feasible adoption flow $N_{t \geq 0}$ such that

- (i). $W_t^N \geq 2p_t^N - 1$ for all t such that $N_t < \rho \bar{N}_t$; and
- (ii). $W_t^N \leq 2p_t^N - 1$ for all t such that $0 < N_t$.

Condition (i) says that if some consumers who receive an adoption opportunity at t decide not to adopt, then the value to waiting, W_t^N , must weakly exceed the expected payoff to immediate adoption, $2p_t^N - 1$. Similarly, condition (ii) requires that if some consumers adopt at time t , then the value to waiting must be weakly less than the payoff to immediate adoption. Thus, at all times, N_t is consistent with consumers optimally trading off the expected payoff to immediate adoption against the value to waiting.¹³

3 Equilibrium Analysis

3.1 Quasi-Single Crossing Property for Equilibrium Incentives

We now proceed to equilibrium analysis. As a preliminary step, we establish a useful structural property of equilibrium incentives under both bad and good news. Suppose that $N_{t \geq 0}$ is an *arbitrary* feasible flow of adopters, with associated no-news posterior $p_{t \geq 0}^N$ and value to waiting $W_{t \geq 0}^N$. In general, the dynamics of the trade-off between immediate adoption at

¹³Definition 3 is essentially Nash equilibrium, i.e., does not impose subgame perfection. The motivation is that in a continuum population any individual consumer's behavior has a negligible impact on aggregate adoption levels, so that any off-path history in which the flow of adopters differs from the equilibrium flow is more than a unilateral deviation from the equilibrium path. Thus, off-path histories do not affect individual consumers' incentives on the equilibrium path and are unimportant for equilibrium analysis.

time t (yielding expected payoff $2p_t^N - 1$) and delaying and behaving optimally in the future (yielding expected payoff W_t^N) can be quite difficult to characterize, with $(2p_t^N - 1) - W_t^N$ changing sign many times. However, when $N_{t \geq 0}$ is an *equilibrium* flow, then for any t ,

$$\begin{aligned} 2p_t^N - 1 < W_t^N &\implies N_t = 0; \text{ and} \\ 2p_t^N - 1 > W_t^N &\implies N_t = \rho \bar{N}_t; \end{aligned}$$

and this imposes considerable discipline on the dynamics of the trade-off. Indeed, the following result shows that $2p_t^N - 1$ and W_t^N satisfy a quasi-single crossing property: There can be at most one transition from strict preference for adoption to strict preference for waiting, or vice versa, with a possible period of indifference in between.

Lemma 1. *Suppose that learning is either via bad news ($\lambda_B > 0 = \lambda_G$) or via good news ($\lambda_G > 0 = \lambda_B$). Let $N_{t \geq 0}$ be an equilibrium, with corresponding no-news posteriors $p_{t \geq 0}^N$ and value to waiting $W_{t \geq 0}^N$. Then $W_{t \geq 0}^N$ and $2p_{t \geq 0}^N - 1$ satisfy quasi-single crossing, in the following sense:*

- If $(\lambda_B - \lambda_G)(W_t^N - (2p_t^N - 1)) < 0$, then $(\lambda_B - \lambda_G)(W_\tau^N - (2p_\tau^N - 1)) < 0$ for all $\tau > t$.
- If $(\lambda_B - \lambda_G)(W_t^N - (2p_t^N - 1)) \leq 0$, then $(\lambda_B - \lambda_G)(W_\tau^N - (2p_\tau^N - 1)) \leq 0$ for all $\tau > t$.

The proof is in Appendix A.1. We briefly illustrate the intuition for the first bullet point when learning is via bad news. Suppose that immediate adoption is strictly better than waiting today, and hence also in the near future provided there are no breakdowns.¹⁴ Then, in equilibrium, in the near future all consumers adopt upon their first opportunity, so the no-news posterior strictly increases while the number of remaining consumers strictly decreases. Because information is generated endogenously, this means that the flow of information must decrease over time. As a result, immediate adoption becomes even more attractive relative to waiting, and hence remains strictly preferable at all times in the future.

Given Lemma 1, we associate two cutoff times $0 \leq t_1^* \leq t_2^* \leq +\infty$ with any equilibrium $N_{t \geq 0}$.¹⁵ If learning is via bad news, set

$$t_1^* := \inf\{t \geq 0 : N_t > 0\} \text{ and } t_2^* := \sup\{t \geq 0 : N_t < \rho \bar{N}_t\}. \quad (2)$$

If learning is via good news, set

$$t_1^* := \inf\{t \geq 0 : N_t < \rho \bar{N}_t\} \text{ and } t_2^* := \sup\{t \geq 0 : N_t > 0\}. \quad (3)$$

¹⁴This follows from the continuity of the equilibrium value to waiting; see Lemma A.1.

¹⁵With the convention that $\inf \emptyset = +\infty$ and $\sup \emptyset = 0$.

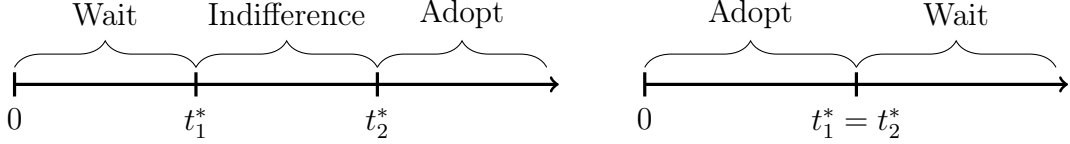


Figure 1: Equilibrium structure under bad news (left) vs. good news (right).

Thus, if $N_{t \geq 0}$ is a bad news equilibrium it features no adoption ($N_t = 0$) for all $t < t_1^*$ and immediate adoption ($N_t = \rho \bar{N}_t$) for all $t > t_2^*$ absent breakdowns; under good news, $N_{t \geq 0}$ features immediate adoption prior to t_1^* and no adoption after t_2^* absent breakthroughs. Lemma 1 implies that at all times $t \in (t_1^*, t_2^*)$, consumers are indifferent ($2p_t^N - 1 = W_t^N$) between adopting and delaying under both bad and good news.¹⁶

In Sections 3.2 and 3.3, we will build on this observation to establish equilibrium existence and uniqueness under both bad and good news. As illustrated in Figure 1, we will see that the good news equilibrium satisfies $t_1^* = t_2^* = t^*$; thus, adoption behavior is **all-or-nothing**, with all consumers adopting upon their first opportunity up to time t^* and adoption ceasing from then on absent breakthroughs. By contrast, the bad news equilibrium features a non-empty region (t_1^*, t_2^*) for suitable parameters. Maintaining indifference throughout (t_1^*, t_2^*) requires a form of informational free-riding, which we term **partial adoption**, whereby only *some* consumers adopt when given the chance. We will see that partial adoption has important implications for the shape of the adoption curve and for the impact of increased opportunities for social learning on welfare, learning, and adoption dynamics.

3.2 Bad News Equilibrium

First, consider learning via bad news. Building on the preceding structural result, the following theorem establishes the existence and uniqueness of equilibrium. The equilibrium flow of adopters N_t is Markovian in the associated no-news posterior p_t and the time- t potential for social learning $\Lambda_t := \lambda \bar{N}_t$:

Theorem 1 (Bad News Equilibrium). *Fix $r, \rho, \lambda, \bar{N}_0 > 0$, $\varepsilon \geq 0$, and $p_0 \in (0, 1)$. There exists a unique equilibrium N_t . Moreover, N_t is Markovian in (p_t, Λ_t) for all t : There exists*

¹⁶Consider the good news case (bad news is analogous). Fix $t \in (t_1^*, t_2^*)$. By the definition of t_1^* and t_2^* , there exist $k \in (t_1^*, t)$ and $l \in (t, t_2^*)$ such that $N_k < \rho \bar{N}_k$ and $N_l > 0$. Since N is an equilibrium, this implies $2p_k - 1 \leq W_k$ and $2p_l - 1 \geq W_l$, whence Lemma 1 yields $2p_t - 1 = W_t$.

a non-decreasing function $\Lambda^* : [0, 1] \rightarrow \mathbb{R} \cup \{\infty\}$ and some $p^* \in [\frac{1}{2}, 1)$ such that

$$N_t = \begin{cases} 0 & \text{if } p_t < p^* \text{ and } \Lambda_t > \Lambda^*(p_t) \\ \frac{r(2p_t-1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda} \in (0, \rho\bar{N}_t) & \text{if } p_t \geq p^* \text{ and } \Lambda_t > \Lambda^*(p_t) \\ \rho\bar{N}_t & \text{if } \Lambda_t \leq \Lambda^*(p_t). \end{cases} \quad (4)$$

As we saw in the previous section, for any equilibrium, (2) yields cutoff times t_1^* and t_2^* such that no consumers adopt at times $t < t_1^*$, all consumers adopt upon an opportunity at times $t > t_2^*$, and consumers are indifferent between immediate adoption and waiting at all intermediate times $t \in (t_1^*, t_2^*)$. Theorem 1 implies that t_1^* and t_2^* are uniquely pinned down by the parameters. Moreover, whenever $t_1^* < t_2^*$ (as we will see is the case for suitable parameter values), the flow of adopters throughout (t_1^*, t_2^*) is also uniquely pinned down. The proof of Theorem 1 is in Appendix A.2. Below we sketch the basic argument:

Partial adoption during (t_1^*, t_2^*) . Lemma A.6 shows that the flow of adopters at all times $t \in (t_1^*, t_2^*)$ must satisfy $N_t = \frac{r(2p_t-1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda} \in (0, \rho\bar{N}_t)$. Thus, adoption throughout (t_1^*, t_2^*) is *partial*, with only some consumers adopting when given a chance and others free-riding on the information generated by the adopters. Heuristically, maintaining consumer indifference requires that the cost and benefit of delaying be equal:

$$\underbrace{(\varepsilon + \lambda N_t)(1 - p_t)dt}_{\text{Probability of breakdown}} \underbrace{(0 - (-1))}_{\text{Benefit: Avoid Bad Product}} = \underbrace{(1 - (\varepsilon + \lambda N_t)(1 - p_t)dt)}_{\text{Probability of no breakdown}} \underbrace{(2p_{t+dt} - 1)r dt}_{\text{Cost: Discounting}}. \quad (5)$$

Delaying one's decision by an instant is beneficial if a breakdown occurs at that instant, allowing a consumer to permanently avoid the bad product. The gain in this case is $(0 - (-1)) = 1$, and this possibility arises with an instantaneous probability of $(\varepsilon + \lambda N_t)(1 - p_t)dt$. On the other hand, if no breakdown occurs, which happens with instantaneous probability $1 - (\varepsilon + \lambda N_t)(1 - p_t)dt$, then consumers incur an opportunity cost of $(2p_{t+dt} - 1)r dt$, reflecting the time cost of delayed adoption. Ignoring terms of order dt^2 and rearranging yields $N_t = \frac{r(2p_t-1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda}$.¹⁷

Determining the cutoff times. Next, Lemma A.7 provides an alternative description of the cutoff times t_1^* and t_2^* in terms of the evolution of the no-news posterior p_t and the potential for social learning Λ_t . To see the idea, let \underline{p} denote the lowest posterior at which a

¹⁷ More precisely, ignoring terms of order dt^2 , the right hand side of (5) is given by $(1 - (\varepsilon + \lambda N_t)(1 - p_t)dt)(2p_t + \dot{p}_t dt - 1)r dt = r(2p_t - 1)dt$. Rearranging yields the desired expression. Note that ρ does not enter into this expression. This is because in the indifference region consumers obtain the same continuation payoff regardless of whether or not they receive an adoption opportunity in the interval $(t, t + dt)$.

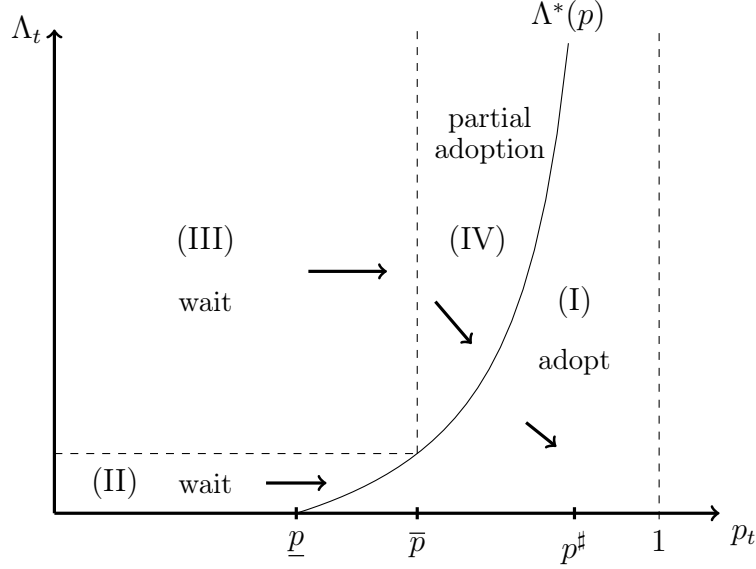


Figure 2: Equilibrium dynamics of (p_t, Λ_t) when $\varepsilon < \rho$.

consumer to whom adoption opportunities arrive at rate ρ is willing to adopt immediately if all information in the future arrives exclusively through the *exogenous* news source (i.e., at rate ε). In the appendix, we show that $\underline{p} = \frac{(\varepsilon+r)(r+\rho)}{2(\varepsilon+r)(r+\rho)-\varepsilon\rho}$. Then

$$\bar{p} := \lim_{\rho \rightarrow \infty} \underline{p} = \frac{\varepsilon + r}{\varepsilon + 2r} \quad \text{and} \quad p^\sharp := \lim_{\varepsilon \rightarrow \infty} \underline{p} = \frac{\rho + r}{\rho + 2r}$$

denote, respectively, the values of \underline{p} in the limit as adoption opportunities arrive continuously or as exogenous information fully resolves all uncertainty by the next instant. Note that $\underline{p} < \bar{p}, p^\sharp$. Moreover, $p^\sharp > \bar{p}$ if and only if $\varepsilon < \rho$ (see Figure 2).

For any belief $p \in (\underline{p}, p^\sharp)$, we show that there is a unique mass $\bar{N}^*(p) \in \mathbb{R}_+$ of remaining consumers with the following property: If *all* these remaining consumers adopt at their first future opportunity, then a consumer with current posterior p is indifferent between immediate adoption and adoption at his next opportunity (absent breakdowns). Based on this, define the function $\Lambda^* : [0, 1] \rightarrow \mathbb{R}_+ \cup \{+\infty\}$, depicted in Figure 2: For $p \in (\underline{p}, p^\sharp)$, set $\Lambda^*(p) = \lambda \bar{N}^*(p)$; for $p \leq \underline{p}$, set $\Lambda^*(p) = 0$; for $p \geq p^\sharp$, set $\Lambda^*(p) = +\infty$. Then, letting $p^* := \min\{\bar{p}, p^\sharp\}$, Lemma A.7 shows that the cutoff times satisfy $t_2^* = \inf\{t \geq 0 : \Lambda_t < \Lambda^*(p_t)\}$ and $t_1^* = \min\{t_2^*, \sup\{t \geq 0 : p_t < p^*\}\}$.

Equilibrium dynamics given initial parameters. The previous two steps imply that any equilibrium must take the Markovian form in (4). It remains to show how (4) uniquely pins down the evolution of N_t as a function of the initial parameters, and to verify that $N_{t \geq 0}$ does indeed constitute an equilibrium (in particular, is feasible). We briefly sketch the

former argument, relegating the latter to the appendix. We impose two conditions that rule out uninteresting values of ε , ρ and p_0 , under which equilibrium adoption is either identically zero or there is never a partial adoption region (i.e., $t_1^* = t_2^*$ regardless of other parameters):¹⁸

Condition 1. *Either $\varepsilon > 0$ or $p_0 \in (\frac{1}{2}, 1)$.*

Condition 2. *The rate at which exogenous information arrives is lower than the rate at which consumers obtain adoption opportunities: $\varepsilon < \rho$.*

Figure 2 illustrates equilibrium dynamics under these conditions. The values of (p_t, Λ_t) in region (I) correspond to the third line of (4), i.e., to immediate adoption. Regions (II) and (III) correspond to the first line of (4); that is, no adoption takes place in these regions. Finally, region (IV) corresponds to partial adoption as given by the second line of (4).

If (p_0, Λ_0) starts off in region (I) or (II), then equilibrium features no partial adoption. In the former case, we have $t_1^* = t_2^* = 0$, so that all consumers adopt upon their first opportunity. In the latter case, initially all consumers delay and the no-news posterior drifts upward according to the law of motion $\dot{p}_t = p_t(1 - p_t)\varepsilon$, while Λ_t remains unchanged at Λ_0 . This yields a unique time $t_1^* = t_2^* > 0$ at which (p_t, Λ_t) hits the boundary between regions (II) and (I); subsequently, all consumers adopt immediately upon an opportunity.

By contrast, if (p_0, Λ_0) is in region (III) or (IV), then $t_1^* < t_2^*$. In the former case, initial dynamics are the same as in region (II). But now this yields a unique time $t_1^* > 0$ at which (p_t, Λ_t) hits the boundary separating regions (III) and (IV), which implies that $t_1^* < t_2^*$. Throughout $[t_1^*, t_2^*)$, the evolution of (p_t, Λ_t) is pinned down by the second line of (4), and t_2^* is uniquely given by the first time at which (p_t, Λ_t) enters region (I).¹⁹ Analogously, if (p_0, Λ_0) is in region (IV), then $t_1^* = 0$ and $t_2^* > t_1^*$ is the first time at which (p_t, Λ_t) , evolving according to the second line of (4), enters region (I).

Conditions for partial adoption. As seen above, whether or not the equilibrium features a period of partial adoption depends on the fundamentals. More specifically, Figure 2 shows that provided consumers are not too optimistic or impatient (i.e., $p_0 < p^\sharp$),²⁰ then $t_1^* < t_2^*$ holds whenever the potential for social learning Λ_0 is sufficiently large:

¹⁸If Condition 1 is violated, then (4) implies that $N_t = 0$ for all t . If Condition 2 is violated, then $p^* := \min\{\bar{p}, p^\sharp\} = p^\sharp$. Thus, Lemma A.7 implies that $N_t = 0$ as long as $\Lambda_t > \Lambda^*(p_t)$ and $N_t = \rho\bar{N}_t$ as soon as $\Lambda_t \leq \Lambda^*(p_t)$.

¹⁹Specifically, combining the second line of (4) with (1) yields the ODE $\dot{p}_t = rp_t(2p_t - 1)$. Given the initial value $p_{t_1^*} = \bar{p}$, this uniquely pins down p_t via $p_t = \frac{p_{t_1^*}}{2p_{t_1^*} - e^{r(t-t_1^*)}(2p_{t_1^*} - 1)}$. Plugging this back into $N_t = \frac{r(2p_t - 1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda}$ pins down $\Lambda_t = \lambda\bar{N}_t$. Note that since $p_{t_1^*} > \frac{1}{2}$, p_t is strictly increasing and reaches p^\sharp in finite time. Thus, $t_2^* = \inf\{t : \Lambda_t < \Lambda^*(p_t)\} < +\infty$.

²⁰Note that p^\sharp is decreasing in the discount rate r .

Lemma 2. Fix ρ , ε and p_0 satisfying Conditions 1 and 2, and $r > 0$. Assume $p_0 < p^\sharp$. Then there exists $\bar{\Lambda}_0 > 0$ such that $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$ if and only if $\Lambda_0 > \bar{\Lambda}_0$.²¹

On the other hand, if as in existing learning-based models of innovation adoption, learning is *purely exogenous* ($\lambda = 0$ and $\varepsilon > 0$) or consumers are *myopic* (“ $r = +\infty$ ”), then there is never any partial adoption, regardless of other parameters. In the former case, $0 = \Lambda_t < \Lambda^*(p)$ for all $p > \underline{p}$, so by Theorem 1 no consumers adopt until the no-news posterior hits \underline{p} , and from then on all consumers adopt immediately when given a chance. The latter case corresponds to $\underline{p} = \bar{p} = \frac{1}{2}$ and $\Lambda^*(p) = +\infty$ for all $p > \frac{1}{2}$, so $t_1^* = t_2^* = \inf\{t : p_t > \frac{1}{2}\}$. Thus, the possibility of partial adoption in equilibrium hinges crucially both on consumers being forward-looking and on there being opportunities for social learning.

3.3 Good News Equilibrium

Next, consider learning via good news. As under bad news, there is a unique equilibrium $N_{t \geq 0}$, and N_t is Markovian in the state variables (p_t, Λ_t) . Notably, however, regardless of the potential for social learning in the economy, the equilibrium is *all-or-nothing*: There is a cutoff belief p^* above which *all* consumers adopt if given an opportunity and below which *no* consumers adopt.

Theorem 2 (Good News Equilibrium). Fix $r, \rho, \lambda, \bar{N}_0 > 0, \varepsilon \geq 0$, and $p_0 \in (0, 1)$. There exists a unique equilibrium N_t . Moreover, N_t is Markovian in (p_t, Λ_t) (or equivalently (p_t, \bar{N}_t)) for all t and satisfies:

$$N_t = \begin{cases} \rho \bar{N}_t & \text{if } p_t > p^* \\ 0 & \text{if } p_t \leq p^*, \end{cases} \quad (6)$$

where

$$p^* = \frac{(\varepsilon + r)(r + \rho)}{2(\varepsilon + r)(r + \rho) - \varepsilon \rho}.$$

We prove Theorem 2 in Appendix A.3. We again invoke the quasi-single crossing property in Lemma 1. As we saw in Section 3.1, this implies that in any equilibrium, there are times $0 \leq t_1^* \leq t_2^* \leq +\infty$ such that absent breakthroughs, $N_t = \rho \bar{N}_t$ if $t < t_1^*$, $N_t = 0$ if $t > t_2^*$, and throughout (t_1^*, t_2^*) consumers are indifferent between adopting immediately and waiting.

The key observation (Lemma A.9) is that we must in fact have $t_1^* = t_2^* =: t^*$. To see the intuition, suppose $t_1^* < t_2^*$. Then consumers would be indifferent between adopting and delaying at each time $t \in (t_1^*, t_2^*)$. Moreover, there is $t \in (t_1^*, t_2^*)$ and $\Delta \in (0, t_2^* - t)$ such that

²¹By the Markovian description in (4), Λ_0 is a sufficient statistic for equilibrium; i.e., holding other fundamentals fixed, Λ_0 pins down the equilibrium adoption flow, beliefs, and cutoff times $t_1^*(\Lambda_0)$ and $t_2^*(\Lambda_0)$.

$N_\tau > 0$ throughout $[t, t + \Delta]$.²² As with bad news learning, we can compare a consumer's payoff to adopting at t with the payoff to delaying his decision by an instant:

$$r(2p_t - 1)dt + p_t(\lambda N_t + \varepsilon)dt \left(1 - \frac{\rho}{r + \rho}\right).$$

The first term represents the gain to immediate adoption if no breakthrough occurs between t and $t+dt$, which happens with instantaneous probability $(1-p_t(\lambda N_t+\varepsilon)dt)$. Just as under bad news, the gain to adopting immediately in this case is $r(2p_{t+dt} - 1)dt$, representing time discounting at rate r and the fact that at $t+dt$ the consumer remains indifferent between adopting and delaying. Ignoring terms of order dt^2 yields $r(2p_t - 1)dt$. The second term represents the gain to immediate adoption if there is a breakthrough between t and $t + dt$, which happens with instantaneous probability $p_t(\lambda N_t + \varepsilon)dt > 0$. Now the situation is very different from the bad news setting: A breakthrough conclusively signals *good* quality, so a consumer who delays his decision by an instant will adopt immediately at his next opportunity. This results in a discounted payoff of $\frac{\rho}{r+\rho}$, reflecting the stochasticity of adoption opportunities. In contrast, by adopting at t , the consumer receives a payoff of $1 > \frac{\rho}{r+\rho}$ immediately. Thus, regardless of whether or not there is a breakthrough between t and $t+dt$, there is a strictly positive gain to adopting immediately at t , contradicting indifference at t .

The above argument illustrates an important difference between bad and good news learning that is new relative to existing strategic experimentation models (see the discussion in Section 1.1). In order to maintain a period of indifference between immediate adoption and waiting, it must be possible to acquire *decision-relevant information* by waiting an instant: Consumers who are prepared to adopt at t will be willing to delay their decision by an instant only if there is a possibility that at the next instant they will no longer be willing to adopt. In the bad news setting, this is indeed possible, because a breakdown might occur. In contrast, if learning is via good news, this cannot happen: A breakthrough between t and $t + dt$ reveals the innovation to be good, so consumers strictly prefer to adopt from $t + dt$ on; and if there is no breakthrough, then consumers remain indifferent at $t + dt$, so in either case the information obtained is not decision-relevant. Of course, breakthroughs do convey decision-relevant information at beliefs where consumers strictly prefer to *delay*. But in the interior of a region of indifference, such beliefs cannot be reached instantaneously.

Given that $t_1^* = t_2^* = t^*$, Theorem 2 follows from the observation that $p_t \leq p^*$ if and only if $t \geq t^*$ (Lemma A.10). It is worth noting that if $\varepsilon = 0$, then $p^* = \frac{1}{2}$, so regardless of the discount rate r , consumers behave *entirely myopically*. If $\varepsilon > 0$, then consumers' forward-

²²By definition of t_2^* , there exists $t \in (t_1^*, t_2^*)$ such that $N_t > 0$. By right-continuity of N , we must then have $N_\tau > 0$ for all $\tau > t$ sufficiently close.

looking nature is reflected by the fact that the cutoff posterior p^* below which consumers are unwilling to adopt is $\frac{(\varepsilon+r)(r+\rho)}{2(\varepsilon+r)(r+\rho)-\varepsilon\rho} > \frac{1}{2}$. In both cases, p^* does not depend on λ or \bar{N}_0 : Social learning only affects the *time* t^* at which adoption ceases conditional on no breakthroughs.

4 Implications

We now study implications of the preceding equilibrium analysis for observed adoption patterns, as well as for the effect of increased opportunities for social learning on welfare, learning, and adoption dynamics.

4.1 Adoption Curves: S-Shaped vs. Concave

Consider the *adoption curve* of the innovation, which plots the share of adopters in the population against time. Conditional on no news up to time t , this is given by

$$A_t := \int_0^t N_s / \bar{N}_0 ds.$$

Theorems 1 and 2 translate into the following predictions for the shape of the adoption curve:

Corollary 1. *Bad News:* *In the unique equilibrium of Theorem 1, A_t has the following shape: For $0 \leq t < t_1^*$, $A_t = 0$. For $t_1^* \leq t < t_2^*$, A_t is strictly increasing and convex in t . For $t \geq t_2^*$, A_t is strictly increasing and concave in t . If the first breakdown occurs at time t , then adoption comes to a standstill from then on.*

Good News: *In the unique equilibrium of Theorem 2, $A_t = 1 - e^{-\rho t}$ is strictly increasing and concave for all $t < t^*$. If there is a breakthrough prior to t^* , then $A_t = 1 - e^{-\rho t}$ for all t . If the first breakthrough occurs at $s > t^*$, then adoption comes to a temporary standstill between t^* and s , and for all $t \geq s$, A_t is strictly increasing and concave and given by $1 - e^{-\rho(t^*+t-s)}$.*

Thus, in bad news markets, the adoption curve exhibits an “S-shaped” (i.e., convex-concave) growth pattern whenever $t_1^* < t_2^*$, where the region of convex growth coincides with the partial adoption region (t_1^*, t_2^*) . By contrast, in good news markets, adoption proceeds in (possibly multiple) “concave bursts.” Concave adoption curves also arise in bad news markets with very optimistic and impatient consumers or with little potential for social learning (in which case $t_1^* = t_2^*$ by Lemma 2). Figures 3 and 4 illustrate the differing adoption patterns.

The fact that the convex growth period of A_t under bad news coincides with the partial adoption region (t_1^*, t_2^*) is tied to consumer indifference in this region: Conditional on no breakdowns during this period, consumers grow increasingly optimistic about the quality

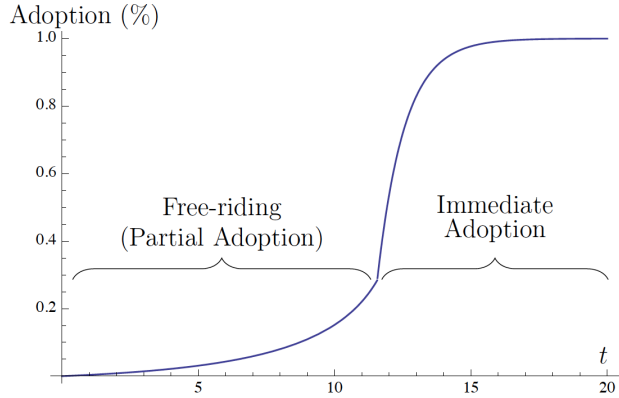


Figure 3: S-shaped adoption curve under bad news conditional on no breakdowns ($t_1^* = 0$).

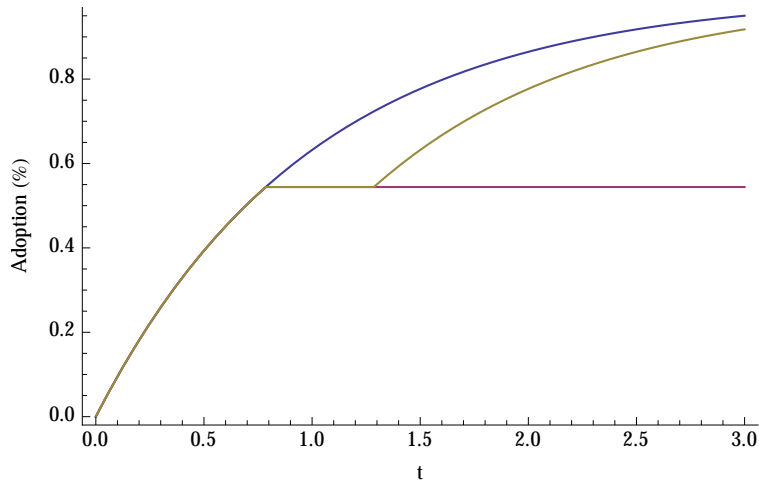


Figure 4: Concave adoption curves under good news: blue = breakthrough before t^* ; yellow = breakthrough after t^* ; pink = bad quality. (Parameters: $\varepsilon = 1/2$, $r = 1$, $\rho = 1$, $\lambda = 0.5$, $p_0 = 0.7$.)

of the innovation, which increases their opportunity cost of delaying adoption. To maintain indifference, the benefit to delaying adoption must then also increase over time: This is achieved by increasing the arrival rate of future breakdowns, which improves the odds that waiting will allow consumers to avoid the bad product. But since the arrival rate of information is increasing in the flow N_t of new adopters, this means that N_t must be strictly increasing throughout (t_1^*, t_2^*) . Since N_t represents the rate of change of A_t , this is equivalent to A_t being convex.²³ By contrast, the concave growth regions under both bad and good news simply reflect the gradual depletion of the population when all consumers adopt immediately upon an opportunity.

²³This argument for convex growth does not rely on the linearity of λN_t ; it remains valid as long as the rate at which the bad product generates breakdowns at t is increasing in N_t .

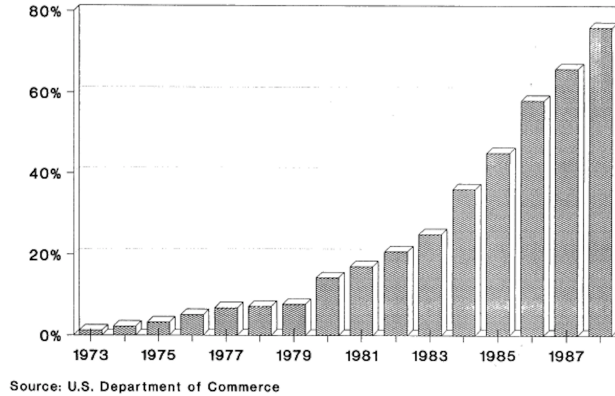


Figure 5: Initial convex growth in US microwave adoption levels through the late 1980s (Guenther et al., 1991). Later growth slowed to reach ownership levels of around 97% in 2011 (Williams, 2014).

As discussed in the Introduction, S-shaped adoption patterns are widely documented for many different innovations. Our model complements existing explanations (see Section 1.1) by identifying a purely informational channel for this regularity: If there is a high enough chance that previous adopters’ experiences might reveal negative information about the innovation, then as long as consumers are forward-looking, S-shaped adoption can arise due to some consumers strategically delaying adoption. This channel may be especially natural for innovations whose introduction was accompanied by substantial safety concerns: A classic example are microwave ovens (Figure 5), whose introduction in the late 1960s faced widespread concerns about possible “radiation leaks” and whose initial adoption levels remained low despite the fact that the entry of Japanese firms into the US market led to substantial price decreases in the early 1970s.²⁴ Safety concerns may also be salient for new medical procedures, for which S-shaped adoption patterns are again commonly documented.²⁵

Though less prevalent than S-shaped curves, concave adoption is another leading pattern documented in the marketing literature (e.g., Keillor, 2007, pp. 51–61), with leisure-enhancing innovations such as movies, books, or beauty and fitness products as examples. While our model abstracts away from many important product-specific details, Corollary 1 suggests some factors that could contribute to concave adoption patterns: In particular, high levels of consumer impatience or optimism, or if social learning about a given product (for

²⁴See Wiersema and Buzzell (1979) for a detailed discussion.

²⁵For example, consider bariatric surgery, a collection of surgical weight loss procedures introduced in the 1990s, for which health advice websites still feature warnings regarding possible serious complications. Consistent with S-shaped adoption, the annual number of procedures worldwide (i.e., the number of *new* adoptions) increased from 40,000 in 1998 to 146,301 in 2003 and to 344,221 in 2008, and then plateaued at 340,768 in 2011 (Buchwald and Oien, 2009, 2013).

example, an essentially side-effect free herbal remedy) is primarily about “whether it actually works” (i.e., good news events or their absence).

4.2 The Effect of Increased Opportunities for Social Learning

Next, we consider an increase in the potential for social learning $\Lambda_0 := \lambda \bar{N}_0$, capturing either a greater ease of information transmission (e.g., due to the introduction of new social networking platforms) or a larger community of consumers. We ask how this affects welfare, learning, and adoption dynamics. Again, informational free-riding in the form of partial adoption has important implications. Indeed, under bad news learning, an economy’s ability to harness its potential for social learning is subject to a *saturation effect*: Whenever the equilibrium features partial adoption, then further increases in the potential for social learning are welfare-neutral, cause learning to slow down over certain periods, and decrease the adoption of (both good and bad) innovations at all times.

Formally, we fix all other parameters and study the effect of increasing Λ_0 on ex-ante equilibrium welfare $W_0(\Lambda_0)$; no-news posteriors $p_t^{\Lambda_0}$; and expected adoption levels $A_t(\Lambda_0, G)$ and $A_t(\Lambda_0, B)$ conditional on good and bad quality, respectively. We assume that the original potential for social learning Λ_0 is such that there is partial adoption, i.e. $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$; under the conditions in Lemma 2, this is the case whenever Λ_0 is large enough.

Proposition 1. *Consider learning via bad news. Fix r, ρ, ε , and p_0 . If Λ_0 is such that $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$, then an increase in the potential for social learning to $\hat{\Lambda}_0 > \Lambda_0$ has the following effect:*

(i). **Welfare Neutrality:** $W_0(\hat{\Lambda}_0) = W_0(\Lambda_0)$.

(ii). **Non-Monotonicity of Learning:** *There exists $\bar{t} > t_2^*(\Lambda_0)$ such that*

$$\begin{cases} p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0} & \text{if } t \leq t_2^*(\Lambda_0) & \text{(learning is **equally fast** under } \Lambda_0 \text{ and } \hat{\Lambda}_0) \\ p_t^{\Lambda_0} > p_t^{\hat{\Lambda}_0} & \text{if } t_2^*(\Lambda_0) < t < \bar{t} & \text{(learning is **slower** under } \hat{\Lambda}_0) \\ p_t^{\Lambda_0} < p_t^{\hat{\Lambda}_0} & \text{if } t > \bar{t} & \text{(learning is **faster** under } \hat{\Lambda}_0). \end{cases}$$

(iii). **Slowdown in Adoption:** *For all t and $\theta = B, G$, we have $A_t(\Lambda_0, \theta) \geq A_t(\hat{\Lambda}_0, \theta)$, with strict inequality for all $t > t_1^*(\Lambda_0)$.*

The proof of Proposition 1 is in Online Appendix B.3. The idea behind welfare neutrality is as follows. Since the equilibrium features partial adoption at Λ_0 , the same is true when the potential for social learning increases to $\hat{\Lambda}_0$. Moreover, both the time t_1^* at which adoption

begins and the posterior $p_{t_1^*}$ at t_1^* are the same under Λ_0 and $\hat{\Lambda}_0$.²⁶ But then, since consumers strictly prefer to delay at all $t < t_1^*$ and are indifferent between delaying and adopting at t_1^* , ex-ante welfare under both Λ_0 and $\hat{\Lambda}_0$ corresponds to the expected payoff to waiting until t_1^* and adopting at t_1^* absent breakdowns. Thus, $W_0(\hat{\Lambda}_0) = W_0(\Lambda_0)$.

This welfare neutrality result contrasts with the cooperative benchmark where consumers coordinate on socially optimal adoption levels: In the latter case, increased opportunities for social learning are strictly beneficial and for any $p_0 > \frac{1}{2}$ the first-best (complete information) payoff of $\frac{\rho}{r+\rho}p_0$ can be approximated in the limit as $\Lambda_0 \rightarrow \infty$.²⁷ The result also contrasts with myopic social learning or forward-looking exogenous learning, where welfare necessarily increases in response to more informative signals (even if consumers are heterogeneous).²⁸

Points (ii) and (iii) further illuminate the forces behind welfare neutrality. By (ii), an increase in Λ_0 affects learning dynamics in a non-monotonic manner. Thus, the impact on a consumer's expected payoff varies with the time t at which he obtains his first adoption opportunity: If $t \leq t_2^*(\Lambda_0)$, his expected payoff is the same under Λ_0 and $\hat{\Lambda}_0$. If $t \in (t_2^*(\Lambda_0), \bar{t})$, he is strictly worse off under $\hat{\Lambda}_0$, because in case the innovation is bad he is less likely to have found out by then than under Λ_0 . Finally, if $t > \bar{t}$, he is strictly better off under $\hat{\Lambda}_0$. Depending on $\hat{\Lambda}_0$, \bar{t} adjusts endogenously to balance out the benefits, which arrive at times after \bar{t} , with the costs incurred at times $(t_2^*(\Lambda_0), \bar{t})$.

Similarly, by (iii), an increase in Λ_0 strictly decreases the adoption $A_t(\Lambda_0, G)$ of good products (which is harmful), but also decreases the adoption $A_t(\Lambda_0, B)$ of bad products (which is beneficial), and welfare-neutrality obtains because these forces balance out in equilibrium. Figure 6 illustrates that the strict slow-down in the adoption of good products is due to two effects: On the extensive margin, the increase in Λ_0 pushes out t_2^* , i.e., prolongs free-riding; on the intensive margin, the increase strictly drives down the growth rate of A_t at all $t < t_2^*(\Lambda_0)$.

Point (iii) yields new testable implications relative to existing models of innovation adoption, suggesting, for example, that the proportion of adopters of an innovation may grow more slowly in larger communities. This prediction is broadly consistent with [Bandiera and Rasul's \(2006\)](#) finding that we discussed in footnote 1: They document that farmers whose network includes many adopters may be less likely to adopt initially themselves; thus, in

²⁶Indeed, t_1^* is the first time at which the posterior exceeds the threshold $\bar{p} = \frac{\varepsilon+r}{\varepsilon+2r}$ and learning up to t_1^* is purely via the exogenous news source.

²⁷In an earlier version of this paper, the cooperative benchmark is shown to take an all-or-nothing form, with no adoption below a cutoff belief p^s and immediate adoption above p^s . Relative to this, equilibrium displays two inefficiencies: First, because $p^s < \underline{p}$, adoption generally begins too late. Second, when $t_1^* < t_2^*$, then once consumers begin to adopt, initial adoption is too low. See [Frick and Ishii \(2014\)](#), Section 5.3.

²⁸To define ex-ante welfare with myopic consumers, assume that consumers' payoffs are discounted at some arbitrary rate $r > 0$, but consumers behave myopically.

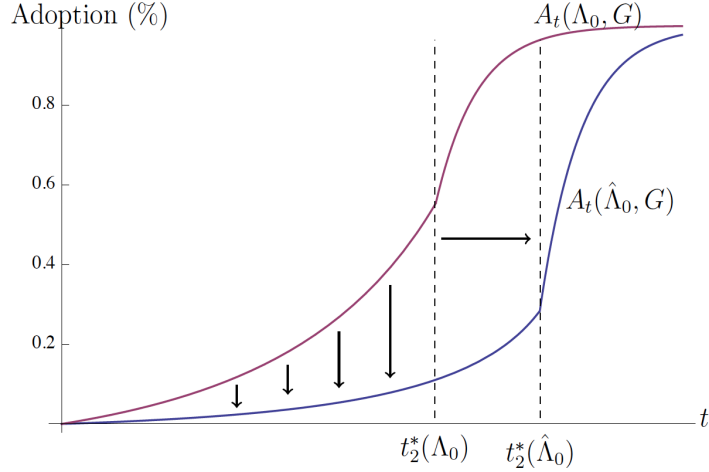


Figure 6: The effect of increased opportunities for social learning on the adoption of a good product under bad news ($\hat{\Lambda}_0 > \Lambda_0$).

equilibrium, larger networks of farmers should feature lower percentages of adoption.

Finally, the logic behind the saturation effect relies crucially on partial adoption/informational free-riding. Correspondingly, as we show in Online Appendix B.5, there is no saturation effect under good news: Since equilibrium adoption is all-or-nothing in this case, increasing the potential for social learning speeds up learning at all times, which strictly improves welfare (provided $\varepsilon > 0$). Likewise, under bad news learning, if initial opportunities for social learning are so low that the equilibrium does not feature partial adoption, then increasing Λ_0 is strictly beneficial; see Online Appendix B.4.

4.3 More Social Learning Can Hurt: An Example

Assuming homogeneous consumers, we saw in Proposition 1 that, beyond a certain level, further increasing the potential for social learning is welfare-neutral in bad news environments. More strongly, we now show that when consumers are *heterogeneous*, increased opportunities for social learning can bring about *Pareto-decreases* in ex-ante welfare.

To illustrate this possibility, we consider bad news learning and introduce some heterogeneity in consumers' patience levels. Suppose the population consists of two types of consumers: There is a mass \bar{N}_0^p of *patient* types with discount rate $r_p > 0$ and a mass \bar{N}_0^i of *impatient* types with discount rate $r_i > r_p$. Assume for simplicity that there is no exogenous news ($\varepsilon = 0$); similar arguments apply when $\varepsilon > 0$.

First, consider the economy consisting only of mass \bar{N}_0^p of patient types (and no impatient types). Then, provided $p_0 \in (\frac{1}{2}, \frac{\rho+r_p}{\rho+2r_p})$, Lemma 2 implies that the equilibrium features partial

adoption whenever the social learning parameter λ exceeds a certain level $\bar{\lambda}^p$. Thus, for any $\lambda > \bar{\lambda}^p$, the patient types are initially indifferent between adopting and waiting, and hence their ex-ante payoff is $2p_0 - 1$.

Now, fix any $\hat{\lambda} > \lambda > \bar{\lambda}^p$, and consider the economy consisting of *both* types of consumers. The following result shows that as long as the mass of impatient types is small, the patient types' ex-ante payoffs continue to be $2p_0 - 1$ under both λ and $\hat{\lambda}$. However, the impatient types' payoffs are *strictly* lower under $\hat{\lambda}$ than λ . Thus, the ex-ante payoff profile at $\hat{\lambda}$ is Pareto-dominated by the payoff profile at λ , despite the fact that $\hat{\lambda} > \lambda$ entails greater opportunities for social learning:

Proposition 2. *Suppose learning is via bad news and $\varepsilon = 0$. Fix $r_i > r_p > 0$, and $p_0 \in (\frac{1}{2}, \frac{\rho+r_p}{\rho+2r_p})$. Consider $\bar{N}_0^p > 0$ and $\hat{\lambda} > \lambda > \bar{\lambda}^p$. There exists $\eta > 0$ such that whenever the mass of impatient types \bar{N}_0^i is at most η , then the patient types' ex-ante payoffs satisfy $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$, while the impatient types' ex-ante payoffs satisfy $W_0^i(\hat{\lambda}) < W_0^i(\lambda)$.*

The proof is in Online Appendix B.6. The basic idea is as follows. First, consider the equilibrium adoption flows under λ and $\hat{\lambda}$ in the game consisting solely of mass \bar{N}_0^p of patient types. What payoffs would a hypothetical impatient type r_i (which does not exist in this game) obtain if he behaved optimally when faced with these adoption flows? Since the patient types are initially indifferent between adopting or delaying in both equilibria, a monotonicity argument shows that the impatient type's optimal strategy in both cases is to adopt upon his first opportunity. Hence, the ex-ante payoff of the hypothetical impatient type under the social learning rate $\gamma \in \{\lambda, \hat{\lambda}\}$ satisfies:

$$W_0^i(\gamma) = \int_0^{\infty} \rho e^{-(r_i+\rho)\tau} \frac{p_0}{p_\tau^\gamma} (2p_\tau^\gamma - 1) d\tau.$$

Given this, we invoke the non-monotonicity result for learning in Proposition 1 to show that $W_0^i(\hat{\lambda}) < W_0^i(\lambda)$. Specifically, by Proposition 1, there exists some time $\bar{t} > t^* := t_2^*(\lambda)$ such that learning under $\hat{\lambda}$ and λ is equally fast up to time t^* , is strictly slower under $\hat{\lambda}$ between t^* and \bar{t} , and is faster under $\hat{\lambda}$ from time \bar{t} on. For the patient types, the cost of the deceleration in learning at times $t \in (t^*, \bar{t})$ and the benefit of the acceleration in learning at times $t > \bar{t}$ exactly balance out, as $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$. But this implies that the hypothetical impatient type must be strictly hurt by these changes: Intuitively, relative to a patient type, he weights the early costs more heavily than the later benefits.

To complete the proof, we show that as long as the mass of impatient types $\bar{N}_0^i > 0$ is sufficiently small, we still have $W_0^i(\hat{\lambda}) < W_0^i(\lambda)$ and $W_0^p(\hat{\lambda}) = W_0^p(\lambda)$. The first inequality

follows from a continuity argument. The second equality reflects the fact that when \bar{N}_0^i is small, patient types continue to partially adopt initially in both equilibria.

An assumption underlying this argument is that adoption opportunities are stochastic and limited. Given that ρ is finite, impatient types may not receive any adoption opportunities for a long time. This is the source of the above welfare loss, because if an impatient type obtains his first adoption opportunity (e.g., arrives to the market) between t^* and \bar{t} , then at that point, he has less information about the innovation under $\hat{\lambda}$ than λ . If consumers were able to adopt freely at any time, then impatient types would incur no loss, as all of them would adopt at time 0 in both the λ and $\hat{\lambda}$ -equilibrium. Thus, Proposition 2 illustrates an interesting interaction between heterogeneity and delays due to limited adoption opportunities.

5 Concluding Remarks

This paper develops a model of innovation adoption when consumers are forward-looking and learning is social. Our analysis isolates the effect of purely informational incentives on aggregate adoption dynamics, learning, and welfare. We highlight how qualitative and quantitative features of the learning environment shape these incentives, most importantly, by determining whether or not there is informational free-riding in the form of partial adoption. The presence or absence of partial adoption has observable implications, suggesting a novel channel for two widespread adoption patterns—S-shaped and concave curves. Moreover, partial adoption has important welfare implications, entailing that increased opportunities for social learning need not benefit consumers and can be strictly harmful.

We conclude by briefly commenting on some modifications and extensions of our model:

Learning from the stock of adopters. In our model, the social learning component of the signal arrival rate at time t , λN_t , depends only on the flow of new adopters, N_t . This effectively assumes that each adopter can generate a signal only once, namely at the time of adoption. This seems an appropriate approximation for innovations such as medical procedures, where “adoption” is a one-time event and the rate at which adopters generate news about quality depreciates rapidly from the time of adoption. In contrast, for some durable goods, it might be more natural to allow adopters to generate signals repeatedly over time, by letting signals at t arrive at rate λS_t , where $S_t := \int_0^t N_s ds$ represents the *stock* of adopters. This would produce similar results. Specifically, similar arguments can be used to establish the existence and uniqueness of equilibrium under both bad and good news. The good news equilibrium is again all-or-nothing while, for appropriate parameters, the bad news equilibrium again features a partial adoption region with behavior pinned down by the indifference condition $S_t = \frac{r(2p_t-1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda}$. Finally, the partial adoption region again

exhibits convex growth in adoption levels.²⁹

Stochastic adoption opportunities. In our model, adoption opportunities arrive stochastically to each consumer at a Poisson rate $\rho \in \mathbb{R}_+$. Section 2.1 motivated this assumption as capturing frictions in the form of cognitive or time constraints that prevent consumers from contemplating the decision whether or not to adopt at every point in continuous time; the first adoption opportunity can also be interpreted as stochastic arrival to the market. Of course, our analysis remains valid for arbitrarily large values of ρ , representing these frictions becoming vanishingly small. To see how equilibrium is affected as ρ grows large, assume for simplicity that $\varepsilon = 0$ and $p_0 \in (1/2, 1)$. Then under good news, Theorem 2 continues to imply that equilibrium is all-or-nothing with cutoff posterior $p^* = 1/2$, but as ρ becomes arbitrarily large, the time t^* it takes to reach p^* becomes arbitrarily short, effectively approximating an initial mass point of adoption.³⁰ Under bad news, for any $p > 1/2$, $\Lambda^*(p) \rightarrow 0$ as $\rho \rightarrow \infty$. Hence, by Theorem 1, there is an initial partial adoption region with flow of adopters given by $N_t = \frac{r(2p_t-1)}{\lambda(1-p_t)}$, and as $\rho \rightarrow \infty$, this partial adoption region grows longer and longer, but its duration t_2^* is bounded above.

More general signal processes. Finally, as in many applications of Poisson learning, we have restricted attention to conclusive bad vs. good news signals. This made the equilibrium analysis quite tractable, yielding closed-form expressions for all key quantities and allowing us to compute several comparative statics. While a careful investigation of more general signal processes is beyond the scope of the paper, the analysis readily extends to hybrid environments with *two* types of conclusive Poisson signals—bad news *and* good news signals with respective arrival rates $\lambda_B N_t$ and $\lambda_G N_t$. In particular, when $\lambda_B > \lambda_G$, the equilibrium is analogous to Theorem 1. Some of our insights also extend beyond environments with conclusive signals: For example, it is worth noting that partial adoption relies crucially on the possibility of news events that trigger discrete downward jumps in beliefs (although such events need not *conclusively* signal bad quality). Without such events (e.g., when learning is based on inconclusive good news Poisson signals or Brownian motion), a similar logic as in Section 3.3 shows that there cannot be continuous regions of partial adoption, because

²⁹The idea behind convexity is as follows: As in Section 3.2, indifference requires that the benefit of avoiding a bad product when a breakdown occurs (i.e., $(1 - p_t)(\lambda S_t + \varepsilon)$) equal the opportunity cost of delaying adoption when no breakdown occurs (i.e., $r(2p_t - 1)$). Since consumers grow more optimistic absent breakdowns, this has two implications throughout the indifference region: (i) beliefs p_t increase convexly, because the growth rate of p_t equals the instantaneous probability of a breakdown ($\frac{\dot{p}_t}{p_t} = (1 - p_t)(\lambda S_t + \varepsilon)$), and the latter must increase over time to balance out the increasing opportunity cost of delay; (ii) the stock of adopters $S_t = S(p_t)$ increases convexly as a function of *beliefs*, to ensure that breakdowns arrive at a rate that counterbalances the convex growth (w.r.t. beliefs) of the ratio $\frac{r(2p_t-1)}{(1-p_t)}$ between the opportunity cost of delay and the probability of facing a bad product. Combining (i) and (ii), it follows that S_t , and hence adoption levels, must increase convexly as a function of time.

³⁰This holds provided Λ_0 is large enough that $\frac{p_0}{p_0 + e^{\Lambda_0}(1-p_0)} < \frac{1}{2}$, so that p^* is reached eventually.

a consumer who is willing to adopt cannot acquire decision-relevant information by delaying his decision by an instant.³¹

A Appendix: Main Proofs

This appendix presents the proofs of Lemma 1 and Theorems 1–2. All remaining proofs are in Online Appendix B.

A.1 Proof of Lemma 1 (Quasi-Single Crossing Property)

In this section, we prove Lemma 1. We will make use of the following five lemmas which are proved in Online Appendix B.1. For an equilibrium adoption flow $N_{t \geq 0}$, denote the associated value to waiting by $W_{t \geq 0}^N$ and the no-news posterior by $p_{t \geq 0}^N$.

Lemma A.1. *If $N_{t \geq 0}$ is an equilibrium, then W_t^N is continuous in t .*

Lemma A.2. *Suppose that $N_{t \geq 0}$ is an equilibrium and that $W_t^N < 2p_t^N - 1$ for some $t > 0$. Then there exists some $\nu > 0$ such that W_t^N is continuously differentiable in t on the interval $(t - \nu, t + \nu)$ and*

$$\begin{aligned} \dot{W}_t^N = & (r + \rho + (\varepsilon_G + \lambda_G \rho \bar{N}_t) p_t^N + (\varepsilon_B + \lambda_B \rho \bar{N}_t)(1 - p_t^N)) W_t^N \\ & - \rho(2p_t^N - 1) - p_t^N (\varepsilon_G + \lambda_G \rho \bar{N}_t) \frac{\rho}{\rho + r}. \end{aligned}$$

Lemma A.3. *Suppose that $N_{t \geq 0}$ is an equilibrium and that $W_t^N > 2p_t^N - 1$ for some $t > 0$. Then there exists some $\nu > 0$ such that W_t^N is continuously differentiable in t on the interval $(t - \nu, t + \nu)$ and*

$$\dot{W}_t^N = (r + p_t^N \varepsilon_G + (1 - p_t^N) \varepsilon_B) W_t^N - p_t^N \varepsilon_G \frac{\rho}{\rho + r}.$$

The final two lemmas focus on learning via bad news:

Lemma A.4. *Let $N_{t \geq 0}$ be an equilibrium under bad news. Suppose that $\varepsilon > 0$ or $p_0 > \frac{1}{2}$. Then $\lim_{t \rightarrow \infty} p_t^N = \mu(\varepsilon, \Lambda_0, p_0)$ and $\lim_{t \rightarrow \infty} W_t^N = \frac{\rho}{\rho + r} (2\mu(\varepsilon, \Lambda_0, p_0) - 1)$, where*

$$\mu(\varepsilon, \Lambda_0, p_0) := \begin{cases} 1 & \text{if } \varepsilon > 0, \\ \frac{p_0}{p_0 + (1 - p_0)e^{-\Lambda_0}} & \text{if } \varepsilon = 0. \end{cases}$$

³¹For this, we assume that there is no exogenous news. Details are available upon request.

Lemma A.5. *Suppose that learning is via bad news. Suppose that $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$. Then the unique equilibrium satisfies $N_t = 0$ for all t .*

Henceforth we drop the superscript N from W and p .

Proof of Lemma 1 under Good News:

Let $\varepsilon = \varepsilon_G \geq 0 = \varepsilon_B$ and $\lambda = \lambda_G > 0 = \lambda_B$.

Step 1: $W_t = 2p_t - 1 \implies W_\tau \geq 2p_\tau - 1$ for all $\tau \geq t$:

Suppose $W_t = 2p_t - 1$ at some time t and suppose for a contradiction that at some time $s' > t$, we have $W_{s'} < 2p_{s'} - 1$. Let

$$s^* = \sup\{s < s' : W_s = 2p_s - 1\}.$$

By continuity, $s^* < s'$, $W_{s^*} = 2p_{s^*} - 1$, and $W_s < 2p_s - 1$ for all $s \in (s^*, s')$. Then by Lemma A.2, the right hand derivative of $W_s - (2p_s - 1)$ at s^* exists and satisfies:

$$\lim_{s \downarrow s^*} \dot{W}_s - 2\dot{p}_s = r(2p_{s^*} - 1) + p_{s^*} (\varepsilon + \lambda \rho \bar{N}_{s^*}) \frac{r}{\rho + r} > 0.$$

This implies that for some $s \in (s^*, s')$ sufficiently close to s^* we have $W_s > 2p_s - 1$, which is a contradiction.

Step 2: $W_t > 2p_t - 1 \implies W_\tau > 2p_\tau - 1$ for all $\tau > t$:

Suppose by way of contradiction that there exists $s' > t$ such that $W_{s'} = 2p_{s'} - 1$. Let

$$s^* = \inf\{s > t : W_s = 2p_s - 1\}.$$

By continuity, $s^* > t$, $W_{s^*} = 2p_{s^*} - 1$, and $W_s > 2p_s - 1$ for all $s \in (t, s^*)$. Note that $p_{s^*} \geq \frac{1}{2}$, because W_{s^*} is bounded below by 0. Moreover, by Lemma A.3 the left-hand derivative of $W_s - (2p_s - 1)$ at s^* exists and is given by:

$$\lim_{s \uparrow s^*} \dot{W}_s - 2\dot{p}_s = r(2p_{s^*} - 1) + p_{s^*} \frac{r}{\rho + r} \varepsilon.$$

If $\varepsilon > 0$, this is strictly positive, implying that for some $s \in (t, s^*)$ sufficiently close to s^* , we have $W_s < 2p_s - 1$, which is a contradiction. If $\varepsilon = 0$, then for all $s \in (t, s^*)$, we have $p_{s^*} = p_s$ and $W_s = e^{-r(s^*-s)} W_{s^*} = e^{-r(s^*-s)} (2p_{s^*} - 1) \leq 2p_s - 1$. Thus, $W_s \leq 2p_s - 1$, again contradicting $W_s > 2p_s - 1$. ■

Proof of Lemma 1 under Bad News:

Let $\varepsilon = \varepsilon_B \geq 0 = \varepsilon_G$ and $\lambda = \lambda_B > 0 = \lambda_G$. If $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$, then by Lemma A.5 $N_t = 0$ for all t , so the proof of Lemma 1 is obvious. We now prove the theorem under the assumption that either $\varepsilon > 0$ or $p_0 > \frac{1}{2}$.

Step 1: $W_t = 2p_t - 1 \implies W_\tau \leq 2p_\tau - 1$ for all $\tau \geq t$:

Suppose that $W_t = 2p_t - 1$ and suppose for a contradiction that $W_{s'} > 2p_{s'} - 1$ for some $s' > t$. Let $\bar{s} := \inf\{s > s' : W_t \leq 2p_s - 1\} < \infty$, since by Lemma A.4 $\lim_{t \rightarrow \infty} 2p_t - 1 > \lim_{t \rightarrow \infty} W_t$. Let $\underline{s} := \sup\{s < s' : W_s \leq 2p_s - 1\}$. Then $\underline{s} < \bar{s}$, $W_{\underline{s}} = 2p_{\underline{s}} - 1$, $W_{\bar{s}} = 2p_{\bar{s}} - 1$, and $W_s > 2p_s - 1$ for all $s \in (\underline{s}, \bar{s})$. Lemma A.3 together with the fact that $N_s = 0$ for all $s \in (\underline{s}, \bar{s})$ implies the following two limits:

$$L_{\underline{s}} := \lim_{s \downarrow \underline{s}} \left(\dot{W}_s - \frac{d}{ds}(2p_s - 1) \right) = (r + (1 - p_{\underline{s}})\varepsilon)(2p_{\underline{s}} - 1) - 2p_{\underline{s}}(1 - p_{\underline{s}})\varepsilon$$

$$L_{\bar{s}} := \lim_{s \uparrow \bar{s}} \left(\dot{W}_s - \frac{d}{ds}(2p_s - 1) \right) = (r + (1 - p_{\bar{s}})\varepsilon)(2p_{\bar{s}} - 1) - 2p_{\bar{s}}(1 - p_{\bar{s}})\varepsilon.$$

Because $W_s > 2p_s - 1$ for all $s \in (\underline{s}, \bar{s})$, we need $L_{\underline{s}} \geq 0$ and $L_{\bar{s}} \leq 0$. Rearranging this implies:

$$r(2p_{\underline{s}} - 1) \geq (1 - p_{\underline{s}})\varepsilon$$

and

$$r(2p_{\bar{s}} - 1) \leq (1 - p_{\bar{s}})\varepsilon.$$

But if $\varepsilon > 0$, then $p_{\bar{s}} > p_{\underline{s}}$, so this is impossible. On the other hand, if $\varepsilon = 0$ and $p_0 > \frac{1}{2}$, then for all $s \in (\underline{s}, \bar{s})$, we have that $p_s = p_{\bar{s}} > \frac{1}{2}$ and $W_s = e^{-r(\bar{s}-s)}W_{\bar{s}}$. Since $W_{\bar{s}} = 2p_{\bar{s}} - 1$, this implies $W_s = e^{-r(\bar{s}-s)}(2p_{\bar{s}} - 1) < 2p_s - 1$, contradicting $W_s > 2p_s - 1$. This completes the proof of Step 1.

Step 2: $W_t < 2p_t - 1 \implies W_\tau < 2p_\tau - 1$ for all $\tau > t$:

Suppose that $W_t < 2p_t - 1$, let $\underline{s} := \inf\{s' > t : W_{s'} \geq 2p_{s'} - 1\}$, and suppose for a contradiction that $\underline{s} < \infty$. By continuity, $W_\tau < 2p_\tau - 1$ for all $\tau \in [t, \underline{s})$ and $W_{\underline{s}} = 2p_{\underline{s}} - 1$. Furthermore, by Lemma A.4, there exists some $\bar{s} \geq \underline{s}$ such that $2p_{\bar{s}} - 1 = W_{\bar{s}}$ and $2p_s - 1 > W_s$ for all $s > \bar{s}$. Lemma A.2 implies the following two limits:

$$H_{\underline{s}} := \lim_{s \uparrow \underline{s}} \left(\dot{W}_s - \frac{d}{ds}(2p_s - 1) \right) = r(2p_{\underline{s}} - 1) - (\varepsilon + \lambda\rho\bar{N}_{\underline{s}})(1 - p_{\underline{s}})$$

$$H_{\bar{s}} := \lim_{s \downarrow \bar{s}} \left(\dot{W}_s - \frac{d}{ds}(2p_s - 1) \right) = r(2p_{\bar{s}} - 1) - (\varepsilon + \lambda\rho\bar{N}_{\bar{s}})(1 - p_{\bar{s}}).$$

As usual, because $W_s < 2p_s - 1$ for all $s \in (t, \underline{s})$ and for all $s > \bar{s}$, we must have $H_{\underline{s}} \geq 0$ and $H_{\bar{s}} \leq 0$. But since $p_{\bar{s}} \geq p_{\underline{s}}$, this is only possible if $\underline{s} = \bar{s} =: s^*$ and $H_{s^*} = H_{\underline{s}} = H_{\bar{s}} = 0$.

Thus,

$$r(2p_{s^*} - 1) = (\varepsilon + \lambda\rho\bar{N}_{s^*})(1 - p_{s^*}).$$

Now consider any $s \in [t, s^*)$. Because $p_s \leq p_{s^*}$ and $\bar{N}_s \geq \bar{N}_{s^*}$, we must have

$$r(2p_s - 1) \leq (\varepsilon + \lambda\rho\bar{N}_s)(1 - p_s).$$

Combining this with the fact that $W_s < 2p_s - 1$ yields

$$rW_s < (\varepsilon + \lambda\rho\bar{N}_s)(1 - p_s) < (2p_s - W_s)(\varepsilon + \lambda\rho\bar{N}_s)(1 - p_s) + \rho(2p_s - 1 - W_s).$$

Rearranging we obtain:

$$0 < -rW_s + \rho(2p_s - 1 - W_s) + (2p_s - W_s)(\varepsilon + \lambda\rho\bar{N}_s)(1 - p_s).$$

By Lemma A.2, the right-hand side is precisely the derivative $\frac{d}{ds}(2p_s - 1) - \dot{W}_s$. But then for all $s \in [t, s^*)$, $2p_s - 1 > W_s$ and $2p_s - 1 - W_s$ is strictly increasing, contradicting continuity and the fact that $2p_{s^*} - 1 = W_{s^*}$. \blacksquare

A.2 Proof of Theorem 1 (Bad News Equilibrium)

In this section we prove Theorem 1. Recall the following beliefs defined in Section 3.2:

$$\begin{aligned} \underline{p} &:= \frac{(\varepsilon + r)(r + \rho)}{2(\varepsilon + r)(r + \rho) - \varepsilon\rho}, \\ \bar{p} &:= \frac{\varepsilon + r}{\varepsilon + 2r}, \\ p^\sharp &:= \frac{\rho + r}{\rho + 2r}. \end{aligned}$$

Let $p^* := \min\{\bar{p}, p^\sharp\}$. Define the function $G : [0, 1] \times \mathbb{R}_+ \rightarrow \mathbb{R}$ by

$$G(p, \Lambda) := \int_0^\infty \rho e^{-(r+\rho)\tau} \left(p - (1-p)e^{-(\varepsilon\tau + \Lambda(1-e^{-\rho\tau}))} \right) d\tau, \text{ for all } (p, \Lambda) \in [0, 1] \times \mathbb{R}_+.$$

We extend G to the domain $[0, 1] \times (\mathbb{R}_+ \cup \{+\infty\})$ by defining:

$$G(p, +\infty) := \frac{\rho}{\rho + r}p.$$

Finally, define the non-decreasing function $\Lambda^* : [0, 1] \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ by

$$\begin{cases} \Lambda^*(p) = 0 & \text{if } p \leq \underline{p}, \\ 2p - 1 = G(p, \Lambda^*(p)) & \text{if } p \in (\underline{p}, p^\sharp), \\ \Lambda^*(p) = +\infty & p \geq p^\sharp. \end{cases}$$

As discussed in the main text, for $p \in (\underline{p}, p^\sharp)$, $\bar{N}^*(p) := \frac{1}{\lambda}\Lambda^*(p)$ has the following property: If $\bar{N}^*(p)$ consumers remain and if *all* these remaining consumers adopt at their first future opportunity, then a consumer with current posterior p is indifferent between immediate adoption and adoption at his next opportunity (absent breakdowns).

The proof of Theorem 1 proceeds in three steps. Suppose that $N_{t \geq 0}$ is an equilibrium with associated cutoff times t_1^* and t_2^* as defined by (2). We first show in Lemma A.6 that if $t_1^* < t_2^*$, then at all $t \in (t_1^*, t_2^*)$, N_t is pinned down by a simple ODE. Second, Lemma A.7 provides a characterization of t_1^* and t_2^* in terms of the evolution of (p_t, Λ_t) . Given these two steps, it is easy to see that if an equilibrium exists, it is unique and must take the Markovian form in (4) of Theorem 1. Finally, to verify equilibrium existence, Lemma A.8 shows that the adoption flow implied by (4) is feasible.

A.2.1 Characterization of Adoption between t_1^* and t_2^*

Lemma A.6. *Suppose $N_{t \geq 0}$ is an equilibrium with associated no-news posterior $p_{t \geq 0}$ and cutoff times t_1^* and t_2^* as defined by Equation (2). Suppose that $t_1^* < t_2^*$. Then at all times $t \in (t_1^*, t_2^*)$,*

$$N_t = \frac{r(2p_t - 1)}{\lambda(1 - p_t)} - \frac{\varepsilon}{\lambda}.$$

Proof. By definition of t_1^* and t_2^* and Lemma 1, we have $2p_t - 1 = W_t^N$ at all $t \in (t_1^*, t_2^*)$. Because p_t is weakly increasing, this implies that p_t and W_t^N are differentiable at almost all $t \in (t_1^*, t_2^*)$ (with respect to Lebesgue measure).

Using again the fact that $2p_t - 1 = W_t^N$ at all $t \in (t_1^*, t_2^*)$ we obtain for all $t \in (t_1^*, t_2^*)$:

$$\begin{aligned} W_t^N &= e^{-r(t_2^* - t)} \left(p_t + (1 - p_t)e^{-\int_t^{t_2^*} (\varepsilon + \lambda N_s) ds} \right) (2p_{t_2^*} - 1) \\ &= e^{-r(t_2^* - t)} \left(p_t - (1 - p_t)e^{-\int_t^{t_2^*} (\varepsilon + \lambda N_s) ds} \right), \end{aligned} \tag{7}$$

where the second equality follows from Equation (1). Consider any $t \in (t_1^*, t_2^*)$ at which W_t^N and p_t are differentiable. Combining the fact that $\dot{p}_t = p_t(1 - p_t)(\varepsilon + \lambda N_t)$ with (7), we

obtain:

$$\dot{W}_t^N = (r + (\varepsilon + \lambda N_t)(1 - p_t)) W_t^N. \quad (8)$$

Furthermore, because $W_t^N = 2p_t - 1$ for all $t \in (t_1^*, t_2^*)$, we must have:

$$\dot{W}_t^N = 2\dot{p}_t = 2p_t(1 - p_t)(\varepsilon + \lambda N_t). \quad (9)$$

Combining (8), (9) and the fact that $W_t^N = 2p_t - 1$ then yields

$$N_t = \frac{r(2p_t - 1)}{\lambda(1 - p_t)} - \frac{\varepsilon}{\lambda}$$

for almost all $t \in (t_1^*, t_2^*)$. By continuity of p_t and right-continuity of N_t , the identity must then hold for all $t \in (t_1^*, t_2^*)$. \blacksquare

As an immediate corollary of Lemma A.6 we obtain:

Corollary A.1. *The posterior at all $t \in (t_1^*, t_2^*)$ evolves according to the following ordinary differential equation:*

$$\dot{p}_t = rp_t(2p_t - 1).$$

Given some initial condition $p = p_{t_1^*}$, this ordinary differential equation admits a unique solution, given by:

$$p_t = \frac{p_{t_1^*}}{2p_{t_1^*} - e^{r(t-t_1^*)}(2p_{t_1^*} - 1)}.$$

A.2.2 Characterization of Cutoff Times

Lemma A.7. *Let $N_{t \geq 0}$ be an equilibrium with corresponding no-news posterior $p_{t \geq 0}$ and cutoff times t_1^* and t_2^* as defined by Equation (2), and let $\Lambda_{t \geq 0} := \lambda \bar{N}_{t \geq 0}$ describe the evolution of the economy's potential for social learning. Then*

- (i). $t_2^* = \inf\{t \geq 0 : \Lambda_t < \Lambda^*(p_t)\}$; and
- (ii). $t_1^* = \min\{t_2^*, \sup\{t \geq 0 : p_t < p^*\}\}$.³²

Proof. We first prove both bullet points under the assumption that either $\varepsilon > 0$ or $p_0 > \frac{1}{2}$. Note that in this case Lemma A.4 implies that $\lim_{t \rightarrow \infty} 2p_t - 1 > \lim_{t \rightarrow \infty} W_t$, whence $t_2^* < +\infty$. Moreover, p_t is strictly increasing for all $t > 0$.

For the first bullet point, note that by definition of $t_2^* := \sup\{t \geq 0 : N_t < \rho \bar{N}_t\}$, we have that $2p_t - 1 \geq W_t = G(p_t, \Lambda_t)$ for all $t \geq t_2^*$. This implies that $\Lambda_{t_2^*} \leq \Lambda^*(p_{t_2^*})$. Moreover,

³²By convention, if $\{t \geq 0 : p_t < p^* = \frac{1}{2}\} = \emptyset$, then $\sup\{t \geq 0 : p_t < p^* = \frac{1}{2}\} := 0$.

for all $t > t_2^*$, $\Lambda_t < \Lambda_{t_2^*}$ and $p_t > p_{t_2^*}$, so since Λ^* is non-decreasing we have $\Lambda_t < \Lambda^*(p_t)$. Suppose that $0 < t_2^*$. Then by continuity we must have $2p_{t_2^*} - 1 = W_{t_2^*} = G(p_{t_2^*}, \Lambda_{t_2^*})$ and so $\Lambda_{t_2^*} = \Lambda^*(p_{t_2^*})$. But since for all $s < t_2^*$ we have $\Lambda_s \geq \Lambda_{t_2^*}$ and $p_s < p_{t_2^*}$, this implies $\Lambda_s \geq \Lambda^*(p_s)$. This establishes (i).

For (ii), it suffices to prove the following three claims:

- (a) If $t_2^* > 0$, then $p_{t_2^*} < p^\sharp$.
- (b) If $t_1^* > 0$, then $p_{t_1^*} \leq \bar{p}$.
- (c) If $t_1^* < t_2^*$, then $p_{t_1^*} \geq \bar{p}$.

Indeed, given (a) and (b), we have that if $0 < t_1^* = t_2^*$, then $p_{t_1^*} \leq p^*$. Given (a)-(c), we have that if $0 < t_1^* < t_2^*$, then $p_{t_1^*} = \bar{p} = p^*$. If $0 = t_1^* < t_2^*$, then (c) implies that $p_0 \geq \bar{p} = p^*$. In all three cases (ii) readily follows. Finally, if $0 = t_1^* = t_2^*$, then there is nothing to prove.

For claim (a), recall from the above that if $t_2^* > 0$, then $\Lambda_{t_2^*} = \Lambda^*(p_{t_2^*})$, whence $p_{t_2^*} < p^\sharp$ because $\Lambda^*(p^\sharp) = +\infty$.

For claim (b), note that if $t_1^* > 0$, then for all $t < t_1^*$, we have $N_t = 0$. Then for all $t < t_1^*$, $W_t \geq 2p_t - 1$ and by the proof of Lemma A.3, $\dot{W}_t = (r + (1 - p_t)\varepsilon)W_t$. Since $W_{t_1^*} = 2p_{t_1^*} - 1$, we must then have

$$\begin{aligned} 0 &\geq \lim_{\tau \uparrow t_1^*} \dot{W}_\tau - 2\dot{p}_\tau = (r + (1 - p_{t_1^*})\varepsilon)(2p_{t_1^*} - 1) - 2p_{t_1^*}(1 - p_{t_1^*})\varepsilon \\ &= r(2p_{t_1^*} - 1) - \varepsilon(1 - p_{t_1^*}), \end{aligned}$$

which implies that

$$p_{t_1^*} \leq \frac{\varepsilon + r}{\varepsilon + 2r} =: \bar{p}.$$

Finally, for claim (c), note that if $t_1^* < t_2^*$, then Lemma A.6 implies that for all $\tau \in (t_1^*, t_2^*)$,

$$0 \leq N_\tau = \frac{r(2p_\tau - 1)}{\lambda(1 - p_\tau)} - \frac{\varepsilon}{\lambda}.$$

This implies that for all $\tau \in (t_1^*, t_2^*)$,

$$p_\tau \geq \frac{\varepsilon + r}{\varepsilon + 2r} =: \bar{p},$$

and hence by continuity $p_{t_1^*} \geq \bar{p}$ as claimed. This proves the lemma when either $\varepsilon > 0$ or $p_0 > \frac{1}{2}$. Finally, if $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$, then by Lemma A.5 $N_t = 0$ for all t . Thus, by definition, $t_1^* = t_2^* = +\infty$. Moreover, $p_t = p_0 \leq \frac{1}{2}$ and $\Lambda_t = \Lambda_0 > 0$ for all t , so $\inf\{t : \Lambda_t < \Lambda^*(p_t) = 0\} = \sup\{t : p_t < p^* = \frac{1}{2}\} = +\infty$, as required. \blacksquare

Given Lemmas A.6 and A.7, it is immediate that if an equilibrium exists, then it must take the form of the adoption flow given by Equation (4) in Theorem 1. Moreover, it is easy to see that given initial parameters, Equation (4) uniquely pins down the times t_1^* and t_2^* as well as the joint evolution of p_t and N_t at all times (we elaborated on this in the main text), and that whenever $t_1^* < t_2^* < +\infty$, then $2p_t - 1 = W_t$ for all $t \in [t_1^*, t_2^*]$. Provided feasibility is satisfied, it is then easy to check that this adoption flow constitutes an equilibrium.

A.2.3 Feasibility

It remains to check feasibility, which is non-trivial only at times $t \in (t_1^*, t_2^*)$.

Lemma A.8. *Suppose $N_{t \geq 0}$ is an adoption flow satisfying Equation (4) in Theorem 1 such that $t_1^* < t_2^*$. Then for all $t \in (t_1^*, t_2^*)$, $N_t \leq \rho \bar{N}_t$.*

Proof. It suffices to show that

$$\lim_{t \uparrow t_2^*} N_t \leq \rho \bar{N}_{t_2^*}.$$

The lemma then follows immediately since $\rho \bar{N}_t - N_t$ is strictly decreasing in t at all times in (t_1^*, t_2^*) .

To see this, suppose by way of contradiction that $\rho \bar{N}_{t_2^*} < \lim_{t \uparrow t_2^*} N_t$. By continuity this means that there exists some $\nu > 0$ such that $\rho \bar{N}_t < N_t$ for all $t \in (t_2^* - \nu, t_2^*)$. Note that from the indifference condition at t_2^* , we have that $2p_{t_2^*} - 1 = G(p_{t_2^*}, \lambda \bar{N}_{t_2^*})$. Furthermore because $\Lambda^*(p_t)$ is increasing in t , $2p_t - 1 < G(p_t, \Lambda_t)$ for all $t < t_2^*$.

Since at all $t \in (t_2^* - \nu, t_2^*)$, $N_t > \rho \bar{N}_t$, this implies that $W_t > G(p_t, \Lambda_t) > 2p_t - 1$. But this is a contradiction since we already checked that the described adoption flow satisfies the condition that $W_t = 2p_t - 1$ for all $t \in (t_1^*, t_2^*)$. ■

A.3 Proof of Theorem 2 (Good News Equilibrium)

Theorem 2 follows immediately from the following two lemmas:

Lemma A.9. *Let $N_{t \geq 0}$ be an equilibrium with associated cutoff times t_1^* and t_2^* given by Equation (3). Then $t_1^* = t_2^* =: t^*$.*

Proof. Suppose for a contradiction that $t_1^* < t_2^*$. From the definition of these cutoffs and Lemma 1, we have that $2p_t - 1 = W_t$ for all $t \in (t_1^*, t_2^*)$. Then for all $t \in (t_1^*, t_2^*)$ and

$\Delta \in (0, t_2^* - t)$ we have:

$$W_t = p_t \int_t^{t+\Delta} (\varepsilon + \lambda N_\tau) e^{-\int_t^\tau (\varepsilon + \lambda N_s) ds} e^{-r(\tau-t)} \frac{\rho}{\rho + r} d\tau + \left((1 - p_t) + p_t e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) e^{-r\Delta} (2p_{t+\Delta} - 1),$$

where the first term represents a breakthrough arriving at some $\tau \in (t, t + \Delta)$ in which case consumers adopt from then on, yielding a payoff of $e^{-r(\tau-t)} \frac{\rho}{\rho+r}$; and the second term represents no breakthrough arriving prior to $t + \Delta$ in which case, due to indifference, consumers' payoff can be written as $e^{-r\Delta} (2p_{t+\Delta} - 1)$.

Note that we must have $p_t \geq \frac{1}{2}$ on (t_1^*, t_2^*) , since W_t is bounded below by 0. Moreover, by the definition of t_2^* , there exists $t \in (t_1^*, t_2^*)$ such that $N_t > 0$. By right-continuity of N , we can pick $\Delta \in (0, t_2^* - t)$ sufficiently small such that $N_\tau > 0$ for all $\tau \in (t, t + \Delta)$. Then,

$$\begin{aligned} & p_t \int_t^{t+\Delta} (\varepsilon + \lambda N_\tau) e^{-\int_t^\tau (\varepsilon + \lambda N_s) ds} e^{-r(\tau-t)} \frac{\rho}{\rho + r} d\tau \\ & < p_t \int_t^{t+\Delta} (\varepsilon + \lambda N_\tau) e^{-\int_t^\tau (\varepsilon + \lambda N_s) ds} \frac{\rho}{\rho + r} d\tau = p_t \left(1 - e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) \frac{\rho}{\rho + r}. \end{aligned}$$

This implies that

$$\begin{aligned} W_t & < p_t \left(1 - e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) \frac{\rho}{\rho + r} \\ & + \left((1 - p_t) + p_t e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) (2p_{t+\Delta} - 1) \\ & \leq p_t \left(1 - e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) + \left((1 - p_t) + p_t e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) (2p_{t+\Delta} - 1) \\ & = 2p_t - 1, \end{aligned}$$

where the final equality comes from Bayesian updating. This contradicts $W_t = 2p_t - 1$.

Thus, $t_1^* = t_2^*$. ■

Lemma A.10. *Let $N_{t \geq 0}$ be an equilibrium with corresponding cutoff time $t^* := t_1^* = t_2^*$ and no-news posterior $p_{t \geq 0}$. Then*

$$p_t \leq p^* \Leftrightarrow t \geq t^*,$$

where

$$p^* = \frac{(\varepsilon + r)(r + \rho)}{2(\varepsilon + r)(r + \rho) - \varepsilon\rho}.$$

Proof. Define

$$H_t := p_t \int_0^\infty (\varepsilon + \lambda N_{t+\tau}) e^{-(\varepsilon\tau + \int_t^{t+\tau} \lambda N_s ds)} \frac{\rho}{r + \rho} e^{-r\tau} d\tau.$$

Thus, H_t represents a consumer's expected value to waiting at time t given that from t on he adopts only if there has been a breakthrough and given that the population's flow of adoption follows $N_{s \geq 0}$. By optimality of W_t , we must have $H_t \leq W_t$ for all t . For any posterior $p \in (0, 1)$, let

$$H(p, 0) := p \int_0^\infty \varepsilon e^{-\varepsilon\tau} \frac{\rho}{r + \rho} e^{-r\tau} d\tau = p \frac{\varepsilon\rho}{(\varepsilon + r)(r + \rho)}.$$

That is, $H(p, 0)$ represents a consumer's expected value to waiting at posterior p , given that he adopts only once there has been a breakthrough and given that breakthroughs are only generated exogenously.

Now note that by definition of t^* , $N_t > 0$ if and only if $t < t^*$. This implies that $H(p_t, 0) < H_t$ if $t < t^*$ and $H(p_t, 0) = H_t = W_t$ if $t \geq t^*$; moreover, $2p_t - 1 \geq W_t$ if $t < t^*$ and $2p_t - 1 \leq W_t$ if $t \geq t^*$. Finally, note that $p^* := \frac{(\varepsilon+r)(r+\rho)}{2(\varepsilon+r)(r+\rho)-\varepsilon\rho}$ has the property that $2p - 1 \leq H(p, 0)$ if and only if $p \leq p^*$.

Combining these observations, we have that if $t < t^*$, then $2p_t - 1 \geq W_t \geq H_t > H(p_t, 0)$, so $p_t > p^*$. And if $t \geq t^*$, then $2p_t - 1 \leq W_t = H(p_t, 0)$, so $p_t \leq p^*$, as claimed. ■

References

- BANDIERA, O. AND I. RASUL (2006): "Social Networks and Technology Adoption in Northern Mozambique," *Economic Journal*, 116, 869–902.
- BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 797–817.
- BAPTISTA, R. (1999): "The Diffusion of Process Innovations: A Selective Review," *International Journal of the Economics of Business*, 6, 107–129.
- BASS, F. M. (1969): "A New Product Growth Model for Consumer Durables," *Management Science*, 15, 215–227.
- BERGEMANN, D. AND J. VÄLIMÄKI (1997): "Market Diffusion with Two-Sided Learning," *RAND Journal of Economics*, 773–795.
- BESLEY, T. AND A. CASE (1993): "Modeling Technology Adoption in Developing Countries," *American Economic Review*, 83, 396–402.
- (1994): "Diffusion as a Learning Process: Evidence from HYV Cotton," *Woodrow Wilson School Discussion Paper #174*.
- BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 992–1026.
- BOARD, S. AND M. MEYER-TER VEHN (2013): "Reputation for Quality," *Econometrica*, 81, 2381–2462.
- (2020): "Learning dynamics in social networks," Working Paper.
- BONATTI, A. AND J. HÖRNER (2011): "Collaborating," *American Economic Review*, 101, 632–663.

- (2017): “Career concerns with exponential learning,” *Theoretical Economics*, 12, 425–475.
- BUCHWALD, H. AND D. M. OIEN (2009): “Metabolic/Bariatric Surgery Worldwide 2008,” *Obesity Surgery*, 19, 1605–1611.
- (2013): “Metabolic/Bariatric Surgery Worldwide 2011,” *Obesity Surgery*, 23, 427–436.
- CABRAL, L. (2012): “Lock in and switch: Asymmetric information and new product diffusion,” *Quantitative Marketing and Economics*, 10, 375–392.
- CHAMLEY, C. AND D. GALE (1994): “Information Revelation and Strategic Delay in a Model of Investment,” *Econometrica*, 1065–1085.
- CHE, Y.-K. AND J. HÖRNER (2018): “Recommender systems as mechanisms for social learning,” *The Quarterly Journal of Economics*, 133, 871–925.
- CONLEY, T. G. AND C. R. UDRY (2010): “Learning about a New Technology: Pineapple in Ghana,” *American Economic Review*, 100, 35–69.
- CRIPPS, M. W. AND C. D. THOMAS (2019): “Strategic experimentation in queues,” *Theoretical Economics*, 14, 647–708.
- DAVID, P. A. (1969): *A Contribution to the Theory of Diffusion*, Research Center in Economic Growth, Stanford University.
- DAVIES, S. (1979): *The Diffusion of Process Innovations*, Cambridge: CUP Archive.
- DUPAS, P. (2014): “Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment,” *Econometrica*, 82, 197–228.
- FARRELL, J. AND G. SALONER (1985): “Standardization, Compatibility, and Innovation,” *RAND Journal of Economics*, 70–83.
- (1986): “Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation,” *American Economic Review*, 940–955.
- FOSTER, A. D. AND M. R. ROSENZWEIG (1995): “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture,” *Journal of Political Economy*, 1176–1209.
- FRICK, M. AND Y. ISHII (2014): “Innovation Adoption by Forward-Looking Social Learners,” *December 2014 version, available at <http://sites.google.com/site/mirafrickecon/research>*.
- GEROSKI, P. A. (2000): “Models of Technology Diffusion,” *Research Policy*, 29, 603–625.
- GORT, M. AND S. KLEPPER (1982): “Time Paths in the Diffusion of Product Innovations,” *Economic Journal*, 92, 630–653.
- GRILICHES, Z. (1957): “Hybrid Corn: An Exploration in the Economics of Technological Change,” *Econometrica*, 25, 501–522.
- GUENTHNER, J., B.-H. LIN, AND A. E. LEVI (1991): “The Influence of Microwave Ovens on the Demand for Fresh and Frozen Potatoes,” *Journal of Food Distribution Research*, 22, 45–52.
- GUO, Y. (2016): “Dynamic Delegation of Experimentation,” *American Economic Review*, 106, 1969–2008.
- HALAC, M., N. KARTIK, AND Q. LIU (2016): “Optimal contracts for experimentation,” *The Review of Economic Studies*, 83, 1040–1091.
- (2017): “Contests for experimentation,” *Journal of Political Economy*, 125, 1523–1569.
- HALAC, M. AND I. KREMER (2020): “Experimenting with career concerns,” *American Economic Journal: Microeconomics*, 12, 260–88.
- HÖRNER, J. AND A. SKRZYPACZ (2016): “Learning, Experimentation and Information Design,” Mimeo.
- HOYER, W. D., D. J. MACINNIS, AND R. PIETERS (2012): *Consumer Behavior*, South-Western, 6th ed.
- JENSEN, R. (1982): “Adoption and Diffusion of an Innovation of Uncertain Profitability,” *Journal of Eco-*

- conomic Theory*, 27, 182–193.
- JOVANOVIC, B. AND S. LACH (1989): “Entry, Exit, and Diffusion with Learning by Doing,” *American Economic Review*, 690–699.
- KARSHENAS, M. AND P. L. STONEMAN (1993): “Rank, Stock, Order, and Epidemic effects in the Diffusion of New Process Technologies: An Empirical Model,” *RAND Journal of Economics*, 503–528.
- KEILLOR, B. D. (2007): *Marketing in the 21st Century*, vol. 4, Greenwood Publishing Group.
- KELLER, G. AND S. RADY (2010): “Strategic Experimentation with Poisson Bandits,” *Theoretical Economics*, 5, 275–311.
- (2015): “Breakdowns,” *Theoretical Economics*, 10, 175–202.
- KELLER, G., S. RADY, AND M. CRIPPS (2005): “Strategic Experimentation with Exponential Bandits,” *Econometrica*, 73, 39–68.
- KHROMENKOVA, D. (2015): “Collective experimentation with breakdowns and breakthroughs,” Working Paper.
- KREMER, M. AND E. MIGUEL (2007): “The Illusion of Sustainability,” *Quarterly Journal of Economics*, 122, 1007–1065.
- LAIHO, T. AND J. SALMI (2018): “Social Learning and Monopoly Pricing with Forward Looking Buyers,” Working Paper.
- MAHAJAN, V. AND R. A. PETERSON (1985): *Models for Innovation Diffusion*, vol. 48, Sage.
- MANSFIELD, E. (1961): “Technical Change and the Rate of Imitation,” *Econometrica*, 29, 741–766.
- (1968): *Industrial Research and Technological Innovation: An Econometric Analysis*, New York: Norton.
- MUNSHI, K. (2004): “Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution,” *Journal of Development Economics*, 73, 185–213.
- MURTO, P. AND J. VÄLIMÄKI (2011): “Learning and Information Aggregation in an Exit Game,” *Review of Economic Studies*, 78, 1426–1461.
- SMITH, L. AND P. SØRENSEN (2000): “Pathological Outcomes of Observational Learning,” *Econometrica*, 68, 371–398.
- STRULOVICI, B. (2010): “Learning while Voting: Determinants of Collective Experimentation,” *Econometrica*, 78, 933–971.
- THOMAS, C. (2019a): “Experimentation with reputation concerns—Dynamic signalling with changing types,” *Journal of Economic Theory*, 179, 366–415.
- (2019b): “Strategic experimentation with congestion,” *American Economic Journal: Microeconomics*, forthcoming.
- WIERSEMA, F. D. AND R. D. BUZZELL (1979): “Note on the Microwave Oven Industry,” *Harvard Business School Case Study*.
- WILLIAMS, A. (2014): “Surveys of Microwave Ovens in US Homes,” *Lawrence Berkeley National Laboratory Paper #5947E*.
- WOLITZKY, A. (2018): “Learning from Others’ Outcomes,” *American Economic Review*, 108, 2763–2801.
- YOUNG, H. P. (2009): “Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning,” *American Economic Review*, 99, 1899–1924.

Online Appendix to “Innovation Adoption by Forward-Looking Social Learners”

Mira Frick and Yuhta Ishii

B Omitted Proofs and Results

B.1 Proofs of Lemmas A.1–A.5

Proof of Lemma A.1: Note the following recursive formulations for W_t^N . If learning is via bad news, then

$$W_t^N = \int_t^\infty \rho e^{-(r+\rho)(s-t)} \frac{p_t^N}{p_s^N} \max \{ (2p_s^N - 1), W_s^N \} ds.$$

If learning is via good news, W_t^N satisfies:

$$W_t^N = \int_t^\infty \rho e^{-(r+\rho)(s-t)} \left(p_t^N \left(1 - e^{-\int_t^s (\varepsilon + \lambda N_k) dk} \right) + \frac{p_t^N e^{-\int_t^s (\varepsilon + \lambda N_k) dk}}{p_s^N} \max \{ (2p_s^N - 1), W_s^N \} \right) ds.$$

From this it is immediate that W_t^N is continuous in t . ■

Proof of Lemma A.2: Suppose that $W_t^N < 2p_t^N - 1$ for some $t > 0$. By Lemma A.1 W_t^N is continuous in t , and so is $2p_t^N - 1$. Hence there exists $\nu > 0$ such that $W_\tau^N < 2p_\tau^N - 1$ for all $\tau \in (t - \nu, t + \nu)$. Because N is an equilibrium this implies that $N_\tau = \rho \bar{N}_\tau$ for all $\tau \in (t - \nu, t + \nu)$. Thus, N_τ is continuous at all $\tau \in (t - \nu, t + \nu)$. From this it is immediate that W_τ^N is continuously differentiable in τ for all $\tau \in (t - \nu, t + \nu)$, because we have that

$$\begin{aligned} W_\tau^N &= \int_\tau^{t+\nu} \rho e^{-(\rho+r)(s-\tau)} \left(p_\tau^N e^{-\int_\tau^s (\varepsilon_G + \lambda_G N_x) dx} - (1 - p_\tau^N) e^{-\int_\tau^s (\varepsilon_B + \lambda_B N_x) dx} \right) ds \\ &\quad + e^{-(r+\rho)(t+\nu-\tau)} \left(p_\tau^N e^{-\int_\tau^{t+\nu} (\varepsilon_G + \lambda_G N_x) dx} + (1 - p_\tau^N) e^{-\int_\tau^{t+\nu} (\varepsilon_B + \lambda_B N_x) dx} \right) W_{t+\nu}^N \\ &\quad + \int_\tau^{t+\nu} \rho e^{-(\rho+r)(s-\tau)} p_\tau^N \left(1 - e^{-\int_\tau^s (\varepsilon_G + \lambda_G N_x) dx} \right) ds \\ &\quad + e^{-(r+\rho)(t+\nu-\tau)} p_\tau^N \left(1 - e^{-\int_\tau^{t+\nu} (\varepsilon_G + \lambda_G N_x) dx} \right) \frac{\rho}{\rho + r}. \end{aligned}$$

The derivative of W_τ^N can be computed using Ito’s Lemma for processes with jumps. Given the perfect Poisson learning structure, the derivation is simple and we provide it here for completeness.

As above, for any $\Delta < t + \nu - \tau$ we can rewrite W_τ^N as

$$\begin{aligned}
W_\tau^N &= \int_\tau^{\tau+\Delta} \rho e^{-(\rho+r)(s-\tau)} \left(p_\tau^N e^{-\int_\tau^s (\varepsilon_G + \lambda_G N_x) dx} - (1 - p_\tau^N) e^{-\int_\tau^s (\varepsilon_B + \lambda_B N_x) dx} \right) ds \\
&+ e^{-(r+\rho)\Delta} \left(p_\tau^N e^{-\int_\tau^{\tau+\Delta} (\varepsilon_G + \lambda_G N_x) dx} + (1 - p_\tau^N) e^{-\int_\tau^{\tau+\Delta} (\varepsilon_B + \lambda_B N_x) dx} \right) W_{\tau+\Delta}^N \\
&+ \int_\tau^{\tau+\Delta} \rho e^{-(r+\rho)(s-\tau)} p_\tau^N \left(1 - e^{-\int_\tau^s (\varepsilon_G + \lambda_G N_x) dx} \right) ds \\
&+ e^{-(r+\rho)\Delta} p_\tau^N \left(1 - e^{-\int_\tau^{\tau+\Delta} (\varepsilon_G + \lambda_G N_x) dx} \right) \frac{\rho}{\rho + r}.
\end{aligned}$$

Since this is true for all $\Delta \in (0, t + \nu - \tau)$, the right hand side of this identity, which we denote R_Δ , is continuously differentiable with respect to Δ and satisfies $\frac{d}{d\Delta} R_\Delta \equiv 0$. Taking the limit as $\Delta \rightarrow 0$ and since $\dot{W}_\tau^N = \lim_{\Delta \rightarrow 0} \frac{d}{d\Delta} W_{\tau+\Delta}^N$ by continuous differentiability, we then obtain:

$$\begin{aligned}
\dot{W}_\tau^N &= (r + \rho + (\varepsilon_G + \lambda_G N_\tau) p_\tau + (\varepsilon_B + \lambda_B N_\tau)(1 - p_\tau)) W_\tau^N \\
&- \rho(2p_\tau - 1) - p_\tau(\varepsilon_G + \lambda_G N_\tau) \frac{\rho}{\rho + r}.
\end{aligned}$$

Plugging in $N_\tau = \rho \bar{N}_\tau$ yields the desired expression. ■

Proof of Lemma A.3: The proof of continuous differentiability of W_t^N follows along the same lines as in the proof of Lemma A.2. Lemma A.1 again implies that if $W_t^N > 2p_t^N - 1$, then there exists $\nu > 0$ such that $W_\tau^N > 2p_\tau^N - 1$ for all $\tau \in (t - \nu, t + \nu)$. By the definition of equilibrium, $N_\tau = 0$ for all $\tau \in (t - \nu, t + \nu)$.

Hence, W_τ^N satisfies

$$\begin{aligned}
W_\tau^N &= e^{-r(t+\nu-\tau)} \left(p_\tau^N e^{-\varepsilon_G(t+\nu-\tau)} + (1 - p_\tau^N) e^{-\varepsilon_B(t+\nu-\tau)} \right) W_{t+\nu}^N \\
&+ p_\tau^N \int_\tau^{t+\nu} \varepsilon_G e^{-(\varepsilon_G+r)s} \frac{\rho}{\rho + r} ds.
\end{aligned}$$

From this it is again immediate that W_τ^N is continuously differentiable in τ .

To compute the derivative, we proceed as above, rewriting W_τ^N as

$$W_\tau^N = e^{-r\Delta} \left(p_\tau^N e^{-\varepsilon_G \Delta} + (1 - p_\tau^N) e^{-\varepsilon_B \Delta} \right) W_{t+\Delta}^N + p_\tau \int_\tau^{\tau+\Delta} \varepsilon_G e^{-(\varepsilon_G+r)s} \frac{\rho}{\rho + r} ds$$

for any $\Delta < t + \nu - \tau$.

Differentiating both sides of the above equality with respect to Δ and taking the limit as $\Delta \rightarrow 0$,

we obtain:

$$\dot{W}_\tau^N = (r + p_\tau^N \varepsilon_G + (1 - p_\tau^N) \varepsilon_B) W_\tau^N - p_\tau^N \varepsilon_G \frac{\rho}{\rho + r},$$

as claimed. ■

Proof of Lemma A.4: Consider first the case in which $\varepsilon > 0$. Then trivially $p_t^N \rightarrow 1$ as $t \rightarrow \infty$. But for any t , $\frac{\rho}{\rho+r} (2p_t^N - 1) \leq W_t^N \leq \frac{\rho}{\rho+r}$. This implies that $\lim_{t \rightarrow \infty} W_t^N = \frac{\rho}{\rho+r}$ as claimed.

Now suppose that $\varepsilon = 0$ and $p_0 > 1/2$. Then note that $W_t^N \leq 2p_t^N - 1$ for all t : Indeed, suppose that $W_t^N > 2p_t^N - 1$ for some t . We can't have that $W_s^N > 2p_s^N - 1$ for all $s \geq t$, since otherwise $W_t^N = 0$, contradicting $W_t^N > 2p_t^N - 1 > 0$. But then we can find $s > t$ such that $W_s^N = 2p_s^N - 1$ and $W_{s'}^N > 2p_{s'}^N - 1$ for all $s' \in (t, s)$. This implies $N_{s'} = 0$ for all s' , and hence $W_t^N = e^{-r(s-t)} W_s^N = e^{-r(s-t)} (2p_s^N - 1) = e^{-r(s-t)} (2p_t^N - 1)$, again contradicting $W_t^N > 2p_t^N - 1 > 0$.

Let $N^* := \lim_{t \rightarrow \infty} \int_0^t N_s ds = \sup_t \int_0^t N_s ds \leq \bar{N}_0$. Let $p^* := \lim_{t \rightarrow \infty} p_t^N = \sup_t p_t^N$. For any $\nu > 0$ we can find t^* such that whenever $t > t^*$, then $e^{-\lambda \int_{t^*}^t N_s ds} > 1 - \nu$. Because $2p_t^N - 1 \geq W_t^N$ for all t , we can then rewrite the value to waiting at time t as:

$$\begin{aligned} W_t^N &= \int_t^\infty \rho e^{-(r+\rho)\tau} \left(p_t^N - (1 - p_t^N) e^{-\lambda \int_t^\tau N_s ds} \right) d\tau \\ &\leq \frac{\rho}{r + \rho} (p_t^N - (1 - p_t^N)(1 - \nu)) \end{aligned}$$

for all $t > t^*$. Moreover, by optimality $W_t^N \geq \frac{\rho}{\rho+r} (2p_t^N - 1)$ for all t , so combining we have

$$\frac{\rho}{\rho + r} (2p^* - 1) \leq \liminf_{t \rightarrow \infty} W_t^N \leq \limsup_{t \rightarrow \infty} W_t^N \leq \frac{\rho}{r + \rho} (p^* - (1 - p^*)(1 - \nu)).$$

Since this is true for all $\nu > 0$, it follows that

$$\lim_{t \rightarrow \infty} W_t^N = \frac{\rho}{r + \rho} (2p^* - 1).$$

But the above is strictly less than $2p^* - 1$, so for all t sufficiently large we must have $2p_t^N - 1 > W_t^N$. Then for all t sufficiently large, we have $N_t = \rho \bar{N}_t$. Thus, $N^* = \bar{N}_0$ and therefore $p^* = \mu(\varepsilon, \Lambda_0, p_0)$. ■

Proof of Lemma A.5: Suppose that $N_{t \geq 0}$ is an equilibrium and suppose for a contradiction that $t_1^* := \inf\{t : N_t > 0\} < \infty$. Pick $t \geq t_1^*$ such that $N_t > 0$. By right-continuity of N , we have $N_\tau > 0$ for all $\tau > t$ sufficiently close to t . This implies that

$$\int_{t_1^*}^\infty \rho e^{-(r+\rho)(s-t)} \left(p_{t_1^*}^N - (1 - p_{t_1^*}^N) e^{-\int_{t_1^*}^s \lambda N_k dk} \right) ds > \frac{\rho}{r + \rho} (2p_{t_1^*}^N - 1) \geq 2p_{t_1^*}^N - 1, \quad (10)$$

where the second inequality holds because $p_{t_1^*}^N = p_0 \leq \frac{1}{2}$. But the integral on the left-hand side is

the expected payoff at time t_1^* to adopting at the first opportunity in the future, conditional on no breakdown having occurred prior to this opportunity. By optimality of the value to waiting, this is weakly less than $W_{t_1^*}^N$. Hence, (10) implies that $W_{t_1^*}^N > 2p_{t_1^*} - 1$. By continuity of W^N and p^N , it follows that for all $s \geq t_1^*$ sufficiently close to t_1^* , $W_s^N > 2p_s^N - 1$ and hence $N_s = 0$, contradicting the definition of t_1^* .

This leaves $N \equiv 0$ as the only candidate equilibrium. In this case $W_t^N = 0 \geq 2p_0 - 1 = 2p_t^N - 1$ for all t , so this is indeed an equilibrium. ■

B.2 Proof of Lemma 2

Proof. Assume that $p_0 \in (0, p^\sharp)$ and impose Conditions 1 and 2. Define $\bar{\Lambda}_0 := \max\{\Lambda^*(p_0), \Lambda^*(\bar{p})\}$. Consider any Λ_0 and let $t_i^* := t_i^*(\Lambda_0)$ for $i = 1, 2$. We show that $t_1^* < t_2^*$ if and only if $\Lambda_0 > \bar{\Lambda}_0$.

Suppose first that $\Lambda_0 > \bar{\Lambda}_0$. By the proof of the first part of Lemma A.7, we must have $t_2^* > 0$ and $\Lambda_{t_2^*} = \Lambda^*(p_{t_2^*})$. If $t_1^* = t_2^* =: t^*$, then by claims (a) and (b) in the proof of Lemma A.7, we must have $p_{t^*} \leq \bar{p}$. But combining these statements, we get

$$\Lambda_{t^*} = \Lambda_0 > \Lambda^*(\bar{p}) \geq \Lambda^*(p_{t^*}) = \Lambda_{t^*},$$

which is a contradiction.

Suppose conversely that $t_1^* < t_2^*$. Then by the proof of Lemma A.7, we have that $\Lambda^*(p_{t_1^*}) < \Lambda_{t_1^*} = \Lambda_0$. That proof also implies that if $0 < t_1^* < t_2^*$, then $p_{t_1^*} = \bar{p} \geq p_0$; and if $0 = t_1^* < t_2^*$, then $p_{t_1^*} = p_0 \geq \bar{p}$. Thus, either way $\Lambda_0 > \bar{\Lambda}_0$, as claimed. ■

B.3 Proof of Proposition 1 (Saturation Effect under Bad News)

Fix r, ρ, ε , and p_0 . Suppose Λ_0 is such that $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$. By Lemma 2 and Theorem 1, this means that Conditions 1 and 2 are satisfied, $p_0 < p^\sharp$, and $\Lambda_0 > \bar{\Lambda}_0$, where $\bar{\Lambda}_0 := \max\{\Lambda^*(p_0), \Lambda^*(\bar{p})\}$ is as given by Lemma 2. Consider any $\hat{\Lambda}_0 > \Lambda_0$.

B.3.1 Proof of part (i) (welfare neutrality)

Write $\Lambda_0^1 := \Lambda_0$ and $\Lambda_0^2 := \hat{\Lambda}_0$, with corresponding cutoff times t_1^i and t_2^i , value to waiting W_t^i , and no-news posteriors p_t^i for $i = 1, 2$. Since $t_1^1 < t_2^1$ and $\Lambda_0^2 > \Lambda_0^1 > \bar{\Lambda}_0$, Lemma 2 implies $t_1^2 < t_2^2$. Moreover, by the proof of Lemma A.7, we have $\max\{p_0, \bar{p}\} = p_{t_1^1}^1 = p_{t_1^2}^2$. Because $N_t^i = 0$ for all $t < t_1^i$ for both $i = 1, 2$, this implies that $t_1^1 = t_1^2 = t_1$. Then

$$W_{t_1}^2 = 2p_{t_1}^2 - 1 = 2p_{t_1}^1 - 1 = W_{t_1}^1.$$

But since there is no adoption until t_1 , we have $W_0^i = e^{-rt_1} \frac{p_{t_1}^i}{p_0} W_{t_1}^i$ for $i = 1, 2$, whence $W_0^1 = W_0^2$, as claimed. ■

B.3.2 Proof of part (ii) (non-monotonicity of learning)

We first prove the following lemma:

Lemma B.1. *Suppose that $\hat{\Lambda}_0 = \hat{\lambda}\hat{N}_0 > \Lambda_0 = \lambda\bar{N}_0 > \bar{\Lambda}_0$, with corresponding equilibrium flows of adoption $\hat{N}_{t \geq 0}$ and $N_{t \geq 0}$. Then*

(i). $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0)$.

(ii). $0 < t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$.

(iii). For all $t < t_2^*(\Lambda_0)$, $\lambda N_t = \hat{\lambda}\hat{N}_t$.

Proof. For (i), note that by the proof of Lemma A.7, time t_1^* under each of Λ_0 and $\hat{\Lambda}_0$ is pinned down by the fact that $\max\{p_0, \bar{p}\} = p_{t_1^*(\Lambda_0)}^{\Lambda_0} = p_{t_1^*(\hat{\Lambda}_0)}^{\hat{\Lambda}_0}$. Because up to time t_1^* learning is purely exogenous under both Λ_0 and $\hat{\Lambda}_0$, this immediately implies that $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0)$.

For (ii) and (iii), note first that by Lemma 2, we have $t_2^*(\hat{\Lambda}_0), t_2^*(\Lambda_0) > 0$. Let $t_2^* = \min\{t_2^*(\hat{\Lambda}_0), t_2^*(\Lambda_0)\}$. Then because $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0)$, the ODE in Corollary A.1 implies that at all times $t < t_2^*$, we have $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0} = p_t$. By Lemma A.6, this implies that for all $t < t_2^*$,

$$\lambda N_t = \hat{\lambda}\hat{N}_t. \quad (11)$$

Note that Equation 11 implies that

$$\Lambda_{t_2^*} = \Lambda_0 - \int_0^{t_2^*} \lambda N_t dt < \hat{\Lambda}_0 - \int_0^{t_2^*} \hat{\lambda}\hat{N}_t dt = \hat{\Lambda}_{t_2^*}.$$

Because $p_{t_2^*}^{\Lambda_0} = p_{t_2^*}^{\hat{\Lambda}_0}$, Lemma A.7 implies that $t_2^* = t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$.

From this and Equation 11, it is then immediate that $\lambda N_t = \hat{\lambda}\hat{N}_t$ for all $t < t_2^*(\Lambda_0)$. ■

Now we prove part (ii) of Proposition 1. By Lemma B.1, $t^* := t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$, $\lambda N_t = \hat{\lambda}\hat{N}_t$, and $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0}$ for all $t \leq t^*$, which proves the first bullet.

For the second bullet, we claim that there exists some $\nu > 0$ such that at all times $t \in (t^*, t^* + \nu)$, we have $p_t^{\Lambda_0} > p_t^{\hat{\Lambda}_0}$. To see this, we prove the following inequality for the equilibrium corresponding to Λ_0 :

$$\lim_{t \uparrow t^*} \lambda N_t < \lim_{t \downarrow t^*} \lambda N_t. \quad (12)$$

In other words, there is a discontinuity in the equilibrium flow of adoption at time t^* . Indeed, because $N_t = \rho\bar{N}_t$ for all $t \geq t^*$ and by continuity of \bar{N}_t , feasibility implies that $\lim_{t \uparrow t^*} \lambda N_t \leq \lim_{t \downarrow t^*} \lambda N_t$. Suppose for a contradiction that $\lim_{t \uparrow t^*} \lambda N_t = \lim_{t \downarrow t^*} \lambda N_t := \lambda N_{t^*}$. Then $\lambda N_{t^*} = \hat{\lambda}\hat{N}_{t^*}$. Moreover, for all $t > t^*$, we have $\lambda N_t = \rho\Lambda_{t^*}e^{-\rho(t-t^*)}$, which is strictly decreasing in t . On the other hand, $\hat{\lambda}\hat{N}_t$ satisfies

$$\hat{\lambda}\hat{N}_t = \begin{cases} \frac{r(2\hat{p}_t-1)}{(1-\hat{p}_t)} - \varepsilon & \text{if } t \in [t^*, t_2^*(\hat{\Lambda}_0)) \\ \rho\Lambda_{t_2^*(\hat{\Lambda}_0)}e^{-\rho(t-t_2^*(\hat{\Lambda}_0))} & \text{if } t \geq t_2^*(\hat{\Lambda}_0). \end{cases}$$

Thus, for $t \in [t^*, t_2^*(\hat{\Lambda}_0))$, $\hat{\lambda}\hat{N}_t$ is strictly *increasing* in t . This implies that $\hat{\lambda}\hat{N}_t > \lambda N_t$ for all $t \in [t^*, t_2^*(\hat{\Lambda}_0))$. But then by Equation 1,

$$p_{t_2^*(\hat{\Lambda}_0)}^{\hat{\Lambda}_0} > p_{t_2^*(\hat{\Lambda}_0)}^{\Lambda_0},$$

which by Lemma A.7 implies

$$\hat{\Lambda}_{t_2^*(\hat{\Lambda}_0)} = \Lambda^*(p_{t_2^*(\hat{\Lambda}_0)}^{\hat{\Lambda}_0}) > \Lambda^*(p_{t_2^*(\hat{\Lambda}_0)}^{\Lambda_0}) > \Lambda_{t_2^*(\hat{\Lambda}_0)}.$$

This yields that for all $t \geq t_2^*(\hat{\Lambda}_0)$

$$\hat{\lambda}\hat{N}_t = \rho e^{-\rho(t-t_2^*(\hat{\Lambda}_0))} \hat{\Lambda}_{t_2^*(\hat{\Lambda}_0)} > \rho e^{-\rho(t-t_2^*(\hat{\Lambda}_0))} \Lambda_{t_2^*(\hat{\Lambda}_0)} = \lambda N_t.$$

Thus, $\hat{\lambda}\hat{N}_t > \lambda N_t$ for *all* $t > t^*$ and hence $p_t^{\hat{\Lambda}_0} > p_t^{\Lambda_0}$ for all $t > t^*$. This implies $W_{t^*}^{\hat{\Lambda}_0} > W_{t^*}^{\Lambda_0}$. But this is a contradiction, because we have

$$W_{t^*}^{\hat{\Lambda}_0} = 2p_{t^*}^{\hat{\Lambda}_0} - 1 = 2p_{t^*}^{\Lambda_0} - 1 = W_{t^*}^{\Lambda_0}.$$

This proves that $\lim_{t \uparrow t^*} \lambda N_t < \lim_{t \downarrow t^*} \lambda N_t$. But then,

$$\lim_{t \downarrow t^*} \hat{\lambda}\hat{N}_t = \lim_{t \uparrow t^*} \hat{\lambda}\hat{N}_t = \lim_{t \uparrow t^*} \lambda N_t < \lim_{t \downarrow t^*} \lambda N_t.$$

Thus, there exists some $\nu > 0$ such that $\hat{\lambda}\hat{N}_t < \lambda N_t$ for all $t \in [t^*, t^* + \nu)$. Together with the fact that $p_{t^*}^{\Lambda_0} = p_{t^*}^{\hat{\Lambda}_0}$, this implies that $p_t^{\Lambda_0} > p_t^{\hat{\Lambda}_0}$ for all $t \in (t^*, t^* + \nu)$, proving the second bullet.

Finally, for the third bullet, observe first that there exists some $t > t^*$ such that $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0}$. If not, then by continuity of beliefs $p_t^{\Lambda_0} > p_t^{\hat{\Lambda}_0}$ for all $t > t^*$, and we have $W_{t^*}^{\hat{\Lambda}_0} < W_{t^*}^{\Lambda_0}$, again contradicting $W_{t^*}^{\hat{\Lambda}_0} = W_{t^*}^{\Lambda_0} = 2p_{t^*} - 1$. Then $\bar{t} := \sup\{s \in (t^*, t) : p_s^{\Lambda_0} > p_s^{\hat{\Lambda}_0}\}$ exists, with $\bar{t} > t^*$ by the second bullet point. Further, by continuity, $p_{\bar{t}}^{\Lambda_0} = p_{\bar{t}}^{\hat{\Lambda}_0}$, which implies $\int_0^{\bar{t}} \lambda N_s ds = \int_0^{\bar{t}} \hat{\lambda}\hat{N}_s ds$. This yields $\Lambda_{\bar{t}} < \hat{\Lambda}_{\bar{t}}$. But this implies that $\hat{\lambda}\hat{N}_t > \lambda N_t$ for all $t > \bar{t}$: Indeed, if $\bar{t} \geq t_2^*(\hat{\Lambda}_0)$, this is obvious. If $\bar{t} \in (t^*, t_2^*(\hat{\Lambda}_0))$, then we must have $\lambda N_s < \hat{\lambda}\hat{N}_s$ for some $s < \bar{t}$, which implies that $\lambda N_{s'} < \hat{\lambda}\hat{N}_{s'}$ for all $s' \in (s, t_2^*(\hat{\Lambda}_0))$, because N is strictly decreasing and \hat{N} is strictly increasing on this domain. To see that we also have $\lambda N_{s'} < \hat{\lambda}\hat{N}_{s'}$ for all $s' \geq t_2^*(\hat{\Lambda}_0)$, note that from the above

$$p_{t_2^*(\hat{\Lambda}_0)}^{\hat{\Lambda}_0} > p_{t_2^*(\hat{\Lambda}_0)}^{\Lambda_0},$$

which as above implies that

$$\hat{\Lambda}_{t_2^*(\hat{\Lambda}_0)} = \Lambda^*(p_{t_2^*(\hat{\Lambda}_0)}^{\hat{\Lambda}_0}) > \Lambda^*(p_{t_2^*(\hat{\Lambda}_0)}^{\Lambda_0}) > \Lambda_{t_2^*(\hat{\Lambda}_0)}.$$

Hence, $\hat{\lambda}\hat{N}_t > \lambda N_t$ for all $t > \bar{t}$. Thus, in either case we get that $p_t^{\hat{\Lambda}_0} > p_t^{\Lambda_0}$ for all $t > \bar{t}$, as claimed

by the third bullet. ■

B.3.3 Proof of part (iii) (slowdown in adoption)

We consider good and bad products separately.

Adoption of Good Products: Recall that by Lemma B.1, $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0) =: t_1^*$ and $\lambda N_t = \hat{\lambda} \hat{N}_t$ for all $t \in (t_1^*, t^*)$, where $t^* := t_2^*(\Lambda_0)$. Then for all $t < t^*$

$$\frac{N_t}{\bar{N}_0} = \frac{\lambda N_t}{\Lambda_0} = \frac{\hat{\lambda} \hat{N}_t}{\Lambda_0} \geq \frac{\hat{\lambda} \hat{N}_t}{\hat{\Lambda}_0} = \frac{\hat{N}_t}{\hat{\bar{N}}_0},$$

with strict inequality for all $t \in (t_1^*, t^*)$. Therefore, $A_t(\Lambda_0, G) \geq A_t(\hat{\Lambda}_0, G)$ for all $t < t^*$, with strict inequality for all $t \in (t_1^*, t^*)$.

Finally note that for all $t \geq t^*$, $N_t = \rho \bar{N}_t$ and so:

$$\begin{aligned} A_t(\Lambda_0, G) &= A_{t^*}(\Lambda_0, G) + (1 - e^{-\rho(t-t^*)}) (1 - A_{t^*}(\Lambda_0, G)) \\ A_t(\hat{\Lambda}_0, G) &\leq A_{t^*}(\hat{\Lambda}_0, G) + (1 - e^{-\rho(t-t^*)}) (1 - A_{t^*}(\hat{\Lambda}_0, G)) \end{aligned}$$

where the second inequality follows from feasibility. But because $A_{t^*}(\Lambda_0, G) > A_{t^*}(\hat{\Lambda}_0, G)$, it follows that $A_t(\Lambda_0, G) > A_t(\hat{\Lambda}_0, G)$ for all $t > t_1^*$, as claimed.

Adoption of Bad Products: Recall that $A_t(\lambda, \bar{N}_0, B)$ denotes the *expected* proportion of adopters at time t conditional on $\theta = B$, that is, letting $N_{t \geq 0}$ denote the associated equilibrium

$$\begin{aligned} A_t(\lambda, \bar{N}_0, B) &:= \int_0^t (\varepsilon + \lambda N_\tau) e^{-\int_0^\tau (\varepsilon + \lambda N_s) ds} \left(\int_0^\tau \frac{N_s}{\bar{N}_0} ds \right) d\tau + e^{-\int_0^t (\varepsilon + \lambda N_s) ds} \int_0^t \frac{N_s}{\bar{N}_0} ds \\ &= \int_0^t \frac{N_\tau}{\bar{N}_0} e^{-\int_0^\tau (\varepsilon + \lambda N_s) ds} d\tau, \end{aligned}$$

where the final equality follows from integration by parts. Moreover, from the Markovian description of equilibrium in Equation (4), it is easy to see that this expression depends on λ and \bar{N}_0 only through $\Lambda_0 = \lambda \bar{N}_0$, so we can denote it by $A_t(\Lambda_0, B)$. Then we can assume without loss of generality that Λ_0 and $\hat{\Lambda}_0$ are of the form $\Lambda_0 = \lambda \bar{N}_0$ and $\hat{\Lambda}_0 = \hat{\lambda} \bar{N}_0$, i.e., that the two environments have the same underlying population size \bar{N}_0 .

Let $N_{t \geq 0}$ and $\hat{N}_{t \geq 0}$ be the equilibrium under λ and $\hat{\lambda}$, respectively. Given an arbitrary strictly positive adoption flow $M_{s \geq 0}$ and $t > 0$, note that the map

$$\lambda \mapsto \int_0^t M_\tau e^{-\int_0^\tau (\varepsilon + \lambda M_s) ds} d\tau$$

is strictly decreasing in λ . Note that since $\hat{\Lambda}_0 > \Lambda_0 > \bar{\Lambda}_0$, we have $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0) =: t_1^*$, and so we get that for all $t > 0$,

$$\int_0^t N_\tau e^{-\int_0^\tau (\varepsilon + \lambda N_s) ds} d\tau \geq \int_0^t N_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau, \quad (13)$$

with strict inequality for all $t > t_1^*$. We now show that

$$\int_0^t N_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau \geq \int_0^t \hat{N}_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} \hat{N}_s) ds} d\tau$$

which together with (13) implies the desired conclusion that $A_t(\hat{\lambda} \bar{N}_0, B) \leq A_t(\lambda \bar{N}_0, B)$ for all $t > 0$, with strict inequality for all $t > t_1^*$.

To prove this, suppose for a contradiction that there exists some $t > 0$ such that

$$\int_0^t N_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau < \int_0^t \hat{N}_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} \hat{N}_s) ds} d\tau. \quad (14)$$

Note that by the above result for *good* products, $\bar{N}_0 A_\tau(\lambda, G) = \int_0^\tau N_s ds \geq \int_0^\tau \hat{N}_s ds = \bar{N}_0 A_\tau(\hat{\lambda}, G)$ for all $\tau \geq 0$ and so

$$\int_0^t \varepsilon e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau \leq \int_0^t \varepsilon e^{-\int_0^\tau (\varepsilon + \hat{\lambda} \hat{N}_s) ds} d\tau \quad (15)$$

for all $t \geq 0$. Inequalities (14) and (15) together imply:

$$\int_0^t (\varepsilon + \hat{\lambda} N_\tau) e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau < \int_0^t (\varepsilon + \hat{\lambda} \hat{N}_\tau) e^{-\int_0^\tau (\varepsilon + \hat{\lambda} \hat{N}_s) ds} d\tau.$$

But this is equivalent to

$$\left(1 - e^{-\int_0^t (\varepsilon + \hat{\lambda} N_s) ds}\right) < \left(1 - e^{-\int_0^t (\varepsilon + \hat{\lambda} \hat{N}_s) ds}\right),$$

which contradicts $\int_0^t N_s ds \geq \int_0^t \hat{N}_s ds$. This completes the proof of part (iii) of Proposition 1. \blacksquare

B.4 Comparative Statics under Bad News without Partial Adoption

Suppose learning is via bad news. This section considers the effect of increased opportunities for social learning when there is no partial adoption. As in Lemma 2, impose Conditions 1 and 2, and

assume that $p_0 \in (0, p^\#)$. Then by Lemma 2, there is no partial adoption whenever $\Lambda_0 \leq \bar{\Lambda}_0 := \max\{\Lambda^*(p_0), \Lambda^*(\bar{p})\}$. The following result shows that, in contrast with Proposition 1, there is no saturation effect: Increased opportunities for social learning lead to strict welfare gains, speed up learning at all times, and reduce the adoption of bad products, while leaving the adoption of good products unchanged.

Proposition B.1. *Consider bad news learning and impose the same assumptions as in Lemma 2. Fix $\Lambda_0 < \hat{\Lambda}_0 \leq \bar{\Lambda}_0$. Then:*

(i). $W_0(\hat{\Lambda}_0) > W_0(\Lambda_0)$.

Moreover, if additionally $\hat{\Lambda}_0 < \Lambda^*(p_0)$,³³ then

(ii). $p_t^{\Lambda_0} < p_t^{\hat{\Lambda}_0}$ for all $t > 0$.

(iii). $A_t(\Lambda_0, G) = A_t(\hat{\Lambda}_0, G)$ and $A_t(\Lambda_0, B) > A_t(\hat{\Lambda}_0, B)$ for all $t > 0$.

B.4.1 Proof of part (i) of Proposition B.1

Denote $\Lambda_0^1 := \Lambda_0$ and $\Lambda_0^2 := \hat{\Lambda}_0$, with corresponding cutoff times t_1^i and t_2^i , value to waiting W_t^i , and no-news posteriors p_t^i for $i = 1, 2$. By Lemma 2, we have $t_1^i = t_2^i =: t^i$. Let $\hat{t} := \min\{t^1, t^2\}$. Then note that for all $t \leq \hat{t}$, $p_t^1 = p_t^2$ and $\Lambda_t^i = \Lambda_0^i$. By Lemma A.7 this implies that either $0 = t^1 = t^2$ or $t^1 < t^2$. If $0 = t^1 = t^2$, then for all $t > 0$, we have $2p_t^i - 1 > W_t^i$ and

$$p_t^i = \frac{p_0}{p_0 + (1 - p_0)e^{-(\varepsilon t + (1 - e^{-\rho t})\Lambda_0^i)}}.$$

Thus, $p_t^1 < p_t^2$ for all $t > 0$ which implies that $W_0^1 < W_0^2$.

If $t^1 < t^2$, then by definition of the cutoff times

$$W_{t^1}^2 > 2p_{t^1}^2 - 1 = 2p_{t^1}^1 - 1 \geq W_{t^1}^1.$$

Since there is no adoption until t^1 , we have

$$W_0^i = e^{-rt^1} \frac{p_{t^1}^i}{p_0} W_{t^1}^i,$$

which again implies that $W_0^1 < W_0^2$, as required. ■

³³If $\bar{\Lambda}_0 > \hat{\Lambda}_0 > \Lambda^*(p_0)$, then equilibrium dynamics under $\hat{\Lambda}_0$ start off in region (II) of Figure 2, i.e., initially all consumers delay and then all adopt upon their first opportunity from time $t_1^*(\hat{\Lambda}_0)$ on. In this case, $t_1^*(\hat{\Lambda}_0) > t_1^*(\Lambda_0)$, so that initially, learning and adoption are weakly slower under $\hat{\Lambda}_0$ than under Λ_0 (even though ex-ante welfare is higher under $\hat{\Lambda}_0$).

B.4.2 Proof of part (ii) of Proposition B.1

Suppose $\Lambda_0 < \hat{\Lambda}_0 < \Lambda^*(p_0)$. Then Lemma A.7 implies that $t_i^*(\Lambda_0) = t_i^*(\hat{\Lambda}_0) = 0$ for $i = 1, 2$. But then for all t ,

$$p_t^{\Lambda_0} = \frac{p_0}{p_0 + (1 - p_0)e^{-(\epsilon t + (1 - e^{-\rho t})\Lambda_0)}},$$

and similarly for $\hat{\Lambda}_0$, which implies $p_t^{\Lambda_0} < p_t^{\hat{\Lambda}_0}$. ■

B.4.3 Proof of part (iii) of Proposition B.1

Adoption of Good Products: Since $t_i^*(\Lambda_0) = t_i^*(\hat{\Lambda}_0) = 0$ for $i = 1, 2$, all consumers adopt upon their first opportunity at all times absent breakdowns. Thus, conditional on a good product, we get $A_t(\Lambda_0, G) = A_t(\hat{\Lambda}_0, G) = 1 - e^{-\rho t}$ for all t .

Adoption of Bad Products: Since $t_i^*(\Lambda_0) = t_i^*(\hat{\Lambda}_0) = 0$ for $i = 1, 2$, this follows from the same argument as in part (iii) of Proposition 1. ■

B.5 Comparative Statics under Good News

Suppose learning is via good news. Consider the effect of increasing the potential for social learning Λ_0 . Under good news, there is no partial adoption. Correspondingly, the following result shows that, in contrast with Proposition 1, there is no saturation effect.³⁴

Proposition B.1. *Consider learning via good news. Fix $r, \rho > 0$, $\varepsilon \geq 0$, and $p_0 \in (p^*, 1)$.³⁵ Suppose $\hat{\Lambda}_0 > \Lambda_0 \geq 0$. Then:*

(i). **Strict Welfare Gains:** *Provided $\varepsilon > 0$, we have $W_0(\hat{\Lambda}_0) > W_0(\Lambda_0)$.*³⁶

(ii). **Learning Speeds Up:**

- $0 < t^*(\hat{\Lambda}_0) < t^*(\Lambda_0)$
- $p_t^{\hat{\Lambda}_0} < p_t^{\Lambda_0}$ for all $t > 0$
- $p_{t^*(\hat{\Lambda}_0)+k}^{\hat{\Lambda}_0} = p_{t^*(\Lambda_0)+k}^{\Lambda_0}$ for all $k \geq 0$.

(iii). **No Initial Slow-Down in Adoption:**

- For all $t \leq t^*(\hat{\Lambda}_0)$, $A_t(\hat{\Lambda}_0; \theta) = A_t(\Lambda_0; \theta) = 1 - e^{-\rho t}$ for $\theta = B, G$.

³⁴Nevertheless, equilibrium behavior is not in general socially optimal, because p^* exceeds the socially optimal cutoff posterior. See Frick and Ishii (2014), sections 3.1 and 6.3.3.

³⁵Recall that $p^* := \frac{(\varepsilon+r)(r+\rho)}{2(\varepsilon+r)(r+\rho)-\varepsilon\rho}$ is the equilibrium cutoff posterior under good news. If $p_0 \leq p^*$, then all consumers rely entirely on the exogenous news source from the beginning, so the potential for social learning is irrelevant. If $\varepsilon = 0$, we assume that $p_0(1 + e^{-\Lambda_0}) < 1$ so that $t^*(\Lambda_0) < \infty$.

³⁶Increasing Λ_0 can increase welfare only if there are histories at which consumers' preference for adoption or delay is affected by information obtained via social learning. If $\varepsilon = 0$, then consumers are (weakly) willing to adopt at all histories, since the equilibrium posterior always remains weakly above $\frac{1}{2}$. Thus, in this case $W(\Lambda_0) = W(\hat{\Lambda}_0)$.

B.5.1 Proof of part (i) of Proposition B.1

If $p_0 > p^*$ and $\varepsilon > 0$, then under both Λ_0 and $\hat{\Lambda}_0$ consumers adopt immediately upon the first opportunity until p^* is reached and from then on delay adoption until there has been a breakthrough. Moreover, the probability π^* of a breakthrough occurring prior to p^* being reached is the same under both Λ_0 and $\hat{\Lambda}_0$: $\pi^* = \frac{p_0 - p^*}{1 - p^*}$. Because learning occurs at the same exogenous rate ε once p^* is reached, the continuation value W^* conditional on p^* being reached is also the same: $W^* = p^* \int_0^\infty \varepsilon e^{-(\varepsilon+r)t} \frac{\rho}{r+\rho} dt = 2p^* - 1$. So the only difference is that conditional on no breakthroughs, the time t^* at which p^* is reached occurs earlier under $\hat{\Lambda}_0$. To see that this is strictly beneficial, note that W_0 is composed of the following two terms:

$$W_0(\Lambda_0) = \left(1 - e^{-(r+\rho)t^*(\Lambda_0)}\right) \frac{\rho}{r+\rho} (2p_0 - 1) \\ + e^{-(r+\rho)t^*(\Lambda_0)} \left(\pi^* \frac{\rho}{r+\rho} + (1 - \pi^*) W^*\right),$$

and similarly for $\hat{\Lambda}_0$. The first term represents the case when a consumer receives an adoption opportunity prior to time t^* , and the second represents the case when a consumer's first adoption opportunity occurs after t^* . Conditional on either of these cases occurring, the expected payoff is the same under both Λ_0 and $\hat{\Lambda}_0$, but the time-discounted probability $e^{-(r+\rho)t^*}$ with which the second case occurs is strictly greater under $\hat{\Lambda}_0$. This is strictly beneficial, because the expected payoff in the second case is strictly greater:

$$\left(\pi^* \frac{\rho}{r+\rho} + (1 - \pi^*) (2p^* - 1)\right) - \frac{\rho}{r+\rho} (2p_0 - 1) \\ = \frac{r}{r+\rho} (1 - \pi^*) (2p^* - 1) > 0. \quad \blacksquare$$

B.5.2 Proof of part (ii) of Proposition B.1

If $p_0 > p^*$, then conditional on no breakthroughs, all consumers adopt immediately upon an opportunity until the time t^* at which the cutoff posterior p^* is reached. By Theorem 2, we have that for all $t < \min\{t^*(\hat{\Lambda}_0), t^*(\Lambda_0)\}$, $\lambda N_t = \rho e^{-\rho t} \Lambda_0 < \rho e^{-\rho t} \hat{\Lambda}_0 = \hat{\lambda} \hat{N}_t$. Since $p^* = \frac{(\varepsilon+r)(\rho+r)}{2(\varepsilon+\rho)(\varepsilon+r) - \varepsilon\rho}$ is independent of the potential for social learning, this implies that $t^*(\hat{\Lambda}_0) < t^*(\Lambda_0)$ and that $p_t^{\hat{\Lambda}_0} < p_t^{\Lambda_0}$ for all $t > 0$. Moreover, once the cutoff posterior is reached, information is generated at the constant exogenous rate ε , which means that conditional on $t > t^*$, beliefs depend only on $t - t^*$, as summarized in the third bullet point. \blacksquare

B.5.3 Proof of part (iii) of Proposition B.1

From Section B.5.2, $t^*(\hat{\Lambda}_0) < t^*(\Lambda_0)$. Thus, at all times $t \leq t^*(\hat{\Lambda}_0)$, all consumers adopt upon the first opportunity in both equilibria. \blacksquare

B.6 Proof of Proposition 2

We first establish the following lemma:

Lemma B.2. *Suppose $\bar{t} > t^* \geq 0$ and consider $f, g : [0, \infty) \rightarrow \mathbb{R}$ such that $f(\tau) = g(\tau)$ for all $\tau \leq t^*$, $f(\tau) < g(\tau)$ for $\tau \in (t^*, \bar{t})$, and $f(\tau) > g(\tau)$ for all $\tau > \bar{t}$. Suppose that $\int_0^\infty e^{-r\tau} f(\tau) d\tau = \int_0^\infty e^{-r\tau} g(\tau) d\tau$ for some $r > 0$. Then for all $\hat{r} > r$,*

$$\int_0^\infty e^{-\hat{r}\tau} f(\tau) d\tau < \int_0^\infty e^{-\hat{r}\tau} g(\tau) d\tau.$$

Proof. We have

$$\begin{aligned} 0 &= \int_0^\infty e^{-r\tau} (g(\tau) - f(\tau)) d\tau \\ &= \int_0^{\bar{t}} e^{-\hat{r}\tau} e^{(\hat{r}-r)\tau} (g(\tau) - f(\tau)) d\tau + \int_{\bar{t}}^\infty e^{-\hat{r}\tau} e^{(\hat{r}-r)\tau} (g(\tau) - f(\tau)) d\tau \\ &< e^{(\hat{r}-r)\bar{t}} \left(\int_0^{\bar{t}} e^{-\hat{r}\tau} (g(\tau) - f(\tau)) d\tau + \int_{\bar{t}}^\infty e^{-\hat{r}\tau} (g(\tau) - f(\tau)) d\tau \right) \\ &< e^{(\hat{r}-r)\bar{t}} \int_0^\infty e^{-\hat{r}\tau} (g(\tau) - f(\tau)) d\tau. \end{aligned}$$

This implies that $\int_0^\infty e^{-\hat{r}\tau} f(\tau) d\tau < \int_0^\infty e^{-\hat{r}\tau} g(\tau) d\tau$, as claimed. \blacksquare

To prove Proposition 2, fix $\varepsilon = 0$, $0 < r_p < r_i$, $\rho > 0$, $\bar{N}_0^p > 0$ and $p_0 \in (\frac{1}{2}, \frac{\rho+r_p}{\rho+2r_p})$. By Lemma 2, there exists $\bar{\lambda}^p$ such that the equilibrium with only patient types features initial partial adoption whenever $\lambda > \bar{\lambda}^p$. Consider $\hat{\lambda} > \lambda > \bar{\lambda}^p$. The following lemma derives the equilibrium of the game with a small enough mass of impatient types:

Lemma B.3. *There exists $\eta > 0$ such that whenever $\bar{N}_0^i < \eta$, the unique equilibrium for $\gamma \in \{\lambda, \hat{\lambda}\}$ takes the following form: There exists some $t^*(\gamma)$ such that the equilibrium flows N^i and N^p of impatient and patient adopters satisfy:*

$$\begin{aligned} N_t^i &= \rho \bar{N}_t^i \text{ for all } t, \\ N_t^p &= \begin{cases} \frac{r_p(2pt-1)}{\gamma(1-pt)} - \rho \bar{N}_t^i & \text{if } t < t^*(\gamma) \\ \rho \bar{N}_t^p & \text{if } t \geq t^*(\gamma). \end{cases} \end{aligned}$$

Proof. Fix $\gamma \in \{\lambda, \hat{\lambda}\}$. Pick $\eta > 0$ sufficiently small that $p_0 > \frac{\eta+r_p}{\eta+2r_p}$. Consider first the game consisting only of mass \bar{N}_0^p consumers of type r_p (and no consumers of type r_i). If there were an

exogenous news source in this game which generated signals at rate $\varepsilon \leq \eta$, then by Theorem 1 type r_p would always weakly prefer to adopt absent breakdowns. Then it is easy to see that in the game with no exogenous news source but with mass $\bar{N}_0^i < \eta$ of types r_i , type r_p will also always weakly prefer to adopt. This implies that type r_i must always *strictly* prefer to adopt.

Thus, $N_t^i = \rho \bar{N}_t^i$ for all t . Given this, the game reduces to one in which patient types view the information generated by the impatient types as a non-stationary exogenous news source which generates signals at rate $\varepsilon_t = \gamma \rho \bar{N}_t^i$. Modifying the arguments in the proof of Lemma 1, there must exist some $t^*(\gamma) > 0$ such that r_p is indifferent between adoption and delay for $t \leq t^*(\gamma)$, and r_p strictly prefers to adopt at all times $t > t^*(\gamma)$. Then the unique equilibrium can be derived in the same manner as in the proof of Theorem 1. ■

Given Lemma B.3, we can follow the arguments in the proof of Proposition 1 to show that $t^*(\lambda) < t^*(\hat{\lambda})$ and that there exists some $\bar{t} > t^*(\lambda)$ such that

$$p_t^\lambda \begin{cases} = p_t^{\hat{\lambda}} & \text{if } t \leq t^*(\lambda) \\ > p_t^{\hat{\lambda}} & \text{if } t \in (t^*(\lambda), \bar{t}) \\ < p_t^{\hat{\lambda}} & \text{if } t > \bar{t}. \end{cases}$$

Note that the ex ante expected payoff of type r_k ($k \in \{p, i\}$) under arrival rate $\gamma \in \{\lambda, \hat{\lambda}\}$ is

$$W_0^k(\gamma) = \int_0^\infty \rho e^{-(r_k + \rho)\tau} \frac{p_0}{p_\tau^\gamma} (2p_\tau^\gamma - 1) d\tau.$$

Since r_p is initially indifferent between adoption and delay under both λ and $\hat{\lambda}$, we have $W_0^p(\lambda) = W_0^p(\hat{\lambda}) = 2p_0 - 1$. Thus, Lemma B.2 yields $W_0^i(\lambda) > W_0^i(\hat{\lambda})$. This completes the proof of Proposition 2. ■