# INNOVATION ADOPTION BY FORWARD-LOOKING SOCIAL LEARNERS

**By**

**Mira Frick and Yuhta Ishii**

**February 2015**

**COWLES FOUNDATION DISCUSSION PAPER NO. 1988**

**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS**
**YALE UNIVERSITY**
**Box 208281**
**New Haven, Connecticut 06520-8281**

**http://cowles.econ.yale.edu/**

# Innovation Adoption by Forward-Looking Social Learners

Mira Frick Yuhta Ishii*

December 20, 2014

**Abstract**

Motivated by the rise of social media, 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), equilibrium dynamics depend non-trivially on qualitative and quantitative features of the informational environment. We identify informational 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. We also suggest a novel, purely informational explanation for different commonly observed adoption curves (S-shaped vs. concave). (*JEL* D81, D83, O33)

---

*Frick: Department of Economics, Harvard University and HBS (mfrick@fas.harvard.edu); Ishii: Cowles Foundation for Research in Economics, Yale University and Centro de Investigación Económica, ITAM (yuhta.ishii@yale.edu). We are extremely grateful to Attila Ambrus, Drew Fudenberg, and Eric Maskin for generous advice and encouragement. For helpful comments, we also thank Dirk Bergemann, Aislinn Bohren, Yeon-Koo Che, Thomas Covert, Martin Cripps, Glenn Ellison, Daniel Garrett, Ben Golub, Brett Green, Marina Halac, Johannes Hörner, Boyan Jovanovic, Scott Kominers, David Laibson, Greg Lewis, César Martinelli, Stephen Morris, Aniko Öry, Sven Rady, Andrew Rhodes, Larry Samuelson, Heather Schofield, Jesse Shapiro, Andy Skrzypacz, Tomasz Strzalecki, Jean Tirole, and seminar audiences at Harvard, Chicago Booth, ITAM, Kellogg, Kansas Workshop on Economic Theory, NASM 2014 (Minneapolis), and EEA/ESEM 2014 (Toulouse).

# 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 (such as Yelp, Amazon, Google, YouTube, and Facebook) has greatly 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 people than ever 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. Our key contribution is a careful analysis of consumers' informational incentives and their dependence on quantitative and qualitative features of the news environment through which social learning occurs. 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 news environments that are subject to a novel *saturation effect*, whereby beyond a certain level, increased opportunities for social learning can slow down adoption and learning and do not increase consumer welfare (possibly even being harmful). Second, at a qualitative level, we show that different news environments give rise to observable differences in aggregate adoption dynamics: This implies a new, *purely informational* explanation for two of the most commonly observed adoption patterns (S-shaped vs. concave curves), which we support with some suggestive evidence.

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 fundamental 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 innovations.[1] However, its informational ramifications have largely remained unexplored theoretically: Existing learning-based models of innovation adoption typically assume either that learning is social but consumers are myopic (e.g. Ellison and Fudenberg, 1993; Young, 2009), or that consumers are forward-looking but information arrives purely exogenously (e.g. Jensen, 1982). In either case, the dependence on the informational environment is trivial, both quantitatively (a greater ease of information transmission is always beneficial) and qualitatively (absent other forces such as consumer heterogeneity, different news environments alone do not give rise to interestingly different adoption dynamics[2]).

**Summary of Model and Results:** 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

[1] Studies of social learning in this domain include Besley and Case (1993, 1994); Foster and Rosenzweig (1995); Conley and Udry (2010). 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.

[2] See the discussion in Footnote 12 under Related Literature.

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 about the innovation 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 employ a variation of the Poisson models of strategic experimentation pioneered by Keller et al. (2005); Keller and Rady (2010, 2014).[3] Individual adopters' experiences generate public signals at a fixed Poisson rate which we use to quantify the potential for social learning. Qualitatively, there is a natural distinction (see also MacLeod, 2007; Board and Meyer-ter Vehn, 2013; Che and Hörner, 2014) 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.

For examples of innovations featuring learning via bad news (or the absence thereof), recall the extensive social media coverage of of a battery fire in a Tesla Model S electric car in October 2013, or the gradual increase of consumers' confidence in microwave ovens in the 1970s (following widespread initial concerns over possible "radiation leaks") or in risky new medical procedures such as gastric bypass surgery.[4] By contrast, learning via good news events (or their absence) is common in award-focused industries (e.g. movies or books); or for (essentially side-effect free) herbal remedies, beauty or fitness products. The news environment may also be determined by limitations or usage practices of the available social learning systems, e.g. the fact that Facebook allows users to "Like" a product's site, but has no "Dislike" but-

[3]Other papers using this "exponential bandits" framework include Bergemann and Hege (1998, 2005); Strulovici (2010); Bonatti and Hörner (2011); Klein and Rady (2011); Hörner and Samuelson (2013); Halac et al. (2013, 2014); Halac and Prat (2014).

[4]The latter two examples are discussed in more detail in Section 4.1.

ton; or that the overwhelming majority of book reviews on Amazon.com and BarnesandNoble.com appear to be positive.[5]

Section 3 analyzes and contrasts equilibrium adoption behavior in bad and good news markets. For tractability, we focus on *perfect* bad (respectively good) news environments, in which a *single* signal arrival *conclusively* indicates bad (respectively good) quality, so that equilibrium dynamics are non-trivial only in the absence of signals. A key insight facilitating our analysis is that consumers' equilibrium incentives across time must satisfy a quasi-single crossing property (Section 3.1): Absent signals, 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. This enables us to establish the existence of unique[6] equilibria. Equilibrium adoption dynamics admit simple closed-form descriptions, which are Markovian in current beliefs and in the mass of consumers who have not yet adopted.

Under perfect bad news (Section 3.2), the unique equilibrium is characterized by two times $0 \le t_1^* \le t_2^*$, which depend on the fundamentals: Until time $t_1^*$, no adoption takes place and consumers acquire information only from exogenous sources; from time $t_2^*$ on, all consumers adopt immediately when given a chance, unless a breakdown occurs, in which case adoption comes to a permanent standstill. If $t_1^* < t_2^*$, then throughout $(t_1^*, t_2^*)$ there is inefficiency in the form of *partial adoption*: Only *some* consumers adopt when given a chance, with others free-riding on the information generated by the adopters. The flow of new adopters on $(t_1^*, t_2^*)$ is uniquely determined by an ODE that guarantees consumers' indifference between adopting and delaying throughout this interval. Given that consumers are forward-looking, $t_1^* < t_2^*$ occurs in economies with a sufficiently large potential for social learning and not too optimistic consumers (on the other hand, if consumers are myopic or if there are no possibilities for social learning, then necessarily $t_1^* = t_2^*$).

By contrast, the perfect good news equilibrium (Section 3.3) features adoption up to some time $t^*$ (which depends on the fundamentals) and no adoption

[5] Cf. Chevalier and Mayzlin (2006), which we discuss in greater detail in Section 4.1.

[6] Uniqueness is in terms of *aggregate* adoption behavior.

from $t^*$ on (unless there is a breakthrough, after which all consumers adopt upon their first opportunity). The key difference with perfect bad news is that equilibrium adoption behavior is *all-or-nothing*: Regardless of the potential for social learning, there are no periods during which only some consumers adopt when given a chance. This highlights a fundamental way in which the nature of the news environment affects consumers' adoption incentives. During a period of time when, absent signals, a consumer is prepared to adopt the innovation, he will be willing to delay his decision only if he expects to acquire *decision-relevant information* in the meantime: Since originally he is prepared to adopt the innovation, such information must make him strictly prefer *not* to adopt. When learning is via bad news, breakdowns have this effect, since they reveal the innovation to be bad. By contrast, under perfect 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 shows that bad news and good news environments give rise to observably different adoption patterns. Under perfect bad news, adoption curves (which plot the percentage of adopters in the population against time) are S-shaped: Up to time $t_1^*$ adoption is flat, on $(t_1^*, t_2^*)$ adoption levels increase convexly (absent breakdowns), and from time $t_2^*$ there is a concave increase. Convex growth throughout $(t_1^*, t_2^*)$ is tied to consumer indifference during this region: As consumers grow increasingly optimistic absent breakdowns, their opportunity cost to delaying goes up. To maintain indifference, this increase is offset by an increase in the flow of new adopters, which raises the odds that waiting will produce information allowing consumers to avoid a bad innovation. By contrast, adoption under perfect good news occurs in concave "bursts": Up to time $t^*$ adoption levels increase concavely, then adoption flattens out, possibly followed by another region of concave growth if a breakthrough occurs. S-shaped and concave curves are arguably the two most widely documented empirical adoption patterns, with the typical marketing textbook devoting a chapter to this "stylized fact".[7] But as we discuss below under Related Literature, our model appears

[7] Cf. Hoyer et al. (2012), Ch. 15, p. 425ff. and Keillor (2007) p. 46–61. The former type

to be the first to point to different market learning environments as a possible source. Focusing on the aforementioned examples of good and bad news markets, we present some suggestive evidence for our predictions, pointing to an opportunity for more systematic empirical work.

Section 4.2 establishes the possibility of a saturation effect: If learning is via perfect bad news and the potential for social learning is great enough that $t_1^* < t_2^*$, then holding fixed other fundamentals, any additional increase in opportunities for social learning has no impact at all on (ex ante) equilibrium welfare levels. This is because any benefits from increasing the potential for social learning are balanced out by an expansion of the period $(t_1^*, t_2^*)$ of informational free-riding. As a result, greater opportunities for social learning 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 non-monotonicity in the speed of learning to construct an example with heterogeneous consumers, where increased opportunities for social learning are not only not beneficial, but in fact give rise to *Pareto-decreases* in ex ante welfare. By contrast, under perfect good news, increasing the potential for social learning is (essentially) always strictly beneficial and speeds up learning at all times.

**Related Literature:** We contribute to a large literature (spanning economics, marketing, and sociology)[8] that seeks to explain why the adoption of innovations is typically a drawn-out process and why different innovations follow different characteristic adoption patterns, notably the widely-documented S-shaped and concave adoption curves.[9] First, we identify a novel, *purely informational* source of these regularities: Forward-looking social learners may delay adoption to gather information about others' experiences, but delay incentives (and hence adoption patterns) are sensitive to the market learning en-

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).

[8]See Geroski (2000) and Baptista (1999) for more comprehensive surveys.

[9]See footnote 7.

vironment.[10] Existing models appear to have overlooked this channel, appealing instead to (a combination of): (i) an assumed heterogeneity of potential adopters, with a distribution of characteristics that is imposed exogenously to fit the desired adoption pattern—as in "probit" models[11] or existing learning-based models;[12] (ii) non-informational "spillover" effects which, *independently of the quality* of the innovation, increase current adoption as a function of past adoption—e.g. by a process of contagion as in "epidemic" models,[13] 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 we abstract away from (i)–(iii), but we do not wish to deny that a combination of these factors is likely often at play as well. Second, however, we investigate the effect of increased opportunities for social learning and obtain predictions (notably the saturation effect) that are outside the scope of existing models.

Our Poisson learning framework borrows from the strategic experimentation literature (Keller et al., 2005; Keller and Rady, 2010, 2014),[14] but we depart in two key respects: First, since our focus is on large market applications, we assume that any individual's influence on public aggregate information is negligible. Second, we assume that adoption is irreversible rather than allowing for continuous back-and-forth switching; this is natural for innovations such as

[10] This message is similar in spirit to Board and Meyer-ter Vehn (2013), who highlight the role of the market learning process in a different setting, viz. a capital-theoretic model of firms' incentives to invest in quality and reputation.

[11] E.g. David (1969); Davies (1979); Karshenas and Stoneman (1993).

[12] E.g. Jensen (1982), where players are forward-looking but information arrives purely exogenously, obtains S-shaped adoption curves by assuming that players' initial beliefs about quality are uniformly distributed over some interval. In his model, and also if learning is social but consumers are myopic (as in Young, 2009), a population of *identical* consumers would follow a cutoff rule, with everyone adopting the innovation at beliefs above a certain threshold and not adopting otherwise, which rules out convex growth in adoption levels regardless of the news environment.

[13] E.g. Mansfield (1961, 1968); Bass (1969, 1980); Mahajan and Peterson (1985); Mahajan et al. (1990).

[14] These papers feature learning via perfect good news, imperfect good news, and perfect and imperfect bad news, respectively. Bolton and Harris (1999), the founding paper of this literature, has learning based on Brownian motion.

medical procedures or movies, for which "consumption" is usually a one-time event, or for technologies with large switching costs. An important theoretical implication is the absence from our model of the *encouragement effect*, which is central to the strategic experimentation literature.[15] This makes our analysis more tractable—e.g., in contrast with the aforementioned papers, our equilibria are unique (at the aggregate level).[16] More substantively, we obtain differences between bad and good news environments that do not arise under strategic experimentation, as well as novel comparative statics with respect to the potential for social learning.[17]

In independent and contemporaneous work, Che and Hörner (2014) employ a similar variation of Keller et al. (2005) to model learning about a new product by a large population of consumers. However, they perform a *normative* analysis: Signals about past adopters' experiences are only visible to a benevolent mediator, who based on his information makes adoption recommendations that maximize social welfare subject to a credibility constraint. To counterbalance informational free-riding, the optimal mechanism under both good and bad news generally features regions of selective over-recommendation.[18]

Finally, informational externalities in social learning are also studied by the observational learning literature (e.g. Banerjee, 1992; Bikhchandani et al.,

[15] According to this effect, individuals have an incentive to increase current experimentation to drive up beliefs and induce more future experimentation by others; it requires crucially that (i) individuals have a direct influence on opponents' information and (ii) they can adjust experimentation as a function of beliefs. There is no encouragement effect in Keller et al. (2005), but again (i) and (ii) are crucial in generating asymmetric switching equilibria, in which players take turns experimenting at different beliefs.

[16] Moreover, we do not need to restrict to Markovian strategies.

[17] Specifically, under *both* perfect good and perfect bad news (resp. Keller et al., 2005; Keller and Rady, 2014), the unique symmetric MPE features mixing throughout an intermediate region of beliefs, whereas in our model partial adoption arises only under perfect bad news. Also, in both Bolton and Harris (1999) and Keller and Rady (2014), an increase in the number of players or signal informativeness makes players willing to experiment at more pessimistic beliefs, whereas we obtain the opposite result under bad news.

[18] This is true when consumers are myopic, which is Che and Hörner's main focus. In section 5, they also consider a version of forward-looking consumers, but this is quite different from our model, because consumers are restricted to choosing a *single* time at which to "check-in" with the mediator and are not able to observe any information prior to this time. Under perfect good news (they do not consider perfect bad news), they show that the optimal policy in this case is sometimes fully transparent.

1992; Smith and Sørensen, 2000, where the timing of players' moves is exogenous; and Chamley and Gale (1994); Rosenberg et al. (2007); Murto and Välimäki (2011), which like our paper feature endogenous timing). The key difference is that in this literature players hold *private* information about a payoff-relevant state variable and make inferences by observing others' *actions*, whereas all relevant news in our model is *public* and derived from previous adopters' *experiences* (which better captures social learning via centralized internet-based review platforms). Given the assumption of private information, a particular focus of this literature is on the possibility of herding and informational cascades. By contrast, none of the cited papers derive adoption curves or study the way in which they are shaped by qualitative features of the news environment.[19]

# 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 according to a Poisson process with exogenous arrival rate $\rho > 0$.[20] 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

[19] At a *quantitative* level, our saturation effect is somewhat reminiscent of Chamley and Gale's (1994) result that as the number of players increases, the rate of investment and the information flow are eventually independent of the number of players.

[20] Stochasticity of adoption opportunities can be seen as capturing the natural assumption that consumers face cognitive and time constraints, making it impossible for them to ponder the decision whether or not to adopt the innovation at every instant in continuous time.

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 of Keller et al. (2005) and Keller and Rady (2010, 2014): Let $N_t$ denote the flow of of consumers newly adopting the innovation at time $t$, which we define more precisely in Section 2.3. Then, conditional on quality $\theta$, 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 $\theta$ of the innovation. The signal process summarizes news events that are generated from two sources:

First, the *social learning* term $\lambda N_t$ 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 $\lambda\, dt$.[21] 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 $\varepsilon\, dt$, representing information generated independently of consumers' behavior, e.g. by professional critics or government watchdog agencies.

[21] By letting the social learning component of the signal arrival rate at time $t$, $\lambda N_t$, depend only on the flow of adopters $N_t$ at time $t$ itself, we are effectively assuming that each adopter can generate a signal only once, namely at the time of adoption. This is appropriate for innovations such as new movies or medical procedures, for which "consumption" is a one-time event and quality is revealed upon consumption. For durable goods (e.g. cars or consumer electronics), it might be more natural to allow adopters to generate signals repeatedly over time, which can be captured by replacing $\lambda N_t$ with $\lambda \int_0^t N_s\, ds$. This would yield results that are qualitatively similar to those presented in the following sections.

For tractability, we focus on learning via *perfect* Poisson processes, where a single signal provides *conclusive* evidence of the quality of the innovation. Qualitatively, we can then distinguish between learning via **perfect bad news**, where $\varepsilon_G = \lambda_G = 0$ and $\varepsilon_B = \varepsilon \geq 0$, $\lambda_B = \lambda > 0$, so that the arrival of a signal (called a *breakdown*) is conclusive evidence that the innovation is bad; and learning via **perfect good news**, where $\varepsilon_B = \lambda_B = 0$ and $\varepsilon_G = \varepsilon \geq 0$, $\lambda_G = \lambda > 0$, so that a signal arrival (called a *breakthrough*) is conclusive evidence that the innovation is good. As motivated in the Introduction and Section 4.1, the distinction between bad news and good news can be seen to reflect the nature of news production in different markets.

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 $\lambda$ 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 perfect bad news, consumers' posterior on $\theta = G$ permanently jumps to 0 at the first breakdown, while under perfect 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 $N_{s\geq 0}$, standard Bayesian updating implies that

$$p_t = \frac{p_0 e^{-\int_0^t (\varepsilon_G + \lambda_G N_s) ds}}{p_0 e^{-\int_0^t (\varepsilon_G + \lambda_G N_s) ds} + (1 - p_0) e^{-\int_0^t (\varepsilon_B + \lambda_B N_s) ds}}.\text{[22]} \tag{1}$$

In particular, if $N_\tau$ is continuous in an open interval $(s, s + \nu)$ for $\nu > 0$, then $p_\tau$ 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 perfect bad news and decreasing under perfect good news.

[22] Definition 2.1 imposes measurability on $N$, so this expression is well-defined.

## 2.3 Equilibrium

Since our main 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, consumers' incentives are non-trivial only in the absence of signals: Under perfect bad news, no new consumers adopt after a breakdown, while under perfect good news all remaining consumers adopt at their first opportunity after there has been a breakthrough. Therefore, 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. Reflecting the assumption that aggregate adoption behavior is predictable with respect to the news process of the economy, we require that $N_t$ be a *deterministic* function of time. We consider all such functions which are feasible in the following sense:

**Definition 2.1.** A feasible flow of adopters is a right-continuous function $N\colon [0, +\infty) \to \mathbb{R}$ 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 $\rho$. Any feasible adoption flow $N_{t\geq 0}$ defines an associated no-news posterior $p_t^N$ as given by Equation (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 that he 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 $\Sigma_t$ denote the set of all right-continuous functions $\sigma : [t, +\infty) \to \{0, 1\}$, each of which defines a potential set of future times at which, absent signals, a given consumer might adopt if given an opportunity. Under the Poisson process generating adoption opportunities, any $\sigma \in \Sigma_t$ defines a random time $\tau^\sigma$ at which, absent signals, the consumer will adopt the innovation and drop out of the game.[23]

Let $W_t^N(\sigma)$ denote the expected payoff to waiting at $t$ and following $\sigma$ in the future, given the aggregate adoption flow $N_{s\geq 0}$. Specifically, if learning is via perfect bad news, $\sigma$ prescribes adoption at the random time $\tau^\sigma$ if and only if there have been no breakdowns prior to $\tau^\sigma$, 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.

If learning is via perfect good news, then following $\sigma$ means that at any adoption opportunity prior to $\tau^\sigma$, adoption occurs only if there has been a breakthrough, and at $\tau^\sigma$ adoption occurs whether or not there has been a breakthrough. For any time $s \geq t$, denote by $\tau_s$ the random time at which the first adoption opportunity after $s$ arrives. Then $W_t^N(\sigma)$ is given by

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

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 2.2.** The value to waiting given a feasible adoption flow $N_{t\geq 0}$ is

[23]Formally, 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 $\rho$, 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 $\inf \varnothing := +\infty$. 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.

the function $W^N : \mathbb{R}_+ \to \mathbb{R}_+$ 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 2.3.** 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.[24]

# 3 Equilibrium Analysis

## 3.1 Quasi-Single Crossing Property for Equilibrium Incentives

We now proceed to equilibrium analysis. As a preliminary step, we first establish a useful property of equilibrium incentives under both perfect bad news and perfect 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 $W_{t\geq 0}^N$ as defined in Definition 2.2. In general, the dynamics of the trade-off between immediate adoption at 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

[24] Note that Definition 2.3 is essentially Nash equilibrium, i.e. we do 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.

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 $t$,

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

and this imposes considerable discipline on the dynamics of the trade-off. Indeed, the following theorem establishes that $2p_t^N - 1$ and $W_t^N$ must satisfy a quasi-single crossing property:

**Theorem 3.1.** *Suppose that learning is either via perfect bad news ($\lambda_B > 0 = \lambda_G$) or via perfect 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 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. We briefly illustrate the intuition for the first bullet point when learning is via perfect bad news. Suppose that immediate adoption is strictly better than waiting today (and hence also in the near future provided there are no breakdowns).[25] Then 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 be decreasing over time. As a result, immediate adoption becomes even more attractive relative to waiting, and consequently immediate adoption continues to be strictly preferable at all times in the future.

[25] This follows from the continuity of the equilibrium value to waiting, which is established in Lemma A.1 in the Online Appendix.

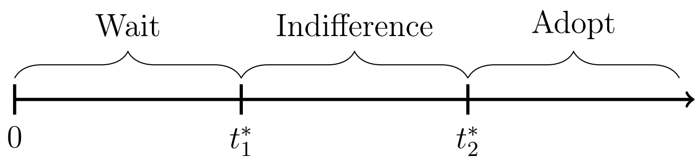


Figure 1: Perfect Bad News

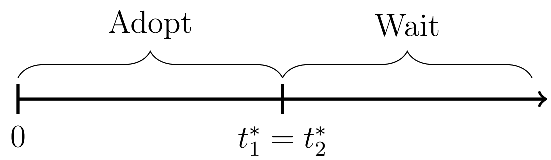


Figure 2: Perfect Good News

With any equilibrium $N_{t\geq 0}$, we associate two cutoff times $0 \leq t_1^* \leq t_2^* \leq +\infty$:[26] If learning is via perfect bad news, set

$$t_1^* := \inf\{t \geq 0 : N_t > 0\} \text{ and } t_2^* := \sup\{t \geq 0 : N_t < \rho \bar{N}_t\};^{27} \tag{2}$$

if learning is via perfect good news, set

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

Thus, if $N_{t\geq 0}$ is a perfect 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; while under perfect good news $N_{t\geq 0}$ features immediate adoption prior to $t_1^*$ and no adoption after $t_2^*$ absent breakthroughs. Moreover, under both perfect bad and good news, Theorem 3.1 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.[28] This is illustrated in Figures 1 and 2. In Sections 3.2 and 3.3 we will build

[26] With the convention that $\inf \varnothing = +\infty$ and $\sup \varnothing = 0$.

[27] Recall that $\bar{N}_t := \bar{N}_0 - \int_0^t N_s \, ds > 0$ denotes the remaining population at time $t$.

[28] Suppose learning is via perfect good news. Consider $t \in (t_1^*, t_2^*)$. 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 by Theorem 3.1 $2p_t - 1 = W_t$. The argument for perfect bad news is analogous.

on this observation to establish the existence of unique equilibria under both perfect bad and good news. The cutoff times, as well as the flow of adopters between $t_1^*$ and $t_2^*$, are fully pinned down by the parameters. Looking ahead to Section 3.3, we will see that the perfect good news equilibrium satisfies $t_1^* = t_2^* = t^*$; thus, adoption behavior is **all-or-nothing**, with all consumers adopting upon first opportunity up to time $t^*$ and adoption ceasing from then on absent breakthroughs. By contrast, for suitable parameters the perfect bad news equilibrium in Section 3.2 features a non-empty region $(t_1^*, t_2^*)$. 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 (i.e. $N_t \in (0, \rho\bar{N}_t)$ at each $t \in (t_1^*, t_2^*)$). 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 Equilibrium under Perfect Bad News

Assume that learning is via perfect bad news. The following theorem builds on the analysis of the previous section to establish the existence of an equilibrium $N_{t\geq 0}$, which is uniquely pinned down by the parameters. At all $t$, $N_t$ is Markovian in the associated no-news posterior $p_t$ and the time-$t$ potential for social learning $\Lambda_t := \lambda\bar{N}_t$:[29]

**Theorem 3.2** (**Equilibrium under PBN**). *Fix $r, \rho, \lambda, \bar{N}_0 > 0$, $\varepsilon \geq 0$, and $p_0 \in (0,1)$. There exists a unique equilibrium. Furthermore, in the unique equilibrium, $N_t$ is Markovian in $(p_t, \Lambda_t)$ for all $t$: There exists a non-decreasing function $\Lambda^* : [0,1] \to \mathbb{R} \cup \{\infty\}$ and some $p^* \in [\frac{1}{2}, 1)$ such that*

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

[29] Recall that $\bar{N}_t := \bar{N}_0 - \int_0^t N_s\,ds$ denotes the remaining population at time $t$.

The proof of Theorem 3.2 is in Online Appendix B.2. Here we sketch the basic argument. Fix parameters $r$, $\rho$, $\bar{N}_0 > 0$, $\varepsilon$, $\lambda \geq 0$, and $p_0 \in (0,1)$, and suppose that $N_{t\geq 0}$ is an equilibrium. By the previous section, Equation (2) defines cutoff times $0 \leq t_1^* \leq t_2^* \leq +\infty$ such that $N_t = 0$ if $t < t_1^*$, $N_t = \rho \bar{N}_t$ if $t > t_2^*$, and at all $t \in [t_1^*, t_2^*)$, consumers are indifferent between adopting immediately and waiting for more information.

**Partial adoption during** $(t_1^*, t_2^*)$**:** Lemma B.1 in Online Appendix B.2 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:

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

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 $(2p_{t+dt}-1)rdt$, reflecting the time cost of delayed adoption.[30] Ignoring terms of order $dt^2$ and rearranging yields $N_t = \frac{r(2p_t-1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda}$.[31]

**Determining the cutoff times:** Next, we derive an alternative descrip-

[30]Note that $\rho$ does not enter into this expression, because in the indifference region consumers obtain the same continuation payoff regardless of whether or not they obtain an adoption opportunity in the time interval $(t, t+dt)$ and hence are indifferent between receiving an opportunity to adopt or not.

[31]A bit more precisely, ignoring terms of order $dt^2$, the right hand side of Equation 5 is given by $(1-(\varepsilon+\lambda N_t)\,(1-p_t)dt)(2(p_t+\dot{p}_t dt)-1)rdt = r(2p_t-1)dt$. Further rearrangement yields the desired expression.

tion of $t_1^*$ and $t_2^*$ in terms of the evolution of the no-news posterior $p_t$ and the potential for social learning $\Lambda_t$. To state this description, we define the following notation. For any $p \in (0,1)$ and $\Lambda \geq 0$, let

$$G(p,\Lambda) := \int_0^\infty \rho e^{-(r+\rho)\tau} \left(p - (1-p)e^{-\left(\varepsilon\tau + \Lambda\left(1-e^{-\rho\tau}\right)\right)}\right) d\tau.$$

$G(p,\Lambda)$ represents the payoff to adopting at the next opportunity absent breakdowns, *given* that the current belief is $p$, that the remaining potential for social learning is $\Lambda$, and that absent breakdowns the remaining $\Lambda/\lambda$ consumers adopt at their first opportunity in the future.

Define cutoff posteriors $\underline{p}$, $\overline{p}$, and $p^\sharp$ as follows. Let $\underline{p}$ be the lowest posterior at which a consumer to whom adoption opportunities arrive at rate $\rho$ is willing to adopt immediately if all information in the future arrives exclusively through the *exogenous* new source; that is,

$$2\underline{p} - 1 = G(\underline{p}, 0) \Leftrightarrow \underline{p} := \frac{(\varepsilon + r)(r+\rho)}{2(\varepsilon + r)(r+\rho) - \varepsilon\rho}.$$

Define $\overline{p} := \lim_{\rho\to\infty} \underline{p} = \frac{\varepsilon+r}{\varepsilon+2r}$ to be the lowest belief at which a hypothetical consumer to whom adoption opportunities arrive *continuously* would be willing to adopt immediately if all information in the future arrives exclusively through the exogenous new source. Define $p^\sharp := \lim_{\varepsilon\to\infty} \underline{p} = \frac{\rho+r}{\rho+2r}$ to be the lowest belief at which a consumer to whom adoption opportunities arrive at rate $\rho$ would be willing to adopt immediately even if all uncertainty were to be *completely resolved* by the next adoption opportunity.[32]

Finally, define the function $\Lambda^* : [0,1] \to \mathbb{R}_+ \cup \{+\infty\}$ as follows. Let $\Lambda^*(p) = 0$ for all $p \leq \underline{p}$, $\Lambda^*(p) = +\infty$ for all $p \geq p^\sharp$, and for all $p \in (\underline{p}, p^\sharp)$, let $\Lambda^*(p) \in \mathbb{R}_+$ be the unique value such that $2p - 1 = G(p, \Lambda^*(p))$.[33] Thus, if the current posterior is $p \in [\underline{p}, p^\sharp)$ and the current potential for social learning in

[32] Thus, for all $p > p^\sharp$, $\lim_{\Lambda\to\infty} G(p,\Lambda) < 2p - 1$ and for all $p < p^\sharp$, $\lim_{\Lambda\to\infty} G(p,\Lambda) > 2p - 1$.

[33] Note that such a value must exist given that $p \in (\underline{p}, p^\sharp)$ and is unique because $G(p,\Lambda) - (2p-1)$ is strictly decreasing in $p$ and strictly increasing in $\Lambda$ on this domain.

the economy is $\Lambda^*(p)$, then consumers are indifferent between adopting now or at their next opportunity absent breakdowns, provided that all remaining $\Lambda^*(p)/\lambda$ consumers also adopt at their first opportunity in the future.

Then, letting $p^* := \min\{\overline{p}, p^\sharp\}$, Lemma B.3 in Online Appendix B.2 shows that $t_2^* = \inf\{t \geq 0 : \Lambda_t < \Lambda^*(p_t)\}$ and $t_1^* = \min\{t_2^*, \sup\{t \geq 0 : p_t < p^*\}\}$.[34]

**Equilibrium dynamics given initial parameters:** From the previous two steps, it is clear that any equilibrium must take the Markovian form in Equation (4), with $\Lambda^*$ and $p^*$ as defined above. It remains to show how Equation (4) uniquely pins down the evolution of $N_t$ as a function of the initial parameters; and to verify that $N_{t\geq 0}$ thus obtained does indeed constitute an equilibrium (in particular, is feasible). Here we sketch the former argument, relegating the latter to Online Appendix B.2.3. Note first the following two special cases: If $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$, then Equation (4) implies that $N_t = 0$ for all $t$. Second, if $\varepsilon \geq \rho$ (so that $p^* := \min\{\overline{p}, p^\sharp\} = p^\sharp$), then because $\Lambda^*(p) = +\infty$ for all $p \geq p^\sharp$, $N_t = 0$ as long as $\Lambda_t > \Lambda^*(p_t)$ and $N_t = \rho \bar{N}_t$ as soon as $\Lambda_t \leq \Lambda^*(p_t)$. Throughout the rest of the paper, we will be particularly interested in equilibria that feature a non-empty partial adoption region $(t_1^*, t_2^*)$. Since the two cases above preclude this regardless of other parameters, we henceforth rule them out (Online Appendix B.2.4 discusses the second case in more detail):

**Condition 3.3.** The rate at which exogenous information arrives is smaller than the rate at which consumers obtain adoption opportunities: $\varepsilon < \rho$.

**Condition 3.4.** Either $\varepsilon > 0$ or $p_0 \in (\frac{1}{2}, 1)$.

Given these conditions, Figure 3 illustrates how the unique equilibrium is obtained as a function of the parameters. Regions (2) and (3) represent values of $(p_t, \Lambda_t)$ corresponding to the first line of Equation (4), so that no adoption takes place in these regions. Region (4) corresponds to partial adoption as given by the second line of Equation (4). Finally, region (1) corresponds to the third line of Equation (4) and thus to immediate adoption.

[34] We impose the convention that if $\{t \geq 0 : p_t < p^* = \frac{1}{2}\} = \varnothing$, then $\sup\{t \geq 0 : p_t < p^* = \frac{1}{2}\} := 0$.

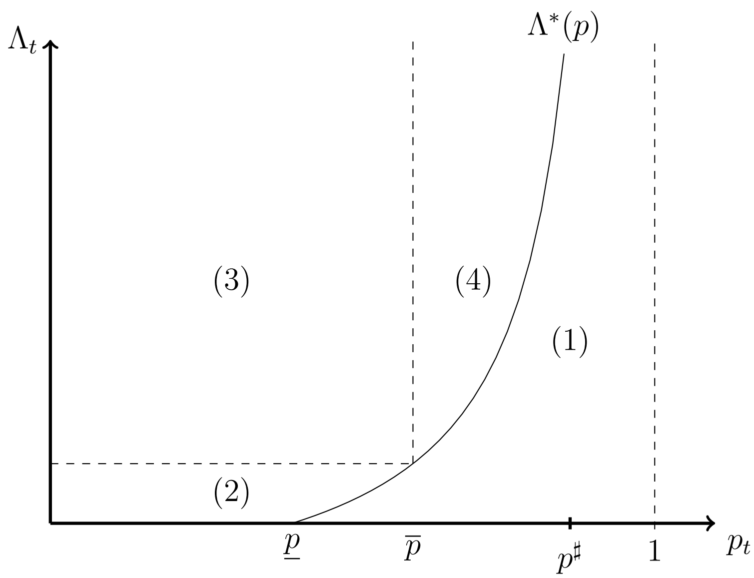


Figure 3: Partition of $(p_t, \Lambda_t)$ when $\varepsilon < \rho$

If $(p_0, \Lambda_0)$ is in region (2), then initially no adoption occurs and the no-news posterior drifts upward according to the law of motion $\dot{p}_t = p_t(1 - p_t)\varepsilon$, while $\Lambda_t$ remains unchanged at $\Lambda_0$. This yields a unique time $0 < t_1^* = t_2^*$ at which $(p_t, \Lambda_t)$ hits the boundary separating regions (2) and (1); subsequently consumers adopt immediately upon an opportunity so that $N_t = \rho e^{-\rho(t-t_2^*)}\bar{N}_{t_2^*}$ uniquely pins down the evolution of $(p_t, \Lambda_t)$. If $(p_0, \Lambda_0)$ is in region (3), then again no initial adoption occurs and the no-news posterior drifts upward according to the law of motion $\dot{p}_t = p_t(1-p_t)\varepsilon$, while $\Lambda_t$ remains unchanged at $\Lambda_0$. However, now this yields a unique time $0 < t_1^*$ at which $(p_t, \Lambda_t)$ hits the boundary separating regions (3) and (4), and at this time $\Lambda_{t_1^*} = \Lambda_0 > \Lambda(p_{t_1^*}) = \Lambda(\overline{p})$, so that we must have $t_1^* < t_2^*$. From $t_1^*$ on the evolution of $(p_t, \Lambda_t)$ is uniquely pinned down by the second line of Equation (4).[35] Thus, $t_2^*$ is uniquely given by the first time $t$ at which $\Lambda_t = \Lambda^*(p_t)$, at which point $(p_t, \Lambda_t)$ enters region

[35]Specifically, combining the second line of Equation (4) with Equation (1) yields the ODE $\dot{p}_t = rp_t(2p_t - 1)$, which pins down $p_t$ uniquely given the initial value $p_{t_1^*} = \overline{p}$:

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

(1). Similar arguments show that when $(p_0, \Lambda_0)$ starts in region (4), we have $t_1^* = 0$ and $t_2^* > t_1^*$ is the first time at which $(p_t, \Lambda_t)$, evolving according to the second line of Equation (4), enters region (1). Finally, if $(p_0, \Lambda_0)$ is in region (1), then $0 = t_1^* = t_2^*$ and absent breakdowns all consumers adopt upon their first opportunity from the beginning.

**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 3 shows that if consumers are forward-looking and not too optimistic ($p_0 < p^\sharp$), then $t_1^* < t_2^*$ holds whenever the potential for social learning $\Lambda_0$ is sufficiently large. The following lemma states this precisely:

**Lemma 3.5.** *Fix $\rho$, $\varepsilon$ and $p_0$ satisfying Conditions 3.3 and 3.4. Assume $p_0 < p^\sharp$. Then for all $r > 0$, there exists $\bar{\Lambda}_0(r) > 0$ such that $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$[36] if and only if $\Lambda_0 > \bar{\Lambda}_0(r)$.*

*Proof.* Set $\bar{\Lambda}_0(r) := \max\{\Lambda^*(p_0), \Lambda^*(\overline{p})\}$ and see Online Appendix B.4 . ■

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 3.2 no consumers adopt until the no-news posterior hits $\underline{p}$ (at $t_1^* = t_2^*$) and from then on all consumers adopt immediately when given a chance. The latter case corresponds to $\underline{p} = \overline{p} = \frac{1}{2}$ 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.

---

Plugging this back into $N_t = \frac{r(2p_t-1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda}$ uniquely pins down $\Lambda_t = \lambda \bar{N}_t$. Note that since $p_{t_1^*} > \frac{1}{2}$, $p_t$ given above is strictly increasing and reaches $p^\sharp$ in finite time. Thus $t_2^* = \inf\{t : \Lambda_t < \Lambda^*(p_t)\} < +\infty$.

[36] Note that by the Markovian description of equilibrium dynamics, $\Lambda_0$ is a sufficient statistic for equilibrium; i.e., holding all other fundamentals fixed, $\Lambda_0$ fully pins down the corresponding no-news equilibrium adoption flow, beliefs and cutoff times $t_1^*(\Lambda_0)$ and $t_2^*(\Lambda_0)$.

## 3.3 Equilibrium under Perfect Good News

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

**Theorem 3.6** (**Equilibrium under PGN**). *Let $r, \rho, \lambda, \bar{N}_0 > 0$, $p_0 \in (0,1)$, and $\varepsilon \geq 0$. There exists a unique equilibrium. Moreover, in the unique equilibrium, $N_t$ is Markovian in $(p_t, \Lambda_t)$ (or equivalently $(p_t, \bar{N}_t)$) for all $t$ and satisfies:*

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

*where*

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

To prove Theorem 3.6 we again invoke the quasi-single crossing property for equilibrium incentives established in Theorem 3.1. As we saw in Section 3.1, this implies that in any equilibrium, there are times $0 \leq t_1^* \leq t_2^* \leq +\infty$ defined by Equation (3) 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 for more information.

The key observation (Lemma B.6 in Online Appendix B.3) 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 $N_\tau > 0$ throughout $[t, t + \Delta)$.[37] As with perfect bad news, we can compare a consumer's payoff to adopting at $t$ with the payoff to delaying his decision by

[37] By 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.

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 with perfect 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 given an opportunity 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 perfect 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. On the other hand, 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 a fundamental difference between the bad news and good news setting. In order to maintain indifference over a period of time between immediate adoption and waiting, it must be possible to acquire *decision-relevant information* by waiting: 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. On the other hand, 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; if there is no breakthrough, then consumers remain indifferent at $t + dt$, so in either case the information obtained is not decision-relevant.[38]

[38] Note that breakthroughs do of course convey decision-relevant information at beliefs where consumers strictly prefer to delay. But during a region of indifference, this cannot be

Given that $t_1^* = t_2^* = t^*$, Theorem 3.6 follows from the observation that $p_t \leq p^*$ if and only if $t \geq t^*$ (Lemma B.7 in Online Appendix B.3). 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-looking nature is reflected by the fact that the cutoff posterior $p^*$ below which consumers are unwilling to adopt is $\frac{(r+\rho)(r+\varepsilon)}{2(r+\rho)(r+\varepsilon)-\rho\varepsilon} > \frac{1}{2}$. In both cases, the cutoff posterior does not depend on $\lambda$ or $\bar{N}_0$: Social learning only affects the *time* $t^*$ at which adoption ceases conditional on no breakthroughs.

# 4 Implications

## 4.1 Adoption Curves: S-Shaped vs. Concave

The differing informational incentives of bad and good news environments have observable implications. Consider the *adoption curve* of the innovation, which plots the percentage 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 3.2 and 3.6 translate directly into different predictions for the shape of the adoption curve, as summarized by the following corollary: Under perfect bad news, $A_t$ exhibits an *S-shaped* (i.e. convex-concave) growth pattern, where the region of convex growth coincides precisely with the partial adoption region $(t_1^*, t_2^*)$. By contrast, under perfect good news, adoption proceeds in *concave "bursts"*:

**Corollary 4.1.** ***Perfect Bad News:*** *In the unique equilibrium of Theorem 3.2, $A_t$ 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.*

***Perfect Good News:*** *In the unique equilibrium of Theorem 3.6, $A_t = 1 - e^{-\rho t}$ for all $t < t^*$, which is strictly increasing and concave. If there is a breakthrough prior to $t^*$, then the proportion of adopters is given by $1 - e^{-\rho t}$*

the case.

*for all $t$; if the first breakthrough occurs at $s > t^*$,*[39] *then adoption comes to a temporary standstill between $t^*$ and $s$, and for all $t \geq s$, the proportion of adopters is strictly increasing and concave and given by $1 - e^{-\rho(t^*+t-s)}$.*

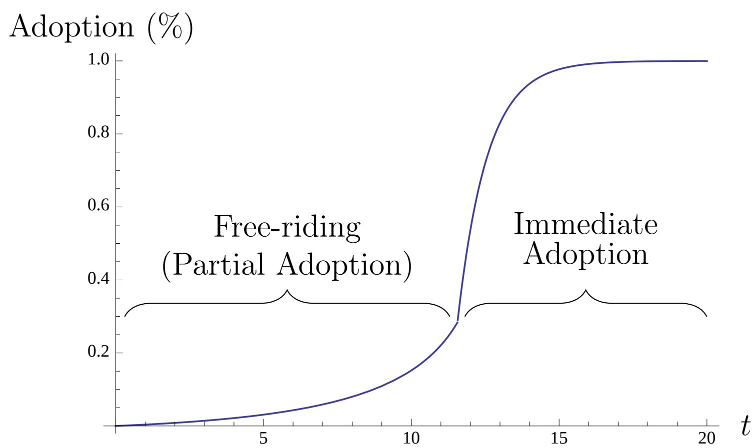


Figure 4: Adoption curve under PBN conditional on no breakdowns ($\varepsilon = 0$)

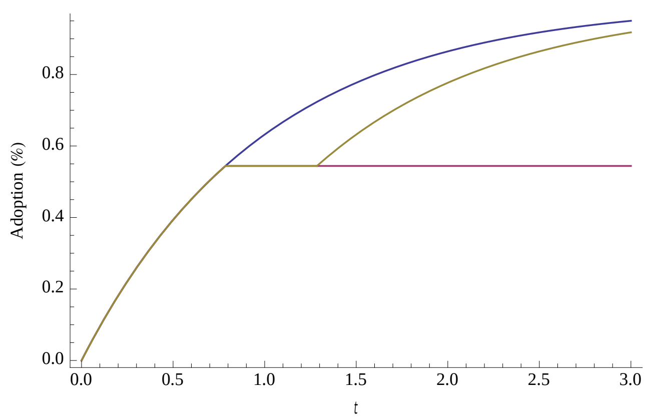


Figure 5: Adoption curves under PGN (blue = breakthrough before $t^*$; yellow = breakthrough after $t^*$; pink = bad quality)

Figures 4 and 5[40] illustrate the differing adopting patterns. As we discussed in the Introduction, both patterns have been widely documented empirically,

[39]This occurs only if $\varepsilon > 0$.

[40]Associated parameter values: $\varepsilon = 1/2$, $r = 1$, $\rho = 1$, $\lambda = 0.5$, and $p_0 = 0.7$.

but our model differs from existing explanations in identifying a purely informational source of this regularity: We predict S-shaped adoption curves in bad news markets with a sufficiently large potential for social learning and sufficiently forward-looking and not too optimistic consumers (so that $t_1^* < t_2^*$ by Lemma 3.5), and concave adoption patterns in good news markets (or in bad news markets with little potential for social learning and with very optimistic and impatient consumers).

The convex[41] growth region of $A_t$ under perfect bad news coincides precisely with the partial adoption region $(t_1^*, t_2^*)$ and is tied to consumer indifference in this region: Conditional on no breakdowns during this period, consumers grow increasingly optimistic about the quality 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.[42] As we discussed following Lemma 3.5, partial adoption depends on the *joint* assumption of forward-looking consumers and social learning. This is why we are able to generate S-shaped adoption curves even when consumers are ex ante identical, whereas existing learning-based models with myopic consumers (Young, 2009) or purely exogenous learning (Jensen, 1982) must appeal to specific distributions of consumer heterogeneity.[43]

Our predictions suggest the need for empirical work that would systemat-

[41]The regions of *concave* growth under both perfect bad and good news result simply from the gradual depletion of the population of remaining consumers.

[42]This argument for convex growth does not rely on linearity of $\lambda N_t$; it remains valid as long as the rate at which the bad product generates breakdowns at $t$ is increasing in $N_t$.

[43]See footnote 12. One exception is Kapur (1995), where a finite number of identical firms engage in a sequence of waiting contests to adopt a new technology and more information is revealed when more firms adopt during a given waiting contest. This can be viewed as a form of forward-looking social learning. He shows that the mean duration of waiting contests shrinks over time, suggesting a crude approximation of convex diffusion.

ically investigate the qualitative and quantitative features of consumer learning about different innovations and compare the associated adoption patterns. Here we provide some suggestive evidence:

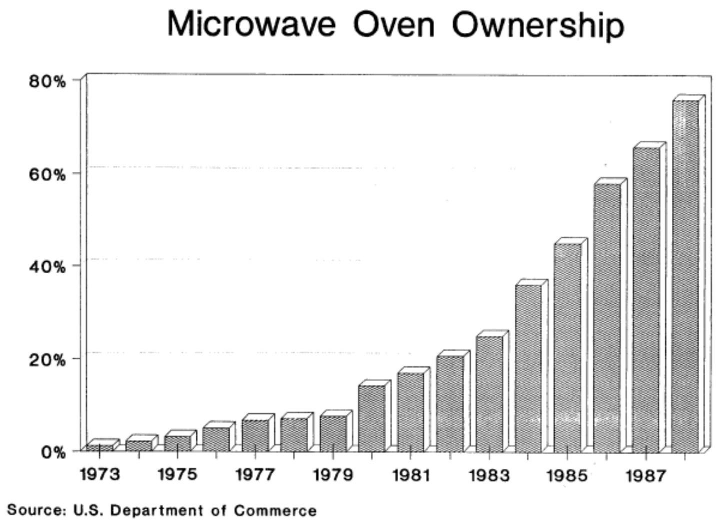


Figure 6: Adoption of microwaves by US households (Guenthner et al. (1991))

Learning via bad news events (or their absence) seems especially plausible in the case of new technologies or medical procedures whose introduction was accompanied by initial safety concerns: For example, following Raytheon's introduction to the US market of the first countertop household microwave oven in 1967, the 1970s were characterized by widespread concerns about possible "radiation leaks", stirred up for instance by a Consumers' Union (1973) report which concluded that "we are not convinced that they are completely safe to use"[44] and by Paul Brodeur's 1977 bestseller *The Zapping of America*.[45] Thus, it seems plausible that some consumers would have delayed their purchase in the hope of learning whether previous adopters experienced any adverse effects, as suggested for instance by Wiersema and Buzzell (1979).[46] Consistent

[44] Consumers' Union (1973), p. 221

[45] The FDA's Bureau of Radiological Health disagreed with the concerns. For details see Wiersema and Buzzell (1979).

[46] Ibid., p. 2. We note that adoption levels remained relatively low throughout the 1970s

with our predictions for bad news markets, the microwave is a textbook example of an innovation with an S-shaped adoption pattern: Figure 6 shows the convex growth in US adoption levels through the late 1980s, with later growth slowing to reach ownership levels of around 97% in 2011.[47] A second example is bariatric surgery, a collection of surgical weight loss procedures (including gastric bypass and gastric band surgery) which began gaining momentum in the mid-1990s. As with any major surgery, complications are possible, with typical health advice websites containing statements such as "a small degree of risk, including death, is inherent to all types of surgery" and "Because bariatric surgery is a relatively new surgical specialty, there are not yet enough medical data to predict with certainty which patients will have better outcomes."[48] Again, consistent with some patients deciding to delay the procedure to learn whether previous adopters suffered serious complications, the available data suggests an S-shaped growth pattern.[49]

Concave adoption patterns have been studied in the marketing literature under the name "fast-break product life cycles", with movies (Figure 7), books, music and similar leisure-enhancing products as canonical examples.[50] Consistent with our predictions, these domains appear to better fit the good news than the bad news model. For example, in a 2003–2004 study of consumer reviews of a representative sample of 6405 books on Amazon.com and BarnesandNoble.com, Chevalier and Mayzlin (2006) find that reviews "are over-

despite the fact that the entry of Japanese firms onto the US market in the mid-1970s brought with it substantial price decreases (from \$550 in 1970 to as low as \$150 in 1978, ibid. p. 2 and p. 5), possibly lending further plausibility to safety concerns as the primary source of delays.

[47] Williams (2014), p. 2.

[48] `http://health.usnews.com/health-conditions/heart-health/information-on-bariatric-surgery/overview#4`.

[49] According to Buchwald and Oien (2009) p. 1609 and Buchwald and Oien (2013) p. 428, the annual number of procedures performed 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. We note that an explanation in terms of reduced costs does not seem possible: For example, in the US the number of annual procedures increased from 13,386 to 121,055 between 1998 and 2004, while the average cost per procedure saw only a limited decrease, from \$10,970 to \$10,395; cf. Zhao and Encinosa (2007) Table 1, p. 6.

[50] Cf. Keillor (2007), pp. 51–61.

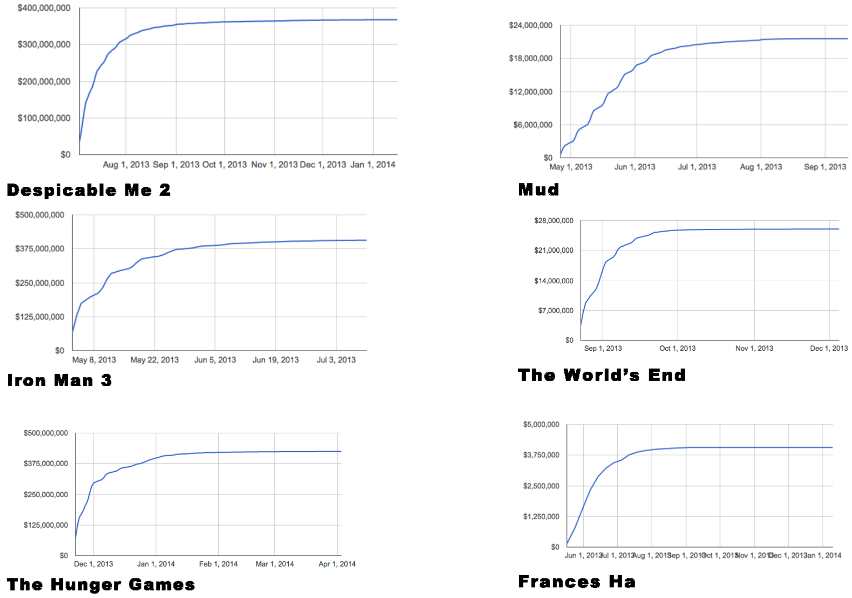


Figure 7: 2013 cumulative box office sales for various blockbuster (left) and independent (right) movies (Source: http://www.the-numbers.com)

whelmingly positive overall at both sites,"[51] suggesting that social learning in this domain proceeds via good news signals (or their absence) rather than via bad news signals: On a scale from one (worst) to five (best) stars, the modal review in the study is 5 stars, the mean star rating exceeds 4, and the fraction of 1-star ratings is in the range of 0.03–0.08.[52] As far as exogenously generated news is concerned, it would again appear that positive events, such as Academy Award, Grammy Award, or Booker Prize wins, receive far greater coverage than the occasional damning review by a critic (for this reason, Board and Meyer-ter Vehn (2013)[53] also cite the movie industry as an example of a good news market). Based on our model, we would also conjecture concave adoption patterns for (essentially side-effect free) herbal remedies and other alternative medical treatments, and for many beauty and fitness products, for

[51] Chevalier and Mayzlin (2006), p. 347.
[52] Ibid., Table 1, p. 347.
[53] Board and Meyer-ter Vehn (2013) footnote 2, p. 2382.

which anecdotal evidence suggests that consumer learning is primarily about "whether they actually work" (i.e. good news events or their absence).

## 4.2 The Effect of Increased Opportunities for Social Learning

How does an increase in the potential for social learning $\Lambda_0 := \lambda \bar{N}_0$ affect welfare, learning, and adoption dynamics? Again, the differing informational incentives of bad and good news environments have important implications.

Under perfect bad news, an economy's ability to harness its potential for social learning is subject to a surprising *saturation effect*: Up to a certain cutoff level, increasing $\Lambda_0$ strictly increases ex-ante welfare, speeds up learning, and decreases expected adoption levels of bad products while leaving adoption levels of good products unaffected; but beyond this cutoff level, further increases in $\Lambda_0$ are ex-ante welfare-neutral, cause learning to slow down over certain periods, and strictly slow down the adoption of good products. By contrast, there is no such saturation effect under perfect good news.

Throughout this section we fix $r$, $\rho$, $\varepsilon$, and $p_0$ and study the effect of increasing $\Lambda_0$ on ex-ante equilibrium welfare $W_0(\Lambda_0)$; equilibrium cutoff times $t_1^*(\Lambda_0)$, $t_2^*(\Lambda_0)$; 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.[54]

**Perfect Bad News:** The following proposition, which we prove in Online Appendix B.4, summarizes the saturation effect.

**Proposition 4.2.** *Consider learning via perfect bad news. Fix* $r, \rho > 0$, $\varepsilon \geq 0$, *and* $p_0$ *satisfying Conditions 3.3 and 3.4 and such that* $p_0 \in (\overline{p}, p^\sharp)$.[55] *Consider* $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. *Then:*

[54]Note that because of the Markovian description of the equilibrium in Theorem 3.2 and Theorem 3.6, $\Lambda_0$ is a sufficient statistic for these quantities when all other parameters are fixed.

[55]We assume $p_0 \in (\overline{p}, p^\sharp)$, so that $t_1^* = 0$, to focus on the inefficiency due to partial adoption without having to take into account the effect of $\Lambda_0$ on $t_1^*$. As we show in Online Appendix B.4.1, the welfare-neutrality result remains valid if $p_0 \in (0, \overline{p}]$, but now the cutoff-level above which it holds is $\Lambda^*(\overline{p})$ rather than $\Lambda^*(p_0)$.

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

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

- *$p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0}$ for all $t \leq t_2^*(\Lambda_0)$,*
- *$p_t^{\Lambda_0} > p_t^{\hat{\Lambda}_0}$ for all $t \in (t_2^*(\Lambda_0), \bar{t})$,*
- *$p_t^{\Lambda_0} < p_t^{\hat{\Lambda}_0}$ for all $t > \bar{t}$.*

(iii). ***Slowdown in Adoption:*** *For all $t$ and $\theta = B, G$, $A_t(\Lambda_0, \theta) > A_t(\hat{\Lambda}_0, \theta)$.*

*On the other hand, if $\Lambda_0 < \hat{\Lambda}_0 \leq \Lambda^*(p_0)$, then $W_0(\Lambda_0) < W_0(\hat{\Lambda}_0)$; $p_t^{\Lambda_0} < p_t^{\hat{\Lambda}_0}$; $A_t(\Lambda_0, G) = A_t(\hat{\Lambda}_0, G)$ and $A_t(\Lambda_0, B) > A_t(\hat{\Lambda}_0, B)$ for all $t$.*

The saturation effect obtains once $\Lambda_0$ exceeds $\Lambda^*(p_0)$. This is precisely the level above which the equilibrium features an initial partial adoption region ($0 = t_1^* < t_2^*(\Lambda_0)$), so that consumers at time 0 are indifferent between delaying and adopting. This immediately implies welfare-neutrality, because $W_0(\Lambda_0) = 2p_0 - 1$ irrespective of the value of $\Lambda_0 \geq \Lambda^*(p_0)$.[56] This result is in stark contrast to the cooperative benchmark in which consumers coordinate on socially optimal adoption levels: Here increased opportunities for social learning are always 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 $\Lambda_0 \to \infty$.[57]

(ii) and (iii) further illuminate the forces behind welfare-neutrality: Because an increase in $\Lambda_0$ affects learning dynamics in a non-monotonic manner, 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 $\Lambda_0$ and $\hat{\Lambda}_0$; if $t \in (t_2^*(\Lambda_0), \bar{t})$, he is strictly worse off under $\hat{\Lambda}_0$,

[56] As discussed in the previous footnote, as long as $\hat{\Lambda}_0 > \Lambda_0 > \max\{\Lambda^*(\bar{p}), \Lambda^*(p_0)\}$, the welfare-neutrality result remains valid even if $p_0 < \bar{p}$ in which case $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0) > 0$.

[57] The cooperative benchmark is derived in Section 3.2 of an earlier version of this paper, Frick and Ishii (2014): It takes 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 types of inefficiency: First, because $p^s < \underline{p}$, adoption generally begins *too late*. Second, whenever $t_1^* < t_2^*$, then once consumers begin to adopt, the initial rate of adoption is *too low*. Cf. Frick and Ishii (2014), section 5.3.

because in case the innovation is bad he is *less* likely to have found out by then than under $\Lambda_0$; 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 $\Lambda_0$ strictly decreases $A_t(\Lambda_0, G)$ (which is harmful), but also decreases $A_t(\Lambda_0, B)$ (which is beneficial), and welfare-neutrality is achieved because these forces balance out in equilibrium. Figure 8 illustrates that the strict slow-down in the adoption of good products is due to two effects: On the *extensive margin*, the increase in $\Lambda_0$ pushes out $t_2^*$ (i.e. prolongs free-riding in the form of partial adoption); on the *intensive margin*, the increase strictly drives down the growth rate of $A_t$ at all $t < t_2^*(\Lambda_0)$.

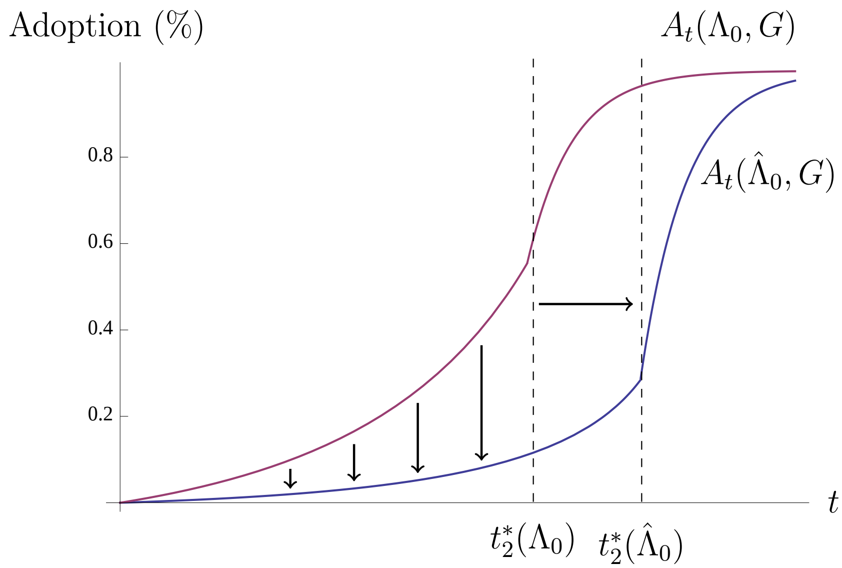


Figure 8: Changes in adoption levels of a good product as a result of increased opportunities for social learning under perfect bad news ($\hat{\Lambda}_0 > \Lambda_0$)

Since it only arises in the presence of partial adoption, the saturation effect once again relies crucially on the interaction between forward-looking consumers and social learning, setting us apart from models of myopic social learning or forward-looking exogenous learning in which ex-ante welfare necessarily increases in response to more informative signals (*even if* consumers are heterogeneous).[58]

**Perfect Good News:** Under perfect good news, there is no partial adoption. Correspondingly, there is no saturation effect:[59]

**Proposition 4.3.** *Consider learning via perfect good news. Fix $r, \rho > 0$, $\varepsilon \geq 0$, and $p_0 \in (p^*, 1)$.*[60] *Suppose $\hat{\Lambda}_0 > \Lambda_0 \geq 0$.*[61] *Then:*

(i). ***Strict Welfare Gains:*** *Provided $\varepsilon > 0$,*[62] *we have $W_0(\hat{\Lambda}_0) > W_0(\Lambda_0)$.*

(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 $\theta = B, G$.*

## 4.3 More Social Learning Can Hurt: An Example

Assuming ex-ante identical consumers, Proposition 4.2 established a saturation effect under perfect bad news: Beyond a certain level of $\Lambda_0$, further increases in the potential for social learning are welfare-neutral. Perhaps even more surprisingly, we show in this section that when consumers are *heterogeneous*,

[58]To define ex-ante welfare with myopic consumers, we assume that consumers' payoffs are discounted at some arbitrary rate $r > 0$, but that consumers behave myopically, i.e. ignore the option value to waiting.

[59]Nevertheless, 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.

[60]Recall that $p^* := \frac{(\varepsilon+r)(\rho+r)}{2(\varepsilon+\rho)(\varepsilon+r)-\varepsilon\rho}$ is the equilibrium cutoff posterior under perfect 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.

[61]If $\varepsilon = 0$ we assume that $p_0 \left(1 + e^{-\Lambda_0}\right) < 1$ so that $t^*(\Lambda_0) < \infty$.

[62] Increasing $\Lambda_0$ 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 $W(\Lambda_0) = W(\hat{\Lambda}_0)$.

increased opportunities for social learning can bring about *Pareto-decreases* in ex-ante welfare. To illustrate this, we introduce some heterogeneity in consumers' patience levels.

Consider a population consisting 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 $r_i > r_p$. Because our aim is simply to construct an example exhibiting welfare loss, we restrict attention to the perfect bad news setting. To simplify the analysis we assume that $\varepsilon = 0$ and $p_0 > 1/2$, but our arguments extend easily to the case where $\varepsilon > 0$.

Recall from Section 3.2 that for any discount rate $r > 0$, we can define the function $\Lambda_r^*$ implicitly for every $p \in (\frac{1}{2}, \frac{\rho+r}{\rho+2r})$:

$$2p - 1 = G_r(p, \Lambda_r^*(p)) := \int_0^\infty \rho e^{-(r+\rho)\tau} \left(p - (1-p)e^{-\Lambda_r^*(p)\left(1-e^{-\rho t}\right)}\right) d\tau.$$

Suppose $p_0 < \frac{\rho+r_p}{\rho+2r_p}$ and $\hat{\lambda}\bar{N}_0^p > \lambda\bar{N}_0^p > \Lambda_{r_p}^*(p_0)$ and consider first the game consisting only of mass $\bar{N}_0^p$ consumers of type $r_p$ (and no consumers of type $r_i$). Then Theorem 3.2 implies that the two equilibria corresponding to information structures $\lambda$ and $\hat{\lambda}$ both feature initial regions of partial adoption, so that $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$.

The following theorem states that provided the mass of impatient types is small, then in the game consisting of *both* types of consumers, the patient types' ex-ante payoffs continue to be $2p_0 - 1$ under both $\lambda$ and $\hat{\lambda}$; however, the impatient types' ex-ante payoffs are *strictly* lower under $\hat{\lambda}$ than under $\lambda$:

**Theorem 4.4.** *Suppose* $0 < r_p < r_i < +\infty$ *and* $p_0 \in (\frac{1}{2}, \frac{\rho+r_p}{\rho+2r_p})$. *Fix* $\bar{N}_0^p > 0$ *and* $\hat{\lambda} > \lambda > 0$ *such that* $\hat{\lambda}\bar{N}_0^p > \lambda\bar{N}_0^p > \Lambda_{r_p}^*(p_0)$. *Then there exists* $\eta > 0$ *such that whenever* $\bar{N}_0^i < \eta$, *then* $W_0^i(\hat{\lambda}) < W_0^i(\lambda)$ *and* $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$. *Thus, whenever* $\bar{N}_0^i < \eta$, *the ex-ante payoff profile* $(W_0^i(\lambda), W_0^p(\lambda))$ *in the* $\lambda$-*equilibrium Pareto-dominates the ex-ante payoff profile* $(W_0^i(\hat{\lambda}), W_0^p(\hat{\lambda}))$ *in the* $\hat{\lambda}$-*equilibrium.*

The proof is in Online Appendix B.6. The basic idea is as follows. Consider first the equilibrium adoption flows that are generated under each of $\lambda$ and $\hat{\lambda}$ in the game consisting solely of mass $\bar{N}_0^p$ of patient consumers of type $r_p$. What are the payoffs that a hypothetical impatient type $r_i$ (which does not exist in this game) would obtain if he were to behave optimally when faced with these adoption flows (and the expected future information they imply)? Since the patient types are initially indifferent between adopting or delaying in both equilibria, a monotonicity argument in types shows that in both cases the optimal strategy of the hypothetical impatient type $r_i$ is to adopt upon first opportunity. Given this, the ex-ante payoff of the hypothetical type $r_i$ under signal arrival rate $\gamma \in \{\lambda, \hat{\lambda}\}$ satisfies:

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

By the non-monotonicity result for learning established in Proposition 4.2, there exists $\bar{t} > t^* := t_2^*(\lambda)$ such that $p_\tau^{\hat{\lambda}} = p_\tau^\lambda$ for all $\tau \leq t^*$, $p_\tau^{\hat{\lambda}} < p_\tau^\lambda$ for all $\tau \in (t^*, \bar{t})$ and $p_\tau^{\hat{\lambda}} > p_\tau^\lambda$ for all $\tau > \bar{t}$. We now exploit the expressions for the value to waiting of the two types together with the deceleration of learning at times just after $t^*$ to obtain the result. Intuitively, since $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$, the cost of the deceleration in learning on $(t^*, \bar{t})$ and the benefit of the acceleration in learning at times after $\bar{t}$ must balance out in such a way that the patient type $r_p$ obtains the same ex-ante payoff under $\lambda$ and $\hat{\lambda}$. But as a result, these adjustments must strictly hurt the less patient hypothetical type $r_i$, because relative to type $r_p$, type $r_i$ weights the early losses due to the slow-down in learning more heavily than the later benefits due to the acceleration.

To complete the proof, we show that as long as $\bar{N}_0^i > 0$ is sufficiently small, we must 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 simple continuity argument. The second equality reflects the fact that provided $\bar{N}_0^i$ is sufficiently small, the patient type must continue to partially adopt initially in both equilibria.

Note that a crucial assumption underlying the above argument is that adoption opportunities are *stochastic and limited.* When $\rho$ is finite, the impatient types may not receive any adoption opportunities for a long time. But as we saw above, if an impatient type obtains his first adoption opportunity between $t^*$ and $\bar{t}$, then the information gained is strictly lower under the equilibrium with information process $\hat{\lambda}$ than under $\lambda$, which is precisely the cause of the impatient type's welfare loss. If on the other hand consumers were able to adopt freely at any time, then the impatient types would incur no losses as all of them would adopt immediately at time 0 in both the $\lambda$ and $\hat{\lambda}$-equilibrium. Thus, the above example illustrates an interesting interaction between heterogeneity and delays due to limited opportunities for adoption.

# 5 Conclusion

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 the role of the news environment in shaping these incentives; most importantly, in 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 explanation for why adoption curves are S-shaped for some innovations and concave for others. Moreover, partial adoption has important welfare implications, entailing that increased opportunities for social learning need not benefit consumers and can be strictly harmful.

To illustrate these points in the simplest possible framework, we have restricted attention to perfect bad and good news Poisson learning. This made our equilibrium analysis very tractable, yielding closed-form expressions for all key quantities and allowing us to compute numerous comparative statics. Nevertheless, many of our conclusions extend to more general information structures: Especially worth noting is the fact 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 as was the case under *perfect* bad news). Without such events (e.g. when learning is based on imperfect 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 a consumer who is willing to adopt cannot acquire decision-relevant information by delaying his decision by an instant.[63]

To highlight the implications of purely informational considerations, we have abstracted away from forces emphasized by existing models of innovation adoption, notably consumer heterogeneity and supply-side factors such as pricing. Nevertheless, exploring the way in which these forces *interact* with informational incentives represents an interesting avenue for future theoretical work. To give a taste, Section 4.3 shows that heterogeneity can further exacerbate the welfare implications of informational free-riding.

Finally, our predictions lend themselves to empirical investigation. Section 4.1 provides some suggestive evidence for the prediction that S-shaped (respectively concave) adoption curves are typical of bad (respectively good) news markets, but a more systematic analysis is called for. The saturation effect implies that the proportion of adopters of an innovation may grow more slowly in communities with more potential consumers or with a greater ease of information transmission. The former could be tested by contrasting the adoption paths of new agricultural technologies across villages with different population sizes,[64] while for the latter one might exploit the staggered introduction of certain social media platforms across different US cities or differences across states in legislation mandating the disclosure of adverse medical events.

# References

Bandiera, O. and I. Rasul (2006): "Social Networks and Technology Adoption in Northern Mozambique," *Economic Journal*, 116, 869–902.

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

[64]This is related to Bandiera and Rasul's (2006) finding which we discussed in footnote 1: They find that an individual farmer's likelihood of adoption is (from a certain point on) decreasing in the number of adopters in his network. But then, in equilibrium, larger networks of farmers should feature lower percentages of adoption.

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.

——— (1980): "The Relationship between Diffusion Rates, Experience Curves, and Demand Elasticities for Consumer Durable Technological Innovations," *Journal of Business*, S51–S67.

Bergemann, D. and U. Hege (1998): "Venture Capital Financing, Moral Hazard, and Learning," *Journal of Banking & Finance*, 22, 703–735.

——— (2005): "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 719–752.

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.

Bolton, P. and C. Harris (1999): "Strategic Experimentation," *Econometrica*, 67, 349–374.

Bonatti, A. and J. Hörner (2011): "Collaborating," *American Economic Review*, 101, 632–663.

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 (2014): “Optimal Design for Social Learning,” Mimeo.

Chevalier, J. A. and D. Mayzlin (2006): “The Effect of Word of Mouth on Sales: Online Book Reviews,” *Journal of Marketing Research*, 43, 345–354.

Conley, T. G. and C. R. Udry (2010): “Learning about a New Technology: Pineapple in Ghana,” *American Economic Review*, 100, 35–69.

Consumers' Union (1973): “Microwave Ovens: Not Recommended,” *Consumer Reports*, 38, 221–228.

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.

Ellison, G. and D. Fudenberg (1993): “Rules of Thumb for Social Learning,” *Journal of Political Economy*, 101, 612–643.

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,” *February 2014 version, available at* http://scholar.harvard.edu/mfrick.

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.

Halac, M., Q. Liu, and N. Kartik (2013): "Optimal Contracts for Experimentation," Mimeo.

——— (2014): "Contests for Experimentation," Mimeo.

Halac, M. and A. Prat (2014): "Managerial Attention and Worker Engagement," Mimeo.

Hörner, J. and L. Samuelson (2013): "Incentives for Experimenting Agents," *RAND Journal of Economics*, 44, 632–663.

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 Economic Theory*, 27, 182–193.

Jovanovic, B. and S. Lach (1989): "Entry, Exit, and Diffusion with Learning by Doing," *American Economic Review*, 690–699.

Kapur, S. (1995): "Technological Diffusion with Social Learning," *Journal of Industrial Economics*, 43, 173–195.

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.

——— (2014): "Breakdowns," *Theoretical Economics*, forthcoming.

Keller, G., S. Rady, and M. Cripps (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73, 39–68.

Klein, N. and S. Rady (2011): "Negatively Correlated Bandits," *Review of Economic Studies*, 1–40.

MacLeod, W. B. (2007): "Reputations, Relationships, and Contract Enforcement," *Journal of Economic Literature*, 595–628.

Mahajan, V., E. Muller, and F. M. Bass (1990): "New Product Diffusion Models in Marketing: A Review and Directions for Research," *Journal of Marketing*, 1–26.

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.

Rosenberg, D., E. Solan, and N. Vieille (2007): "Social Learning in One-Arm Bandit Problems," *Econometrica*, 75, 1591–1611.

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.

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.*

Young, H. P. (2009): "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *American Economic Review*, 99, 1899–1924.

Zhao, Y. and W. Encinosa (2007): "Bariatric Surgery Utilization and Outcomes in 1998 and 2004," *Agency for Healthcare Research and Quality*, Statistical Brief #23, 1–7.

# A Proof of Theorem 3.1

This appendix establishes the quasi-single crossing property for equilibrium incentives (Theorem 3.1). All remaining proofs are in Online Appendix B. 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 $W^N_{t\geq 0}$ and the no-news posterior by $p^N_{t\geq 0}$.

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

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

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

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

$$\dot{W}^N_t = (r+p^N_t\varepsilon_G+(1-p^N_t)\varepsilon_B)W^N_t - p^N_t\varepsilon_G\frac{\rho}{\rho+r}.$$

The final two lemmas focus on learning via perfect bad news (PBN):

**Lemma A.4.** *Let $N_{t\geq 0}$ be an equilibrium under PBN. Suppose that $\varepsilon > 0$ or $p_0 > \frac{1}{2}$. Then $\lim_{t\to\infty} p^N_t = \mu(\varepsilon, \Lambda_0, p_0)$ and $\lim_{t\to\infty} W^N_t = \frac{\rho}{\rho+r}(2\mu(\varepsilon,\Lambda_0,p_0)-1)$, where*

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

**Lemma A.5.** *Suppose that learning is via PBN. 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 Theorem 3.1 under Perfect Good News:***
Let $\varepsilon = \varepsilon_G \geq 0 = \varepsilon_B$ and $\lambda = \lambda_G > 0 = \lambda_B$.

**Step 1:** $W_t = 2p_t - 1 \Longrightarrow W_\tau \geq 2p_\tau - 1$ 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:

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

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:** $W_t > 2p_t - 1 \Longrightarrow W_\tau > 2p_\tau - 1$ 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 Theorem 3.1 under Perfect 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 Theorem 3.1 is obvious. We now prove the theorem under the assumption that either $\varepsilon > 0$ or $p_0 > \frac{1}{2}$.

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

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

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

$$r(2p_{\underline{s}} - 1) \geq (1 - p_{\underline{s}})\varepsilon$$
$$\text{and}$$
$$r(2p_{\overline{s}} - 1) \leq (1 - p_{\overline{s}})\varepsilon.$$

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

**Step 2:** $W_t < 2p_t - 1 \Longrightarrow W_\tau < 2p_\tau - 1$ 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

$\overline{s} \geq \underline{s}$ such that $2p_{\overline{s}} - 1 = W_{\overline{s}}$ and $2p_s - 1 > W_s$ for all $s > \overline{s}$. Lemma A.2 implies the following two limits:

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

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

Thus,

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

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

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

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

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

Rearranging we obtain:

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

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 $2p_{s^*} - 1 = W_{s^*}$. ∎

# B Online Appendix: Not For Publication

## 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 perfect bad news, then

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

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

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

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, $N_\tau$ is continuous at all $\tau \in (t-\nu, t+\nu)$. From this it is immediate that $W_\tau^N$ is continuously differentiable in $\tau$ for all for all $\tau \in (t-\nu, t+\nu)$, because we have that

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

The derivative of $W_\tau^N$ 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 $W_\tau^N$ 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 $R_\Delta$, is continuously differentiable with respect to $\Delta$ and satisfies $\frac{d}{d\Delta} R_\Delta \equiv 0$. Taking the limit as $\Delta \to 0$ and since $\dot{W}_\tau^N = \lim_{\Delta \to 0} \frac{d}{d\tau} W_{\tau+\Delta}^N$ 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, $W_\tau^N$ satisfies

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

From this it is again immediate that $W_\tau^N$ is continuously differentiable in $\tau$.

To compute the derivative, we proceed as above, rewriting $W_\tau^N$ 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 $\Delta$ and taking the limit as $\Delta \to 0$, 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 $p_t^N \to 1$ as $t \to \infty$. But for any $t$, $\frac{\rho}{\rho+r}\left(2p_t^N - 1\right) \leq W_t^N \leq \frac{\rho}{\rho+r}$. This implies that $\lim_{t\to\infty} W_t^N = \frac{\rho}{\rho+r}$ 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 $N^* := \lim_{t\to\infty}\int_0^t N_s ds = \sup_t \int_0^t N_s ds \leq \bar{N}_0$. Let $p^* := \lim_{t\to\infty} p_t^N = \sup_t p_t^N$. 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:

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

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

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

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

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

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

$$
\int_{t_1^*}^\infty \rho e^{-(r+\rho)(s-t)} \left( p_{t_1^*}^N - (1-p_{t_1^*}^N) e^{-\int_{t_1^*}^s \lambda N_k\, dk} \right) ds > \frac{\rho}{r+\rho} \left( 2p_{t_1^*}^N - 1 \right) \geq 2p_{t_1^*}^N - 1, \tag{7}
$$

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, (7) 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 Equilibrium under Perfect Bad News (Theorem 3.2)

In this section we prove Theorem 3.2. For this we do *not* impose Conditions 3.3 or 3.4. Recall the following definitions which we motivated in Section 3.2: Define

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

and define $p^* := \min\{\overline{p}, p^\sharp\}$. Define $G : [0,1] \times \mathbb{R}_+ \to \mathbb{R}$ by

$$G(p, \Lambda) := \int_0^\infty \rho e^{-(r+\rho)\tau} \left(p - (1-p)e^{-\left(\varepsilon\tau + \Lambda\left(1 - e^{-\rho\tau}\right)\right)}\right) d\tau.$$

We extend the function 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 $\Lambda^* : [0,1] \to \mathbb{R}_+ \cup \{+\infty\}$ by

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

The proof of Theorem 3.2 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 Equation (2). We first show in Lemma B.1 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 B.3 provides a characterization of $t_1^*$ and $t_2^*$ in terms of the evolution of $(p_t, \Lambda_t)$. Given these two steps, it is easy to see that if an equilibrium exists, it is unique and must take the Markovian form in Equation (4) of Theorem 3.2. Finally, to verify equilibrium existence, Lemma B.4 shows that the adoption flow implied by Equation (4) is feasible.

### B.2.1 Characterization of Adoption between $t_1^*$ and $t_2^*$

**Lemma B.1.** *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 Theorem 3.1, 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 $t \in (t_1^*, t_2^*)$:

$$\begin{aligned} W_t^N &= e^{-r(t_2^*-t)}\left(p_t + (1-p_t)e^{-\int_t^{t_2^*}(\varepsilon+\lambda N_s)ds}\right)(2p_{t_2^*} - 1) \\ &= e^{-r(t_2^*-t)}\left(p_t - (1-p_t)e^{-\int_t^{t_2^*}(\varepsilon+\lambda N_s)ds}\right), \end{aligned} \tag{8}$$

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 (8), we obtain:

$$\dot{W}_t^N = (r + (\varepsilon + \lambda N_t)(1-p_t))\, W_t^N. \tag{9}$$

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

$$\dot{W}_t^N = 2\dot{p}_t = 2p_t(1-p_t)(\varepsilon + \lambda N_t). \tag{10}$$

Combining (9), (10) 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 $t \in (t_1^*, t_2^*)$. ■

As an immediate corollary of Lemma B.1 we obtain:

**Corollary B.2.** *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:*

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

### B.2.2 Characterization of Cutoff Times

**Lemma B.3.** *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). $t_1^* = \min\{t_2^*, \sup\{t \geq 0 : p_t < p^*\}\}$.[65]

[65] We impose the convention that if $\{t \geq 0 : p_t < p^* = \frac{1}{2}\} = \varnothing$, then $\sup\{t \geq 0 : p_t < p^* = \frac{1}{2}\} := 0$.

*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 $\lim_{t\to\infty} 2p_t - 1 > \lim_{t\to\infty} W_t$, 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, for all $t > t_2^*$, $\Lambda_t < \Lambda_{t_2^*}$ and $p_t > p_{t_2^*}$, so since $\Lambda^*$ 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 $p_{t_1^*} \leq \overline{p}$.

(c) If $t_1^* < t_2^*$, then $p_{t_1^*} \geq \overline{p}$.

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 $p_{t_1^*} = \overline{p} = p^*$. If $0 = t_1^* < t_2^*$, then (c) implies that $p_0 \geq \overline{p} = p^*$. 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

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

Finally, for claim (c), note that if $t_1^* < t_2^*$, then Lemma B.1 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 $\tau \in (t_1^*, t_2^*)$,

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

and hence by continuity $p_{t_1^*} \geq \overline{p}$ 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. ∎

Given Lemmas B.1 and B.3, it is immediate that if an equilibrium exists, then it must take the form of the adoption flow given by Equation (4) in Theorem 3.2. 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.

### B.2.3 Feasibility

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

**Lemma B.4.** *Suppose $N_{t\geq 0}$ is an adoption flow satisfying Equation (4) in Theorem 3.2 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 $t \in (t_1^*, t_2^*)$. ■

#### B.2.4 Equilibrium under Perfect Bad News without Condition 3.3

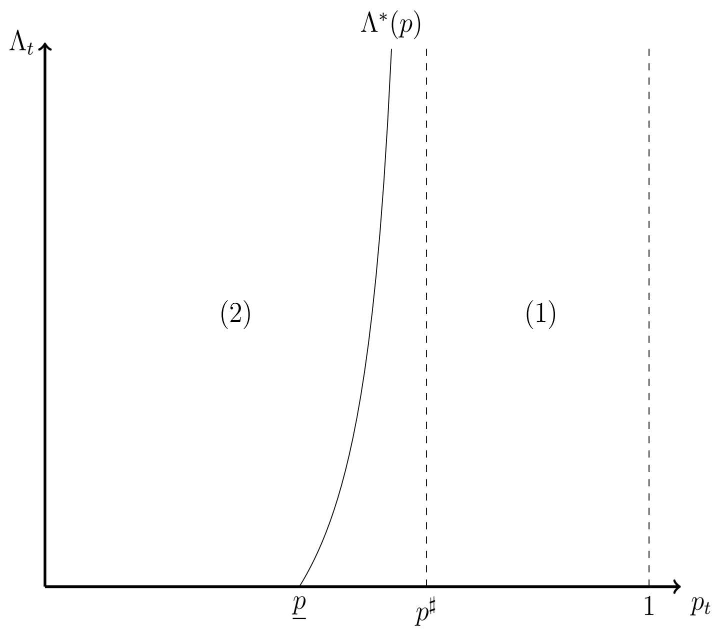


Figure 9: Partition of $(p_t, \Lambda_t)$ when $\varepsilon \geq \rho$

In this section, we discuss the case where $\rho \leq \varepsilon$, so that Condition 3.3 is violated. The previous sections established the equilibrium characterization of Theorem 3.2 without assuming Condition 3.3 . If $\rho \geq \varepsilon$, then $p^* = p^\sharp$, so because $\Lambda^*(p) = +\infty$ for all $p > p^\sharp$, we have:

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

Thus, there is no region of partial adoption. As a result, it is easy to see that the saturation effect discussed in Section 4.2 is no longer present and welfare always strictly increases in response to an increase in the potential for social learning:

**Proposition B.5.** *Fix $r > 0$ and $p_0 \in (0, 1)$ and suppose that $\varepsilon \geq \rho > 0$. Then $W_0$ is strictly increasing in $\Lambda_0$.*

## B.3 Equilibrium under Perfect Good News (Theorem 3.6)

Theorem 3.6 follows immediately from Lemma B.6 and Lemma B.7:

**Lemma B.6.** *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 Theorem 3.1 , 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:

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

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 $e^{-r\Delta} \left(2p_{t+\Delta} - 1\right)$.

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

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

where the final equality comes from Bayesian updating of beliefs. This contradicts $W_t = 2p_t - 1$. Thus, $t_1^* = t_2^*$. ∎

**Lemma B.7.** *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*

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

*Proof.* Define

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

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

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

$H(p,0)$ 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 $p^* := \frac{(\varepsilon+r)(\rho+r)}{2(\varepsilon+\rho)(\varepsilon+r)-\varepsilon\rho}$ 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. ■

## B.4 Comparative Statics under PBN (Proposition 4.2)

As in the text, we impose Conditions 3.3 and 3.4 throughout this section. Define $\overline{\Lambda}_0 := \max\{\Lambda^*(p_0), \Lambda^*(\overline{p})\}$. We first prove Lemma 3.5 from Section 3.2:

***Proof of Lemma 3.5:*** We show that $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$ if and only if $\Lambda_0 > \overline{\Lambda}_0$.

Suppose first that $\Lambda_0 > \overline{\Lambda}_0$. Then by the proof of the first part of Lemma B.3, 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 B.3, we must have $p_{t^*} \leq \overline{p}$. But combining these statements, we get

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

which is a contradiction.

Suppose conversely that $t_1^* < t_2^*$. Then by the proof of Lemma B.3, we have that $\Lambda^*(p_{t_1^*}) < \Lambda_{t_1^*} = \Lambda_0$. That proof also implies that if $0 < t_1^* < t_2^*$, then $p_{t_1^*} = \overline{p} \geq p_0$; and if $0 = t_1^* < t_2^*$, then $p_{t_1^*} = p_0 \geq \overline{p}$. Thus, either way $\Lambda_0 > \overline{\Lambda}_0$, as claimed. ■

The following three subsections prove Proposition 4.2, by considering the effect of an increase in $\Lambda_0$ on welfare, learning, and adoption behavior, respectively.

### B.4.1 Comparative Statics of Welfare

We prove a slightly more general result than in Proposition 4.2: We allow for any $p_0 \in (0, 1)$ and show that

- if $\hat{\Lambda}_0 > \Lambda_0 > \overline{\Lambda}_0$, then $W_0(\hat{\Lambda}_0) = W_0(\Lambda_0)$;
- if $\overline{\Lambda}_0 \geq \hat{\Lambda}_0 > \Lambda_0$, then $W_0(\hat{\Lambda}_0) > W_0(\Lambda_0)$.

If $p_0 \in (\bar{p}, p^\sharp)$ as in Proposition 4.2, then $\overline{\Lambda}_0 = \Lambda^*(p_0)$, so we get the result in Proposition 4.2.

To prove the first bullet point, consider $\hat{\Lambda}_0 := \Lambda_0^2 > \Lambda_0 := \Lambda_0^1 > \overline{\Lambda}_0$ 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 3.5, we have $t_1^i < t_2^i$ for $i = 1, 2$. Moreover, by the proof of Lemma B.3, we have $\max\{p_0, \overline{p}\} = p^1_{t_1^1} = p^2_{t_1^2}$. 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 $W_0^i = e^{-rt_1}\frac{p_{t_1}}{p_0}W_{t_1}^i$ for $i = 1, 2$, whence $W_0^1 = W_0^2$.

For the second bullet point, suppose $\Lambda_0^1 < \Lambda_0^2 \leq \overline{\Lambda}_0$. By Lemma 3.5, we must have $t_1^i = t_2^i =: t^i$. Let $\hat{t} := \min\{t^1, t^2\}$. Then note that for all $t \leq \hat{t}$, $p_{\hat{t}}^1 = p_{\hat{t}}^2$ and $\Lambda_{\hat{t}}^i = \Lambda_0^i$. By Lemma B.3 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

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

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

### B.4.2 Comparative Statics of Learning

In this section and Section B.4.3, we assume as in Proposition 4.2 that $p_0 \in (\bar{p}, p^\sharp)$. This implies that $t_1^* = 0$ and $\overline{\Lambda}_0 = \Lambda^*(p_0) < +\infty$.

Note first that $p_t^{\Lambda_0}$ is strictly increasing in $\Lambda_0$ for all $\Lambda_0 \in (0, \Lambda^*(p_0))$ since in this case $t_2^*(\Lambda_0) = 0$ so that

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

Suppose next that $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. To prove the non-monotonicity result in item (ii) of Proposition 4.2, we first prove the following lemma:

**Lemma B.8.** *Suppose that $\hat{\Lambda}_0 = \hat{\lambda}\hat{\bar{N}}_0 > \Lambda_0 = \lambda \bar{N}_0 > \Lambda^*(p_0)$, with corresponding equilibrium flows of adoption $\hat{N}_{t\geq 0}$ and $N_{t\geq 0}$. Then*

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

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

*Proof.* Suppose that $\hat{\Lambda}_0 > \Lambda_0 > \overline{\Lambda}_0 = \Lambda^*(p_0)$. Then by Lemma 3.5, 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 $p_0 = p_0^{\Lambda_0} = p_0^{\hat{\Lambda}_0}$, the ODE in Corollary B.2 implies that at all times $t < t_2^*$, we have $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0} = p_t$. By Lemma B.1, this implies that for all $t < t_2^*$,

$$\lambda N_t = \hat{\lambda}\hat{N}_t. \tag{11}$$

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 B.3 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 $t < t_2^* = t_2^*(\Lambda_0)$. ■

***Proof of item (ii) of Proposition 4.2:*** Suppose that $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. By Lemma B.8, $t^* := t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$, $\lambda N_t = \hat{\lambda}\hat{N}_t$, and $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0}$ for all $t \leq t^*$, which proves the first bullet point.

To prove the second bullet point, 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 $\Lambda_0$:

$$\lim_{t \uparrow t^*} \lambda N_t < \lim_{t \downarrow t^*} \lambda N_t. \tag{12}$$

In other words, there is necessarily a discontinuity in the equilibrium flow of adoption at exactly $t^*$. 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

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

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 B.3 implies

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

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

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

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

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

Therefore there must exist 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 point.

Finally, for the third bullet point, observe first that there must exist 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 once again get that $W_{t^*}^{\hat{\Lambda}_0} > W_{t^*}^{\Lambda_0}$, which is false. 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. On the other hand, 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

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

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 point. ■

### B.4.3 Comparative Statics of Adoption Behavior

**Adoption of Good Products:** *For all $t$, $A_t(\Lambda_0, G)$ is constant in $\Lambda_0$ for all $\Lambda_0 \leq \Lambda^*(p_0)$ and strictly decreasing in $\Lambda_0$ for all $\Lambda_0 > \Lambda^*(p_0)$.*

*Proof.* First note that because $p_0 \geq \overline{p}$, $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0) = 0$.

Then at all $\Lambda_0 < \Lambda^*(p_0)$, the adoption flow absent breakdowns satisfies $N_t = \rho\bar{N}_t$ for all $t$. 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 $t$ and all pairs $\Lambda_0, \hat{\Lambda}_0 \leq \Lambda^*(p_0)$.

Now suppose that $\hat{\Lambda}_0 > \Lambda_0 > \Lambda^*(p_0)$. Note that $N_t, \hat{N}_t > 0$ for all $t > 0$ (recall Condition 3.4). Let $t^* = t_2^*(\Lambda_0)$. By Lemma B.8, $\lambda N_t = \hat{\lambda}\hat{N}_t$ for all $t < t^*$. Then for all $t < t^*$

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

. Therefore for all $t < t^*$, we have $A_t(\Lambda_0, G) > A_t(\hat{\Lambda}_0, G)$.

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

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

where the second inequality follows from feasibility. But because $A_{t^*}(\Lambda_0, G) > A_{t^*}(\hat{\Lambda}_0, G)$, $A_t(\Lambda_0, G) > A_t(\hat{\Lambda}_0, G)$ for all $t > 0$. ■

**Adoption of Bad Products**: *For all $t > 0$, $A_t(\Lambda_0, B)$ is strictly decreasing in $\Lambda_0$.*

*Proof.* 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

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

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 $\lambda$ and $\bar{N}_0$ only through $\Lambda_0 = \lambda \bar{N}_0$, so we can denote it by $A_t(\Lambda_0, B)$. Then it suffices to prove the claim when $\lambda$ is increased to $\hat{\lambda} > \lambda$ holding fixed $\bar{N}_0$, because for any $\hat{\Lambda}_0 > \Lambda_0$ there exists $\bar{N}_0$ and $\hat{\lambda} > \lambda$ such that $\hat{\Lambda}_0 = \hat{\lambda}\bar{N}_0$ and $\Lambda_0 = \lambda \bar{N}_0$.

Let $N_{t\geq 0}$ and $\hat{N}_{t\geq 0}$ be the equilibrium under $\lambda$ and $\hat{\lambda}$, respectively. Note that when $\overline{p} \leq p_0$, $N_t > 0$ for all $t > 0$. 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 $\lambda$. This implies that for all $t > 0$,

$$\int_0^t N_\tau e^{-\int_0^\tau (\varepsilon+\lambda N_s)ds} d\tau > \int_0^t N_\tau e^{-\int_0^\tau \left(\varepsilon+\hat{\lambda} N_s\right)ds} d\tau. \tag{13}$$

We now show that

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

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

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

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

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

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

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

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

But this is equivalent to

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

which contradicts $\int_0^t N_s ds \geq \int_0^t \hat{N}_s ds$. This shows that for all $\hat{\lambda} > \lambda$ and $t > 0$, $A_t(\hat{\lambda}\bar{N}_0, B) < A_t(\lambda\bar{N}_0, B)$, as required. ■

## B.5 Comparative Statics under PGN (Proposition 4.3)

We prove Proposition 4.3.

### B.5.1 Strict Welfare Gains

*Proof.* If $p_0 > p^*$ and $\varepsilon > 0$, then under both $\Lambda_0$ and $\hat{\Lambda}_0$ consumers adopt immediately upon first opportunity until $p^*$ is reached and from then on delay adoption until there has been a breakthrough. Moreover, the probability $\pi^*$

of a breakthrough occurring prior to $p^*$ being reached is the same under both $\Lambda_0$ and $\hat{\Lambda}_0$: $\pi^* = \frac{p_0 - p^*}{1-p^*}$. Because learning occurs at the same exogenous rate $\varepsilon$ 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:

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

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 $\Lambda_0$ 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:

$$
\begin{aligned}
& \left(\pi^* \frac{\rho}{r+\rho} + (1-\pi^*)(2p^* - 1)\right) - \frac{\rho}{r+\rho}(2p_0 - 1) \\
= \; & \frac{r}{r+\rho}(1-\pi^*)(2p^* - 1) > 0.
\end{aligned}
$$

■

### B.5.2 Learning Speeds Up

*Proof.* 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 3.6, 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 $\varepsilon$, which means that conditional on $t > t^*$, beliefs depend only on $t - t^*$, as summarized in the third bullet point. ■

### B.5.3 No Initial Slow-Down in Adoption

*Proof.* 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 first opportunity in both equilibria. ■

## B.6 Proof of Theorem 4.4

We first establish the following basic mathematical fact:

**Lemma B.9.** *Suppose $\bar{t} > t^* \geq 0$ and consider $f, g : [0, \infty) \to \mathbb{R}$ 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

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

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. ■

To prove Theorem 4.4, fix $0 < r_p < r_i$, $\rho > 0$, $\bar{N}_0^p > 0$ and $p_0 \in (\frac{1}{2}, \frac{\rho+r_p}{\rho+2r_p})$. Consider $\hat{\lambda} > \lambda > 0$ such that $\hat{\lambda}\bar{N}_0^p > \lambda\bar{N}_0^p > \Lambda^*_{r_p}(p_0)$. As in the text, we assume that there is no exogenous news source. The following lemma derives the equilibrium of the game with a sufficiently small mass of impatient types:

**Lemma B.10.** *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:*

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

*Proof.* Fix $\gamma \in \{\lambda, \hat{\lambda}\}$. Pick $\eta > 0$ such that $p_0 > \frac{\eta+r}{\eta+2r}$. 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 3.2 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 Theorem 3.1, 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 3.2. ■

Given Lemma B.10, we can follow the arguments in the proof of Proposition 4.2 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 given by

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

Since $r_p$ is initially indifferent between adoption and delay under both $\lambda$ and $\hat{\lambda}$, we have $W_0^p(\lambda) = W_0^p(\hat{\lambda}) = 2p_0 - 1$. But then applying Lemma B.9 yields $W_0^i(\lambda) > W_0^i(\hat{\lambda})$. This completes the proof of Theorem 4.4.