Innovation Adoption by Forward-Looking Social Learners*

Mira Frick† Yuhta Ishii‡

(**JOB MARKET PAPER**)

First version: October 2, 2013
This version: February 9, 2014

Please click (here) for the latest version.

Abstract

A large population of long-lived consumers faces stochastic opportunities to adopt an innovation of uncertain quality. Consumers are *social learners*: Over time, news about the product's quality is generated endogenously, based on the experiences of past adopters. We analyze how the potential for social learning in an economy affects consumers' informational incentives and how these in turn shape the aggregate adoption dynamics of an innovation. Our main results highlight the importance of two features of the economy: The extent to which consumers are *forward-looking* and the nature of news events through which social learning occurs. When consumers are forward-looking social learners, the trade-off between the benefit of adopting the innovation at any given time and the option value of waiting for endogenous news can generate rich aggregate adoption dynamics, even in the absence of any consumer heterogeneity. The dynamics of this trade-off and the extent to which it is affected by increased opportunities for social learning interact in interesting ways with the news process of the economy. For a class of Poisson learning processes, we establish the existence and uniqueness of equilibria. In line with empirical findings, equilibrium adoption patterns are either S-shaped or feature successions of concave bursts. In the former case, our analysis predicts a novel *saturation effect*: Due to informational free-riding, increased opportunities for social learning necessarily lead to temporary slow-downs in learning and do not produce welfare gains.

*We are extremely grateful to Attila Ambrus, Drew Fudenberg, and Eric Maskin for their invaluable advice and unwavering support for this project. Many thanks also to Aislinn Bohren, Thomas Covert, Glenn Ellison, Ben Golub, Marina Halac, Johannes Hörner, Sabrina Howell, Ryota Iijima, Sonia Jaffe, Sangram Kadam, Divya Kirti, Scott Kominers, David Laibson, Greg Lewis, Annie Liang, Jonathan Libgober, Morgan McClellon, Stephen Morris, Joana Naritomi, Arash Nekoei, Gleb Romanyuk, Heather Schofield, Ran Shorrer, Tomasz Strzalecki, Dmitry Taubinsky, Laura Trucco, and Lilei Xu for helpful comments and discussions that improved both the content and exposition of the paper.

†Department of Economics, Harvard University and HBS. Email: mfrick@fas.harvard.edu.

‡Department of Economics, Harvard University. Email: ishii2@fas.harvard.edu, Tel: (703)-623-2039.

1 Introduction

Suppose an innovation of uncertain quality, such as a novel medical treatment or a new piece of software, 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* about the innovation: An individual's treatment success story or discovery of a bug in the software is much more likely 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.

We analyze how the potential for social learning in an economy affects consumers' informational incentives and how these in turn shape the aggregate adoption dynamics of an innovation. Our main results highlight the importance of two features of the economy: The extent to which consumers are *forward-looking* and the nature of news events through which social learning occurs. In choosing whether to adopt an innovation, forward-looking consumers recognize the option value of waiting for more information. With social learning, information is created *endogenously*, based on the consumption experiences of past adopters. In equilibrium, adoption levels must therefore strike a balance: If too many consumers adopt at any given time, then too much information is available in the future and all consumers would rather wait; conversely, if too few consumers adopt, it might not be worthwhile for anyone to wait. We show that the dynamics of this trade-off and the extent to which it is affected by increased opportunities for social learning depend crucially on the kind of information consumers expect to acquire by waiting. In line with numerous empirical findings, our analysis predicts adoption patterns that are either S-shaped or feature successions of concave bursts, suggesting novel micro-foundations for these observations. We also make new predictions regarding the impact of *increased* opportunities for social learning on consumer welfare, on equilibrium learning dynamics, and on observed adoption behavior.

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

Learning 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. Formally, we employ a variation of the Poisson learning models pioneered by Keller et al. (2005), Keller and Rady (2010), and Keller and Rady (2013). As in these models, our analysis distinguishes between *bad news* markets, in which signal arrivals (*breakdowns*) indicate bad quality and the absence of signals makes consumers more optimistic about the innovation; and *good news* markets, in which signal arrivals (*breakthroughs*) suggest good quality and the absence of signals makes consumers more pessimistic. To capture social learning, we assume that the informativeness of signals is increasing in the number of previous adopters.

The automobile industry is an example of a market in which learning is predominantly via bad news events, as evidenced by the wide-spread social media coverage of a battery fire in a Tesla Model S electric car in October 2013 or of the 2009-2011 Toyota vehicle recalls. By contrast, in the market for (essentially side-effect free) herbal remedies or other alternative medical treatments, learning is mostly via good news: Occasional reports of success stories boost consumers' confidence in a treatment, while consumers grow more skeptical of its effectiveness in the absence of any such reports.[1]

The heart of our paper, Sections 5 and 6, 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 (Theorem 4.1): 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[2] 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.

Section 5 studies the perfect bad news case. In the absence of breakdowns, consumers grow increasingly *optimistic* about the innovation over time. As a result of the single-crossing property, the unique equilibrium is then characterized by two times $0 \leq t_1^* \leq t_2^*$, which depend on the fundamentals (Theorem 5.1): 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

[1]Cf. Board and Meyer–ter–Vehn (2013) and MacLeod (2007) for additional examples of bad news and good news markets.

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

standstill. If $t_1^* < t_2^*$, then throughout $[t_1^*, t_2^*)$ only *some* consumers adopt whenever given a chance, with the flow of new adopters 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 (by contrast, if consumers are myopic or if there are no possibilities for social learning, then necessarily $t_1^* = t_2^*$).

We highlight two key implications for aggregate adoption dynamics and consumer welfare:

First, provided $t_1^* < t_2^*$, the innovation's adoption curve (which plots the percentage of adopters in the population against time) has the characteristic S-shaped growth pattern that has been widely observed in empirical studies:[3] Up to time t_1^* adoption is flat, on $[t_1^*, t_2^*)$ adoption levels increase convexly, and from time t_2^* there is a concave increase. Moreover, an increase in the potential for social learning prolongs the period of convex growth and leads to strictly lower expected adoption levels across time. The possibility of S-shaped adoption curves in our model is notable because we assume consumers to be (ex ante) identical, whereas most alternative explanations in the literature rely on specific distributions of consumer heterogeneity to generate a region of convex growth. In our model, convex growth is driven by informational incentives: As consumers grow increasingly optimistic, their opportunity cost to delaying goes up. To maintain indifference between adopting and delaying throughout $[t_1^*, t_2^*)$, 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.

Second, we predict a *saturation effect*: If 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 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 can in fact be used to approximate first-best (complete information) payoffs in the limit. Relative to the cooperative benchmark, equilibrium adoption behavior displays two inefficiencies: First, adoption generally begins too late; second, once adoption begins it initially occurs at an inefficiently low rate, because during $[t_1^*, t_2^*)$ consumers who do not adopt when given a chance effectively free-ride on the information generated by consumers who do adopt. Increased opportunities for social learning exacerbate the second inefficiency by prolonging the period of free-riding. As a result, greater opportunities for social learning do not translate into uniformly faster learning about the quality of the innovation, but rather lead to strictly slower learning

[3] See, for example, Griliches (1957), Mansfield (1961), Mansfield (1968), Davies (1979), and Gort and Klepper (1982), among many others.

over some periods and faster learning over others. These two effects balance out to produce the saturation effect. In Section 7, we further build on this non-monotonicity in the speed of learning to construct an example involving consumers with heterogeneous discount rates, where increased opportunities for social learning are not only not beneficial, but in fact strictly *hurt* aggregate welfare.

In Section 6 we study learning via perfect good news. Here consumers grow increasingly *pessimistic* about the innovation in the absence of breakthroughs. Hence, the single-crossing property for equilibrium incentives implies adoption up to some time t^* (which depends on the fundamentals) and no adoption from t^* on, unless there is a breakthrough, after which all consumers adopt upon their first opportunity (Theorem 6.1). Interestingly, in contrast with the perfect bad news case, 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 information transmission in an economy 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, breakthroughs in the perfect good news environment conclusively reveal the innovation to be good and hence cannot be decision-relevant to a consumer who is already willing to adopt.

The all-or-nothing nature of the good news equilibrium has the following implications for adoption dynamics and welfare:

First, adoption 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. While less commonly observed than S-shaped growth, this pattern is reminiscent of the "fast-break" product life cycles studied in the marketing literature[4], with movies, music, and other "leisure-enhancing" products as canonical examples.[5] We predict that increased opportunities for social learning bring forward t^*, compressing the initial period of concave growth, but do not affect the probability of adoption picking up again after coming to a temporary standstill.

Second, even in economies with rich opportunities for social learning, an increase in the potential for social learning is (essentially) always strictly beneficial and speeds up learning at

[4]Cf. Keillor (2007)

[5]For additional examples in the context of industrial process innovations, see Davies (1979).

all times. Nevertheless, equilibrium behavior is generally socially inefficient: Relative to the cooperative benchmark, adoption takes place at an optimal rate until time t^*, but consumers stop adopting too soon.

1.1 Related Literature and Outline

Our paper proposes a model of innovation adoption by consumers who learn from each other's experiences and are forward-looking. Having a tractable model that can incorporate these two assumptions, examine the informational externalities they give rise to, and derive predictions for the effect of increased opportunities for social learning is desirable, as there is considerable empirical evidence for both assumptions. For example, a growing literature in development economics documents the effect of learning from others' experiences on the adoption of new agricultural technologies, as in Foster and Rosenzweig (1995) or Conley and Udry (2010). This literature also finds evidence for forward-looking behavior: Bandiera and Rasul (2006) analyze the decision of farmers in Mozambique to adopt a new crop, sunflower. They find that farmers whose network of friends and family contains *many* adopters of the new crop are *less* likely to initially adopt it themselves. Relatedly, 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.

At a theoretical level, the key feature of our model is that social learning and forward-looking incentives jointly give rise to informational externalities that do not arise in the absence of either assumption. In relation to existing models of innovation adoption, this has at least two interesting implications.

First, many models of innovation adoption rely on consumer heterogeneity as a key ingredient in fitting observed adoption data. Our analysis suggests that in existing *learning-based* models[6] heterogeneity is only crucial because of the common assumption that either consumers are forward-looking but news is generated purely exogenously, as in Jensen (1982), or that learning is social but consumers are myopic, as in Young (2009) or Ellison and Fudenberg (1993):[7] In

[6]For comprehensive surveys of the literature, including also non-learning based explanations of innovation adoption, such as the epidemic model and the probit model of firm characteristics, see for example Geroski (2000) and Baptista (1999).

[7]Two exceptions are Persons and Warther (1997) and Kapur (1995), who feature a form of forward-looking social learning, but differ substantially from our paper in terms of both setup and focus. Persons and Warther (1997) focuses on the combination of forward-looking incentives, endogenously generated news, and firm heterogeneity to provide rational foundations for seemingly irrational, fad-like patterns in the adoption of financial innovations. In Kapur (1995), a finite number of firms engage in a sequence of waiting contests to adopt a

either case, a population of identical consumers would behave according to a simple cutoff rule, adopting the innovation at beliefs above a certain threshold and not adopting otherwise, and this rules out convex growth in adoption levels.[8] By contrast, in our model consumers are assumed to be ex ante identical, but the combination of forward-looking behavior and social learning allows us to provide an alternative micro-foundation for convex growth in terms of purely informational incentives.

The literature also commonly appeals to variations in consumer heterogeneity in order to explain qualitative differences in adoption patterns across different products. For example, in his study of the diffusion of 22 post-war industrial process innovations in the UK, Davies (1979) uses symmetrical logistic distributions to fit the S-shaped adoption patterns characteristic of expensive and complex innovations, but lognormal distributions to fit the rapid, essentially concave growth in adoption levels he observes for less expensive and simpler innovations. Again, our analysis shows that when consumers are forward-looking social learners, these contrasting patterns can instead be explained through differences in the informational environment: S-shaped curves arise in bad news markets with a relatively large potential for social learning, while concave adoption patterns are characteristic of good news markets (or of bad news markets with little potential for social learning or with very optimistic consumers). Our focus on the role of the market learning process in shaping consumers' informational incentives and generating varied aggregate adoption dynamics is similar in spirit to Board and Meyer–ter–Vehn (2013), who in the context of a capital-theoretic model of quality and reputation, highlight the dependence of firms' reputational incentives on the news process and contrast reputational dynamics across different markets.

Second, in addition to providing an alternative explanation for observed data, the informational externalities that arise from the interaction between forward-looking behavior and endogenously generated information are important because they suggest caution in evaluating the effect of increased opportunities for social learning. In contrast to existing models, we predict that increased opportunities for social learning need not produce welfare gains and may lead

new technology, with each contest ending once a firm adopts. Restricting to MPE, he finds that if more information is revealed when more firms adopt during a given waiting contest, then the mean duration of waiting contests shrinks over time, suggesting a crude approximation of convex diffusion. Since both models are set in discrete time, they are less tractable and not suited to performing comparative statics analyses with respect to the potential for social learning in an economy. In addition, discrete time is less suited to highlighting the role of the market learning process in shaping aggregate adoption dynamics, because when the information process is sufficiently informative relative to the period length, adoption behavior is qualitatively similar across many news processes. By contrast, when the period length becomes short as in our continuous time model, differences become transparent.

[8] Adoption patterns can exhibit concave growth simply as a result of gradual depletion of the population of remaining consumers.

to a temporary slowdown in learning and a strict decline in initial adoption levels. On the other hand, if learning is modeled as purely exogenous or consumers are assumed to be myopic, then increased opportunities for social learning necessarily speed up learning and are unambiguously welfare-improving.

The techniques and framework of this paper are closest to those employed in the strategic experimentation literature, e.g. Bolton and Harris (1999), Keller et al. (2005), Keller and Rady (2010), and Keller and Rady (2013). However, our paper differs in two key respects: First, in our model any individual consumer's influence on the information seen by others is negligible; second, adoption of the innovation is irreversible. The first assumption is natural in the context of the large market applications we have in mind, and for many new products (for example movies or books, for which consumption is usually a one-time event, or technologies that entail large switching costs) irreversibility is also more reasonable than the possibility of consumers continuously switching back and forth between the innovation and the status quo as in the strategic experimentation literature.[9] In the strategic experimentation literature, consumers' direct influence on opponents' information and their ability to adjust their experimentation levels as a function of beliefs produces the so-called *encouragement effect*: There is an incentive to increase current experimentation in order to drive up beliefs and induce more future experimentation by others.[10] As a result of the encouragement effect, many comparative statics in those models differ substantially: For example, an increase in the rate of information transmission may cause consumers to begin to adopt earlier, whereas in our model, we observe that initially adoption rates always weakly decrease in response to such a change. Without the encouragement effect, we are more easily able to study comparative statics on adoption behavior, speed of learning, and welfare with respect to changes in the social learning environment. Moreover, we obtain equilibrium uniqueness (at the aggregate level) without any Markovian restriction on strategies.

A number of papers, including Rosenberg et al. (2007), Chamley and Gale (1994), and Murto and Välimäki (2011), also study the impact of informational externalities on adoption, investment, or exit behavior, but rely on the assumption that agents hold *private* information. Notably, Chamley and Gale (1994) obtain a result somewhat resembling our saturation effect, according to which in the limit, an increase in the number of players has no effect on the rate of investment or flow of information. In the context of a two-armed bandit problem in which

[9]Moreover, if consumers could continuously switch back and forth between the two options, then under the large market assumption, consumers' equilibrium strategies would effectively reduce to myopic best response with respect to beliefs.

[10]There is no encouragement effect in the perfect good news environment of Keller et al. (2005), but consumers' ability to influence each other's beliefs as well as the reversibility of experimentation are once again crucial in generating asymmetric switching equilibria, in which consumers take turns in experimenting at different beliefs.

the decision to switch to the *safe* arm is irreversible, Rosenberg et al. (2007) obtain a similar uniqueness result to ours in the limit as the number of players becomes large. However, the specifics of all these models differ substantially from ours, as agents obtain private information and make inferences about the quality of the product by observing others' *actions*, while in our model all relevant news is *public* and actions do not reveal additional information.

Finally, Bergemann and Välimäki (1997) and Bergemann and Välimäki (2000) study innovation adoption in the presence of pricing motives by sellers when learning is social. In these papers, prices that dynamically adjust through time act as an additional instrument through which the seller can affect the endogenous information generation process. Bergemann and Välimäki (1997) study a model in which one established firm (with known technology) and a new firm with a risky innovation compete through prices. They derive the Markov perfect equilibrium pricing strategies and adoption behavior and demonstrate that adoption is too fast (relative to the social optimum) when consumers are pessimistic and too slow when consumers are optimistic. The main difference with our paper is that consumers in their model best respond myopically at each point in time, so that adoption dynamics are driven purely by sellers' informational and pricing motives. By contrast, in our model consumers are more sophisticated and consider the option value to waiting, producing interesting adoption dynamics even in the absence of pricing motives.[11] Bergemann and Välimäki (2000) analyze a similar model in which consumers display forward-looking behavior. As in Bergemann and Välimäki (1997), they find that pricing motives cause experimentation to be excessive, which is in contrast to our finding that in the absence of pricing motives there is too little (and, under perfect bad news, too slow) adoption. They find additionally that when the innovation is launched in many markets simultaneously, adoption rates become socially optimal in the limit as the number of markets grows large. Much of the focus in our paper is on analyzing the effect of increased opportunities for social learning on consumers' informational incentives. In order to isolate the effect on the consumer side, our baseline model therefore abstracts away from pricing considerations.

The rest of the paper is organized as follows. Section 2 describes the model, defining formally the perfect bad news and perfect good news signal processes that we use throughout the paper as well as the equilibrium concept. Section 3 analyzes the cooperative (socially optimal) benchmark which selects an aggregate flow of adoption so as to maximize ex ante aggregate welfare. Section 4 establishes a quasi-single crossing property for equilibrium incentives that simplifies the equilibrium analysis in the following sections. Section 5 establishes existence of a unique

[11]The key distinction is again due to the assumption that adoption is irreversible in our model, so that potentially adopting a bad product incurs a cost on consumers. On the other hand, in Bergemann and Välimäki (1997), consumers adopt at every point in time and the adoption decision is freely reversible.

equilibrium under perfect bad news and studies comparative statics with respect to changes in the potential for social learning. Section 6 performs the analogous exercise under perfect good news. Section 7 provides an example, involving consumers with heterogeneous discount rates, where an increase in the potential for social learning strictly hurts ex ante welfare. Section 8 concludes. Appendix A - I contains proofs omitted from the main text.

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 $\bar{N}_0 \in \mathbb{R}_+$. 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 $\rho > 0$.[12] Upon an adoption opportunity, a consumer must choose whether to adopt the innovation ($a_t = 1$) or to wait ($a_t = 0$). If a consumer adopts, he receives an expected lump sum payoff of $\mathbb{E}_t[\theta]$, conditioned on information available up to time t, and drops out of the game. If the consumer chooses to wait or does not receive an adoption opportunity at t, he receives a flow payoff of 0 until his next adoption opportunity, where he faces the same decision again.

2.2 Learning

Over time, consumers observe public signals that convey information about the quality of the innovation. To capture the idea of social learning, the informativeness of the public signal at time t is increasing in the flow N_t of consumers newly adopting the innovation at t, which we define more precisely in Section 2.3.

Formally, we employ a variation of the Poisson learning model pioneered by Keller et al. (2005), Keller and Rady (2010), and Keller and Rady (2013).[13] Conditional on quality θ,

[12]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.

[13]Keller et al. (2005) have learning via perfect good news Poisson signals, Keller and Rady (2010) study imperfect good news learning, and Keller and Rady (2013) study perfect and imperfect bad news learning. For

public signals arrive according to an inhomogeneous Poisson process with arrival rate $(\varepsilon_\theta + \lambda_\theta N_t)dt$, where $\lambda_\theta > 0$ and $\varepsilon_\theta \geq 0$ are exogenous parameters that depend on the quality θ of the innovation. The signal process summarizes news events that are generated from two sources. First, the social learning term λN_t represents news generated endogenously, based on the experiences of other consumers: It captures the idea of a flow N_t of new adopters each generating signals at rate $\lambda\, dt$.[14] 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 $\varepsilon\, dt$, which represents information generated independently of consumers' behavior, for example by professional critics or government watchdog agencies.

For tractability, we focus on learning via *perfect* Poisson processes, where a single signal provides *conclusive* evidence of the quality of the innovation. Learning is via *perfect bad news* if $\varepsilon_G = \lambda_G = 0$ 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. Learning is via *perfect good news* if $\varepsilon_B = \lambda_B = 0$ and $\varepsilon_G = \varepsilon \geq 0$, $\lambda_G = \lambda > 0$, so that a signal arrival (called a *breakthrough*) is conclusive evidence for the innovation being good. As motivated in the Introduction, the distinction between bad news and good news can be seen to reflect the nature of news production in different markets. In addition, $\Lambda_0 := \lambda \bar{N}_0$ can be seen as a simple measure of the *potential for social learning* in an economy, summarizing both the likelihood λ with which individual adopters' experiences find their way into the public domain and the size $\bar{N}_0$ of the population which can contribute to and access the common pool of information.

We briefly summarize the evolution of consumers' beliefs under bad and good news:

2.2.1 Learning via Perfect Bad News

Under perfect bad news, consumers' posterior on $\theta = G$ permanently jumps to 0 at the first breakdown. 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, standard Bayesian

---

other recent work that prominently features learning via Poisson signals, see for example Che and Hörner (2013); Board and Meyer–ter–Vehn (2013); Halac et al. (2013).

[14] Note that by letting the social learning component of the signal arrival rate at time t, λN_t, depend only on the flow of adopters N_t at time t itself, we are effectively assuming that each each adopter can generate a signal only once, namely at the time of adoption. This assumption is natural 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, such as cars or consumer electronics, it might be more natural to allow adopters to generate signals repeatedly over time, which can be captured by replacing λN_t with $\lambda \int_0^t N_s\, ds$. This would yield results that are qualitatively similar to those presented in the following sections.

updating implies that

$$p_t = \frac{p_0}{p_0 + (1-p_0)e^{-\int_0^t (\varepsilon + \lambda N_s) ds}}.^{15} \tag{1}$$

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

$$\dot{p}_\tau = (\varepsilon + \lambda N_\tau)\, p_\tau (1 - p_\tau).$$

Note that the no-news posterior is continuous and *increasing*.

2.2.2 Learning via Perfect Good News

Under perfect good news, consumers' posterior on $\theta = G$ permanently jumps to 1 at the first breakthrough. Given a flow of adopters N, Bayes' rule now implies that consumers' no-news posterior satisfies

$$p_t = \frac{p_0 e^{-\int_0^t (\varepsilon + \lambda N_s) ds}}{p_0 e^{-\int_0^t (\varepsilon + \lambda N_s) ds} + (1-p_0)}. \tag{2}$$

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

$$\dot{p}_\tau = -\,(\varepsilon + \lambda N_\tau)\, p_\tau (1 - p_\tau).$$

In contrast to the perfect bad news case, the no-news posterior is now continuous and *decreasing*.

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 $(N_t)_{t \geq 0}$ 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 will adopt after a breakdown, while under perfect good news all remaining consumers will adopt when given a chance 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:

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

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 ρ.

Any feasible adoption process N defines an associated no-news posterior p_t^N as given by Equation 1 if learning is via perfect bad news and by Equation 2 if learning is via perfect good news.

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: If a consumer receives an adoption opportunity at t, he 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. For this we must first define the value to waiting at t.

Let Σ_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 τ^σ at which, absent signals, the consumer will adopt the innovation and drop out of the game.[16]

Let $W_t^N(\sigma)$ denote the expected payoff to waiting at t and following σ in the future, given the aggregate adoption process N. Specifically, if learning is via perfect bad news, σ prescribes adoption at the random time τ^σ if and only if there have been no breakdowns prior to τ^σ, yielding

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

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

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

[16] Formally, we define τ^σ as follows. Let $(X_s)_{s\geq t}$ denote the stochastic process representing the number of arrivals generated on $[t, s]$ by a Poisson process with arrival rate ρ, and let $(X_{s^-})_{s>t}$ denote the number of arrivals on $[t, s)$. Then,

$$\tau^\sigma := \inf\{s \geq t : \sigma_s \times (X_s - X_{s^-}) > 0\},$$

where, as per convention, $\inf \varnothing := +\infty$. 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.

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

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

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

Definition 2.2. The value to waiting given a feasible adoption process N is the function $W_t^N : \mathbb{R}_+ \to \mathbb{R}_+$ defined by $W_t^N := \sup_{\sigma\in\Sigma_t} W_t^N(\sigma)$ for all t.

We are now ready to define our equilibrium concept:

Definition 2.3. An equilibrium is a feasible adoption process $(N_t)_{t\geq 0}$ such that

(i). $W_t^N \geq 2p_t^N - 1$ for all t such that $\rho\bar{N}_t > N_t$

(ii). $W_t^N \leq 2p_t^N - 1$ for all t such that $0 < N_t$.

Thus, Definition 2.3 requires that at any time t, the aggregate flow of new adopters N_t be consistent with the remaining $\bar{N}_t$ consumers optimally trading off the expected payoff to immediate adoption, $2p_t^N - 1$, against the value to waiting, W_t^N.

Note that our definition of equilibrium is essentially Nash equilibrium, i.e. we do not require subgame perfection. The motivation for this is that in a continuum population any individual consumer's behavior has a negligible impact on the aggregate adoption levels so that any history not on the equilibrium path (in which a different number of consumers than expected previously adopted) 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.

As usual, the equilibrium value to waiting W_t^N admits an alternative characterization as the solution to a functional equation, which we note here for use in future sections:

Lemma 2.4. *Suppose N is an equilibrium. If learning is via perfect bad news, W_t^N satisfies the functional equation*

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

If learning is via perfect good news, W_t^N satisfies the functional equation

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

Proof. The proof is standard. ■

3 Cooperative Benchmark

To establish a socially optimal benchmark, we first consider the cooperative problem: This selects an aggregate flow N of adopters that maximizes ex ante aggregate welfare, taking into account the effect of N on the public information process; we impose feasibility, but do not impose the incentive compatibility requirements of the equilibrium in Definition 2.3.[17]

Clearly, under perfect good news it is optimal to require adoption at the maximal possible rate once there has been a breakthrough. Similarly, under perfect bad news it is optimal to terminate adoption as soon as there has been a breakdown. Thus, the objective of the cooperative problem under perfect good news is:

$$\sup_N p_0 \int_0^\infty \left(\varepsilon_G + \lambda_G N_\tau\right) e^{-\int_0^\tau (\varepsilon_G + \lambda_G N_s) ds} \left(\int_0^\tau e^{-rs} N_s ds + e^{-r\tau} \frac{\rho}{\rho + r} \left(\bar{N}_0 - \int_0^\tau N_s ds \right) \right) d\tau$$
$$+ p_0 e^{-\int_0^\infty (\varepsilon_G + \lambda_G N_s) ds} \int_0^\infty e^{-rs} N_s\, ds - (1 - p_0) \int_0^\infty e^{-rs} N_s\, ds, ^{18}$$

subject to the feasibility constraint that $N_t \in [0, \rho\bar{N}_t]$ for all t.

Under perfect bad news, the objective is:

$$\sup_N p_0 \int_0^\infty e^{-rs} N_s ds - (1 - p_0) \int_0^\infty \left(\varepsilon_B + \lambda_B N_\tau\right) e^{-\int_0^\tau (\varepsilon_B + \lambda_B N_s) ds} \int_0^\tau e^{-rs} N_s\, ds\, d\tau$$
$$- (1 - p_0) e^{-\int_0^\infty (\varepsilon_B + \lambda_B N_s) ds} \int_0^\infty e^{-rs} N_s ds, ^{19}$$

[17] We are not concerned with implementation here, but because beliefs are publicly observed, as long as we allow for transfers, the solution that we provide will be implementable while respecting individual rationality.

[18] We impose the convention that $e^{-\infty} = 0$. Thus whenever $\varepsilon_G > 0$, $e^{-\int_0^\infty (\varepsilon_G + \lambda_G N_s) ds} = 0$.

again subject to the feasibility constraint that $N_t \in [0, \rho \bar{N}_t]$ for all t.

Standard techniques show that the solution to both cooperative problems has an *all-or-nothing* form:[20] In each problem, there is a cutoff time t^s (depending on the parameters) such that conditional on no signals, there is no (respectively maximal) adoption until time t^s under perfect bad (respectively good) news, and maximal (respectively no) adoption from t^s on:

Proposition 3.1. *In both problems, there exists an adoption flow N that attains the maximum. Furthermore, there exists an optimal adoption flow with the property that there exists t^s such that*

- $N_t = \rho \bar{N}_t$ *for all t such that* $(\lambda_G - \lambda_B)(t^s - t) > 0$;
- $N_t = 0$ *for all t such that* $(\lambda_G - \lambda_B)(t^s - t) < 0$.

Proof. See Appendix Section I. ■

We now solve for the cutoff time, or equivalently the cutoff belief, under both signal structures.

3.1 Cooperative Benchmark under Perfect Good News

Under perfect good news, letting $\varepsilon := \varepsilon_G$ and $\Lambda_0 := \lambda_G \bar{N}_0$, the cutoff time t^s solves

$$\sup_{t^s \geq 0} \frac{\rho}{r+\rho}\left(1 - e^{-(r+\rho)t^s}\right)\bar{N}_0(2p_0 - 1) + e^{-(r+\rho)t^s}\frac{\rho}{r+\rho}\bar{N}_0\left(\pi^s + (1-\pi^s)p^s\frac{\varepsilon}{\varepsilon + r}\right) \tag{3}$$

where π^s and p^s denote, respectively, the probability of a breakthrough prior to time t^s and the no-news posterior at time t^s; that is,

$$\pi^s := p_0\left(1 - e^{-\varepsilon t^s - \Lambda_0(1-e^{-\rho t^s})}\right),$$
$$p^s := \frac{p_0 e^{-\varepsilon t^s - \Lambda_0(1-e^{-\rho t^s})}}{p_0 e^{-\varepsilon t^s - \Lambda_0(1-e^{-\rho t^s})} + (1-p_0)}.$$

Taking the first order condition of the above, we obtain:

$$(r+\rho)(1-\pi^s)\left(\left(2 - \frac{\varepsilon}{\varepsilon + r}\right)p^s - 1\right) + p^s(1-\pi^s)\left(\varepsilon + \Lambda_0 \rho e^{-\rho t^s}\right)\frac{r}{\varepsilon + r} = 0 \tag{4}$$

[19] Again we assume that whenever $\varepsilon_B > 0$, $e^{-\int_0^\infty (\varepsilon_B + \lambda_B N_s)\, ds} = 0$.

[20] This is due to the linearity of the signal arrival rate in N_t.

if an interior solution exists.

If the left-hand side of Equation 4 is non-positive at all times, then the cooperative cutoff satisfies $t^s = 0$, so that there is no adoption until a breakthrough. This happens if and only if

$$(r+\rho)\left((2p_0-1) - p_0\frac{\varepsilon}{\varepsilon+r}\right) + p_0\left(\varepsilon+\rho\Lambda_0\right)\frac{r}{\varepsilon+r} \leq 0. \tag{5}$$

On the other hand, if the left-hand side of Equation 4 is strictly positive at all times, then $t^s = +\infty$ and the cooperative solution calls for maximal adoption irrespective of whether or not there has been a breakthrough. This happens if and only if $\varepsilon = 0$ and $p_0\left(1+e^{-\Lambda_0}\right) \geq 1$.

We summarize this in the following proposition:

Proposition 3.2. *Under perfect good news, the cooperative cutoff time is as follows:*

- *If Inequality (5) holds, then* $t^s = 0$.
- *If* $\varepsilon = 0$ *and* $p_0\left(1+e^{-\Lambda_0}\right) \geq 1$, *then* $t^s = +\infty$.
- *Otherwise,* t^s *satisfies Equation (4).*

Note that the cutoff posterior p^s depends on the prior. This is in contrast to the strategic experimentation literature because of our assumption that the stock of remaining consumers is depleted as consumers drop out following adoption. In strategic experimentation, the cooperative solution only depends on the current belief and does not depend on the initial conditions since experimenters remain in the game to potentially experiment further in the future.

3.2 Cooperative Benchmark under Perfect Bad News

Under perfect bad news, letting $\varepsilon := \varepsilon_B$ and $\Lambda_0 := \lambda_B \bar{N}_0$, the cutoff time t^s solves:

$$\sup_{t^s \geq 0} e^{-rt^s}\bar{N}_0\left(p_0\frac{\rho}{\rho+r} - (1-p_0)e^{-\varepsilon t^s}\int_0^\infty \rho e^{-\varepsilon\tau - \Lambda_0(1-e^{-\rho\tau})}e^{-(r+\rho)\tau}d\tau\right).$$

Taking the first order condition, we obtain:

$$e^{-\varepsilon t^s}K(\Lambda_0) = \frac{r}{\varepsilon+r}\frac{\rho}{r+\rho}\frac{p_0}{1-p_0}$$

where

$$K(\Lambda_0) := \int_0^\infty \rho e^{-\varepsilon\tau-\Lambda_0(1-e^{-\rho\tau})}e^{-(r+\rho)\tau}d\tau < \frac{\rho}{\varepsilon+\rho+r}.$$

Then an easy calculation yields the cutoff posterior:

$$p^s = \frac{K(\Lambda_0)}{\frac{r}{\varepsilon+r}\frac{\rho}{\rho+r} + K(\Lambda_0)} < \frac{\varepsilon + r}{\varepsilon + 2r}.$$

We summarize this in the following proposition:

Proposition 3.3. *Under perfect bad news, the cooperative solution is given by:*

$$N_t = \begin{cases} 0 & \text{if } p_t < p^s \\ \rho \bar{N}_t & \text{if } p_t \geq p^s, \end{cases}$$

where

$$p^s = \frac{K(\Lambda_0)}{\frac{r}{\varepsilon+r}\frac{\rho}{\rho+r} + K(\Lambda_0)}.$$

4 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 $p^N_{t\geq 0}$ and value to waiting $W^N_{t\geq 0}$ as defined in Definition 2.2. In general, the dynamics of the trade-off between immediate adoption at time t (yielding expected payoff $2p^N_t - 1$) and delaying and behaving optimally in the future (yielding expected payoff W^N_t) can be quite difficult to characterize, with $(2p^N_t - 1) - W^N_t$ changing sign many times. However, when $N_{t\geq 0}$ is an *equilibrium* flow, then for any t,

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

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

Theorem 4.1. *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 $p^N_{t\geq 0}$ and value to waiting $W^N_{t\geq 0}$. Then $W^N_{t\geq 0}$ and $2p^N_{t\geq 0} - 1$ satisfy single-crossing, in the following sense:*

- *Whenever $(\lambda_B - \lambda_G)(W^N_t - (2p^N_t - 1)) < 0$, then $(\lambda_B - \lambda_G)(W^N_\tau - (2p^N_\tau - 1)) < 0$ for all $\tau > t$.*

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

Proof. See Appendix Section B. ■

The basic intuition is as follows. Consider first the case of learning via perfect bad news and suppose that immediate adoption is strictly better than waiting today (and hence also in the near future provided there are no breakdowns).[21] Then all consumers adopt upon an opportunity in the near future, 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 in the future.

Similarly, suppose that learning is via perfect good news and that waiting is strictly more attractive than immediate adoption today (and hence also in the near future). Then in the near future, no consumers adopt and information is generated purely via the exogenous news source (or not at all if $\varepsilon = 0$). As a result, the no-news posterior decreases (weakly) while the number of remaining consumers does not change. This makes waiting even more attractive relative to adopting immediately, so that waiting continues to be strictly preferable in the future.

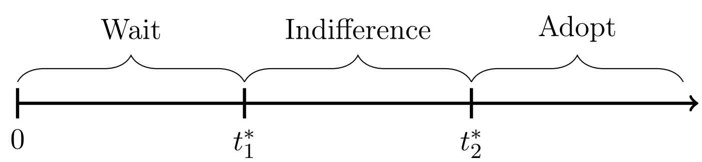

Figure 1: Perfect Bad News

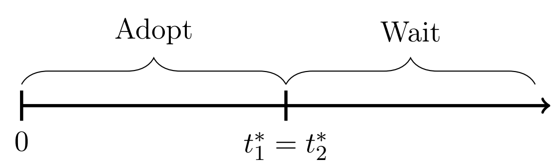

Figure 2: Perfect Good News

Theorem 4.1 implies that any equilibrium features two threshold times $0 \leq t_1^* \leq t_2^* \leq +\infty$

[21]The latter implication follows from the continuity of the equilibrium value to waiting, which is established in the Appendix.

given by[22]

$$t_1^* := \inf\{t : (\lambda_B - \lambda_G)\left(2p_t^N - 1 - W_t^N\right) \geq 0\},$$
$$t_2^* := \inf\{t : (\lambda_B - \lambda_G)\left(2p_t^N - 1 - W_t^N\right) > 0\},$$

such that if there are no signal arrivals, then under perfect bad (respectively good) news, waiting (respectively adoption) is strictly preferable before t_1^*, and adoption (respectively waiting) is strictly preferable after t_2^*, with indifference in between, as illustrated in Figures 1 and 2. In Sections 5 and 6 we will build on this observation to establish the existence of unique equilibria under both perfect bad news and good news. The threshold 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 6, we will see that under perfect good news, any equilibrium must in fact satisfy $t_1^* = t_2^*$.[23] Depending on parameters, the equilibrium takes three possible forms: (i) $0 = t_1^* = t_2^*$; (ii) $0 < t_1^* = t_2^* < +\infty$; or (iii) $0 < t_1^* = t_2^* = +\infty$.[24] By contrast, under perfect bad news in Section 5, the equilibrium takes one of six forms depending on parameters: (i) $0 = t_1^* = t_2^* < +\infty$; (ii) $0 = t_1^* < t_2^* < +\infty$; (iii) $0 < t_1^* = t_2^* < +\infty$; (iv) $0 < t_1^* < t_2^* < +\infty$; (v) $0 < t_1^* = t_2^* = +\infty$;[25] or (vi) $0 = t_1^* < t_2^* = +\infty$.[26] The possibility of a non-empty interval (t_1^*, t_2^*) of indifference will emerge as a key feature distinguishing bad news markets from good news markets. Maintaining indifference at times (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 not just from an efficiency standpoint, but also for the shape of equilibrium adoption curves and for the impact of increased opportunities for social learning on welfare, learning, and adoption dynamics.

5 Perfect Bad News

5.1 Equilibrium Characterization

We now build on the analysis of the previous section to establish the existence of a unique equilibrium when learning is via perfect bad news. Fix parameters r, ρ, $\bar{N}_0 > 0$, $\varepsilon = \varepsilon_B$, $\lambda = \lambda_B \geq 0$, and $p_0 \in (0,1)$. Suppose $N_{t\geq 0}$ is an equilibrium flow of adopters. Let $p_{t\geq 0}$

[22] With the usual convention that $\inf \varnothing = +\infty$.

[23] With the sole exception of $\varepsilon = 0$ and $p_0 = \frac{1}{2}$, in which case it is easy to see that $N \equiv 0$ and $t_1^* = 0 < t_2^* = \infty$.

[24] This possibility will arise iff $\varepsilon = 0$ and $p_0\left(1 + e^{-\lambda\bar{N}_0}\right) \geq 1$.

[25] This possibility will arise iff $\varepsilon = 0$ and $p_0 < \frac{1}{2}$.

[26] This possibility will arise iff $\varepsilon = 0$ and $p_0 = \frac{1}{2}$.

and $W_{t\geq 0}$ be the corresponding no-news posterior and value to waiting, and let $\Lambda_{t\geq 0} := \lambda\bar{N}_{t\geq 0}$ describe the evolution of the economy's potential for social learning.[27] From Theorem 4.1, we know that there are times $0 \leq t_1^* \leq t_2^* \leq +\infty$ given by

$$t_1^* := \inf\{t : 2p_t - 1 \geq W_t\},$$
$$t_2^* := \inf\{t : 2p_t - 1 > W_t\},$$

such that (appealing also to right-continuity) N must satisfy

$$\begin{cases} N_t = 0 & \text{if } t < t_1^*, \\ 2p_t - 1 = W_t & \text{if } t \in [t_1^*, t_2^*) \\ N_t = \rho\bar{N}_t & \text{if } t \geq t_2^*. \end{cases}$$

In the following we will show that t_1^*, t_2^*, and the evolution of N_t between t_1^* and t_2^* are uniquely pinned down by the parameters. We first introduce some 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 if there have been no breakdowns by then, *given* that the current belief is p, that the remaining potential for social learning is Λ, and that absent breakdowns the remaining $\frac{\Lambda}{\lambda}$ consumers adopt at their first opportunity in the future.

Define the posteriors $\underline{p}$, $\overline{p}$, and $p^\sharp$ as follows. Let $\underline{p}$ be the posterior given by $2\underline{p}-1 = G(\underline{p}, 0)$; that is,

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

Thus, $\underline{p}$ is the lowest belief at which a consumer is willing to adopt given that he could also delay, obtain more information at rate ε and reevaluate his decision at his next adoption opportunity which is generated at rate ρ.

Define $\overline{p} := \lim_{\rho\to\infty} \underline{p}$, that is,

$$\overline{p} = \frac{\varepsilon + r}{\varepsilon + 2r};$$

$\overline{p}$ is the lowest belief at which a consumer would be willing to adopt given that he could also delay and obtain more information at rate ε and given that adoption opportunities arrive continuously

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

in the future.

Define $p^\sharp := \lim_{\varepsilon\to\infty} \underline{p}$, that is,

$$p^\sharp = \frac{\rho + r}{\rho + 2r}.$$

$p^\sharp$ is the lowest belief at which a consumer would be willing to adopt given that he could also delay until his next opportunity, which is generated at rate ρ, and given that all uncertainty is completely resolved by then.[28]

Finally, define the function $\Lambda^* : (0,1) \to \mathbb{R}_+ \cup \{+\infty\}$ as follows. Let $\Lambda^*(p) \equiv 0$ for all $p \le \underline{p}$, $\Lambda^*(p) = +\infty$ for all $p \ge p^\sharp$, and for all $p \in (\underline{p}, p^\sharp)$, let $\Lambda^*(p) \in \mathbb{R}_+$ be the unique value such that

$$2p - 1 = G(p, \Lambda^*(p)).^{29}$$

Thus, if the current posterior is $p \in [\underline{p}, p^\sharp)$ and the current potential for social learning in the economy is $\Lambda^*(p)$, then consumers are indifferent between adopting now or at their next opportunity absent breakdowns, provided that all remaining $\frac{\Lambda^*(p)}{\lambda}$ consumers also adopt at their first opportunity in the future.

We are now ready to state the equilibrium characterization theorem:

Theorem 5.1. *Fix $r, \rho > 0$, $\varepsilon \ge 0$, and $p_0 \in (0,1)$. Let $p^* := \min\{\overline{p}, p^\sharp\}$. For every $\lambda, \bar{N}_0 > 0$, there is a unique equilibrium. Furthermore, in the unique equilibrium, N_t is Markovian in (p_t, Λ_t) for all t and satisfies*

$$N_t = \begin{cases} 0 & \text{if } p_t \le p^* \text{ and } \Lambda_t > \Lambda^*(p_t), \\ \frac{r(2p_t - 1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda} \in (0, \rho\bar{N}_t) & \text{if } p_t > p^* \text{ and } \Lambda_t > \Lambda^*(p_t) \\ \rho\bar{N}_t & \text{if } \Lambda_t \le \Lambda^*(p_t). \end{cases} \tag{6}$$

A detailed proof of Theorem 5.1 is provided in Appendix Section C.1. Here we sketch the basic idea. Before we proceed, however, note the following two special cases of the theorem: First, if $\rho \le \varepsilon$, so that $p^* := \min\{\overline{p}, p^\sharp\} = p^\sharp$, then by Equation (6) and because $\Lambda^*(p) = +\infty$ for all $p \ge p^\sharp$, Theorem 5.1 asserts that regardless of the other parameters, N_t takes an all-or-nothing form with cutoff belief $p^\sharp$: $N_t = 0$ whenever $p_t < p^\sharp$ and $N_t = \rho\bar{N}_t$ whenever $p_t \ge p^\sharp$. Second, if $\varepsilon = 0$ and $p_0 \le \frac{1}{2}$, then it is easy to see that Theorem 5.1 asserts that regardless of the other parameters, the unique equilibrium is given by $N_t = 0$ for all t.

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

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

Throughout Section 5, we will be particularly interested in the implications of N featuring a *partial adoption* region, in which $N_t \in (0, \rho\bar{N}_t)$ is as described by the second line of Equation (6). Since the two special cases above preclude the existence of such a region regardless of other parameters, we rule out these cases for the remainder of Section 5 by imposing the following two conditions:[30]

Condition 5.2. The rate at which exogenous information arrives is small relative to the rate at which consumers obtain adoption opportunities: $\varepsilon < \rho$. Thus, $p^* = \overline{p} < p^\sharp$.

Condition 5.3. Either $\varepsilon > 0$ or $p_0 \in (\frac{1}{2}, 1)$.

Given these two conditions, we now sketch the derivation of Theorem 5.1. In order to obtain the Markovian description of N_t in Equation (6), we note the following lemma, which we prove in the Appendix. This provides an alternative characterization of the threshold times t_1^* and t_2^*, relating these times to the evolution of (p_t, Λ_t):

Lemma 5.4. *Fix $r, \rho > 0$, $\varepsilon \geq 0$ and $p_0 \in (0,1)$ satisfying Conditions 5.2 and 5.3. Let $N_{t\geq 0}$ be an equilibrium with corresponding no-news posterior $p_{t\geq 0}$ and threshold times t_1^* and t_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). $t_2^* = \inf\{t : \Lambda_t < \Lambda^*(p_t)\}$*; and*

(ii). $t_1^* = \min\{t_2^*, \sup\{t : p_t < \overline{p}\}\}$.[31]

Proof. See Appendix Section C.1.2. ■

By Lemma 5.4 the first line of Equation (6) corresponds to times $t \leq t_1^*$, the second line to $t \in (t_1^*, t_2^*)$, and the third line to $t \geq t_2^*$. Thus, the first and third lines are immediate from the definition of these threshold times. We now give a heuristic argument outlining the derivation of the second line, i.e. the equilibrium flow of adoption at times $t \in (t_1^*, t_2^*)$, where adoption is partial. At all these times, consumers must be exactly indifferent between adopting today and waiting for more information. Maintaining consumer indifference at these times requires that the cost and benefit of delaying be equal:

$$\overbrace{\underbrace{(\varepsilon + \lambda N_t)(1-p_t)dt}_{\substack{\text{Probability of} \\ \text{breakdown}}} \quad \underbrace{(0-(-1))}_{\substack{\text{Benefit:} \\ \text{Avoid Bad Product}}}}^{\text{Benefit of Delay}} = \overbrace{\underbrace{(1-(\varepsilon + \lambda N_t)(1-p_t)dt)}_{\substack{\text{Probability of} \\ \text{no breakdown}}} \underbrace{(2p_{t+dt}-1)rdt}_{\substack{\text{Cost:} \\ \text{Discounting}}}}^{\text{Cost of Delay}}. \tag{7}$$

[30]In Section D in the Appendix, we discuss in more detail the case where $\rho \leq \varepsilon$.
[31]With the convention that if $\{t \geq 0 : p_t < \overline{p}\} = \varnothing$, then $\sup\{t : p_t < \overline{p}\} = 0$.

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 $(0-(-1))$, 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.[32] Ignoring terms of order dt^2 and rearranging yields $N_t = \frac{r(2p_t-1)}{\lambda(1-p_t)} - \frac{\varepsilon}{\lambda}$, as in Equation (6).[33]

Finally, Figure 3 illustrates how from Equation 6, we obtain a unique equilibrium as a function of the parameters. Regions (2) and (3) represent values of (p_t, Λ_t) corresponding to the first line of Equation (6), so that no adoption takes place in these regions. Region (4) corresponds to partial adoption as given by the second line of Equation (6). Finally, region (1) corresponds to the third line of Equation (6) and thus to immediate adoption.

If (p_0, Λ_0) is in region (2), then initially no adoption takes place and the no-news posterior drifts upward according to the law of motion $\dot{p}_t = p_t(1-p_t)\varepsilon$, while Λ_t remains unchanged at Λ_0. This yields a unique time $0 < t_1^* = t_2^*$ at which (p_t, Λ_t) hits the boundary separating regions (2) and (1); from then on 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, Λ_t). If (p_0, Λ_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 Λ_t remains unchanged at Λ_0. However, now this yields a unique time $0 < t_1^*$ at which (p_t, Λ_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, Λ_t) is uniquely pinned down by the second line of Equation (6).[34] Thus, t_2^* is uniquely given by the first time t at which $\Lambda_t = \Lambda^*(p_t)$, at which point (p_t, Λ_t) enters region (1). Similar arguments show that when (p_0, Λ_0) starts in region (4), we have $t_1^* = 0$ and $t_2^* > t_1^*$ is the first time at

[32] Note that ρ 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.

[33] A bit more precisely, ignoring terms of order dt^2, the right hand side of Equation 7 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.

[34] Specifically, combining the second line of Equation (6) with Equation (1) yields the ODE:

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

which pins down p_t uniquely given the initial value $p_{t_1^*} = \overline{p}$:

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

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 $t_2^* = \inf\{t : \Lambda_t < \Lambda^*(p_t)\} < +\infty$.

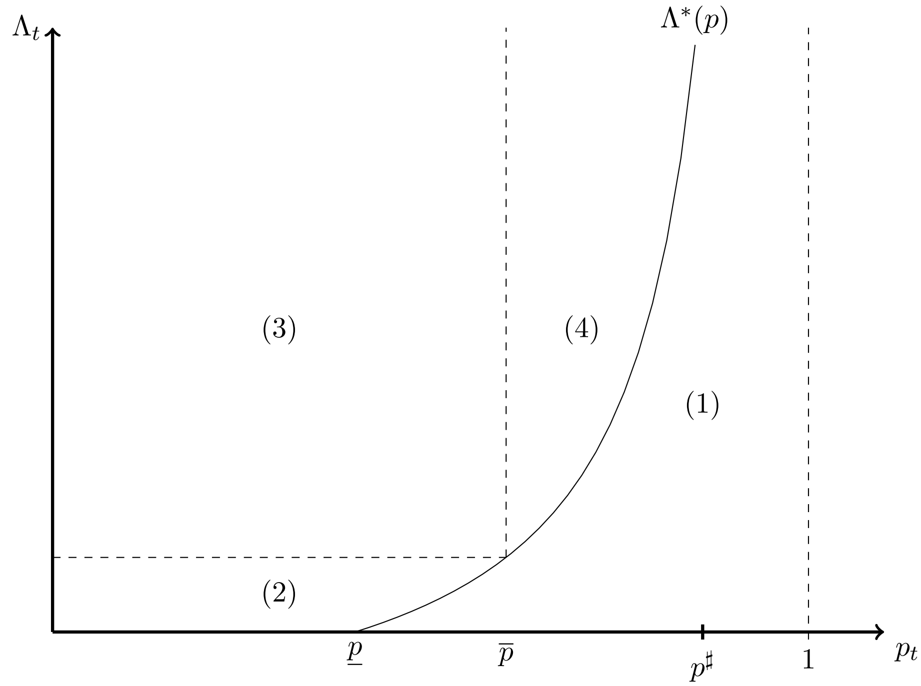

Figure 3: Partition of (p_t, Λ_t) when $\varepsilon < \rho$

which (p_t, Λ_t), evolving according to the second line of Equation (6), enters region (1). Finally, if (p_0, Λ_0) is in region (1), then $0 = t_1^* = t_2^*$ and absent breakdowns all consumers adopt upon their first opportunity from the beginning. This completes the description of the equilibrium.

As seen above, whether or not the equilibrium features a period of partial adoption depends on the fundamentals. More specifically, we can show that if consumers are forward-looking and not too optimistic, then $t_1^* < t_2^*$ arises whenever the potential for social learning in the economy is sufficiently large. To state this precisely, first note that from the Markovian description of equilibrium dynamics, it is easy to see that $\Lambda_0 = \lambda \bar{N}_0$ is a sufficient statistic for equilibrium when other fundamentals are fixed:

Lemma 5.5. *Fix $r, \rho > 0$, $p_0 \in (0,1)$, and $\varepsilon \geq 0$. Suppose that $\hat{\lambda}\hat{\bar{N}}_0 = \lambda_0 \bar{N}_0$. Let $\hat{N}_t$ and N_t denote the unique equilibrium adoption flows under $(\hat{\lambda}, \hat{\bar{N}}_0)$ and $(\lambda, \bar{N}_0)$, respectively, and let $\hat{p}_t$, $\hat{t}_1^*$, $\hat{t}_2^*$ and p_t, t_1^*, t_2^* denote the corresponding equilibrium beliefs and cutoff times. Then*

(i). *$\hat{t}_i^* = t_i^*$ for $i = 1, 2$;*

(ii). *$\hat{p}_t = p_t$ for all t*

(iii). *and* $\hat{\lambda}\hat{N}_t = \lambda N_t$ *for all* t.

Proof. Immediate from the proof of Theorem 5.1. ■

With this, the condition for partial adoption to arise in equilibrium can be stated as follows:

Lemma 5.6. *Fix* ρ, ε *and* p_0 *satisfying Conditions 5.2 and 5.3. Assume* $p_0 < p^\sharp$. *Then for all* $r > 0$, *there exists* $\bar{\Lambda}_0(r) > 0$ *such that* $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$ *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 Section F.1 in the Appendix. ■

On the other hand, if learning is *purely exogenous* ($\lambda = 0$ and $\varepsilon > 0$) or if 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 5.1 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.

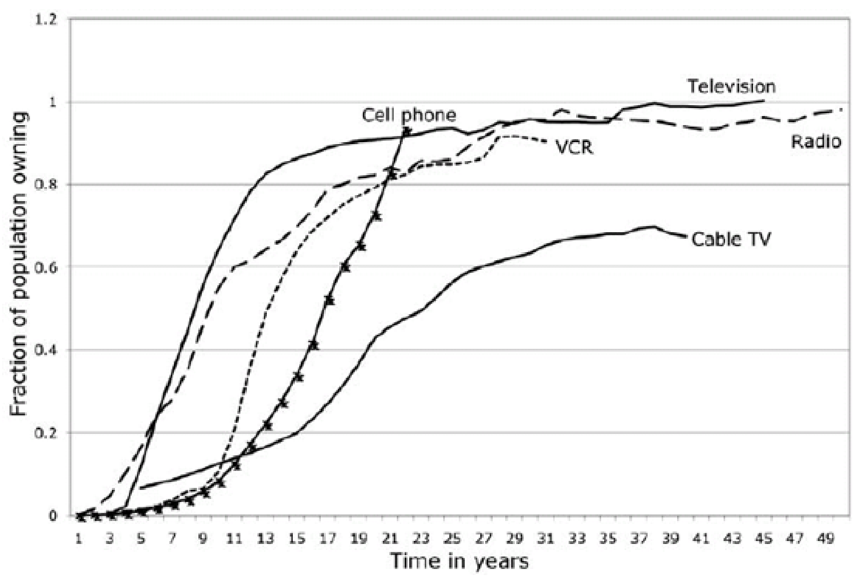

Figure 4: Examples of S-shaped adoption curves (Source: Narayanan and O'Connor (2010), Figure 2.1.)

5.2 Shape of Adoption Curve

With the equilibrium characterization in place, we can explore implications for the shape of an innovation's *adoption curve*, which plots the percentage of adopters in the population against time. Conditional on no breakdowns up to time t, this is given by

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

Conditional on the innovation being good, observed adoption levels at t will be exactly A_t. If the innovation is bad, then observed adoption levels follow A_t until the first breakdown (which occurs at a stochastic time), and remain constant from then on. As a result of the equilibrium characterization in Theorem 5.1, we obtain the following prediction for the shape of the adoption curve:

Corollary 5.7. *In the unique equilibrium of Theorem 5.1, the adoption curve A_t conditional on no breakdowns up to time 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.*

In particular, if $t_1^* < t_2^*$, then Corollary 5.7 predicts that, possibly after an initial period of no adoption, the adoption curve conditional on no breakdowns exhibits an *S-shaped* (i.e. convex-concave) growth pattern. In the empirical literature on innovation adoption,[35] S-shaped adoption patterns have been widely documented for many different innovations over the past century, including new agricultural seed varieties, such as hybrid corn; household electronics, such as refrigerators and color television; and industrial and medical innovations, such as the diesel locomotive and electrocardiographs. Figure 4 illustrates this for a selection of household technologies. Figure 5 represents a typical adoption curve generated in our model when $\varepsilon = 0$.

The intuition for S-shaped adoption curves in our model is as follows: There is no adoption before time t_1^*, because initially consumers are pessimistic about the quality of the innovation and strictly prefer to wait for information from the exogenous news source rather than risk adopting a bad product. The adoption curve is concave from time t_2^* on, because now consumers are

[35]See, for example, Griliches (1957), Mansfield (1961), Mansfield (1968), Davies (1979), and Gort and Klepper (1982), among many others.

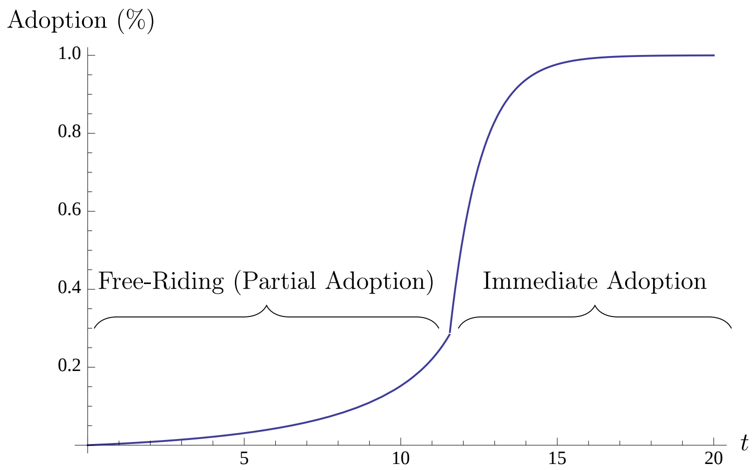

Figure 5: Adoption curve conditional on no breakdowns ($\varepsilon = 0$)

sufficiently optimistic to strictly prefer adopting the innovation when given the chance, so that the flow of new adopters is depleted at the rate ρ at which adoption opportunities are generated.

More interestingly, *the period of convex growth coincides precisely with the period of informational free-riding (in the form of partial adoption)*. The reason for this is the fundamental trade-off between adopting now and waiting for more information that arises when consumers are forward-looking social learners. During the period (t_1^*, t_2^*) of partial adoption, consumers are indifferent between adopting immediately and delaying. 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. In order to maintain indifference as captured by equation (7), the benefit to delaying adoption must also increase over time: Since consumers are forward-looking, this can be achieved by increasing the arrival rate of future breakdowns, which improves the odds that waiting will allow consumers to avoid the bad product. Since consumers are social learners, the arrival rate of information is increasing in the flow N_t of new adopters. Thus, whenever there is informational free-riding, N_t is strictly increasing over time. Since N_t represents the rate of change of the proportion A_t of adopters in the population, this is equivalent to A_t being convex.

Once again, this result relies crucially on our two modeling assumptions that learning is social and that consumers are forward-looking. As we pointed out following Lemma 5.6, if learning is purely exogenous or if consumers are myopic, then $t_1^* = t_2^*$, in which case the adoption curve does not feature a region of convex growth. In order to generate S-shaped adoption patterns in the absence of either of our assumptions, alternative models appeal to specific distributions of consumer heterogeneity, for example Jensen (1982) (in a model of exogenous learning with forward looking consumers) or Young (2009) (in a model of myopic social learning). The interplay of social learning and forward-looking consumers allows us to explain convex growth in terms of purely *informational* incentives, thus suggesting a novel micro-foundation for S-shaped curves that remains valid even when consumers are fully homogeneous.

5.3 Welfare

We now examine ex ante consumer welfare, as captured by the time 0 equilibrium value to waiting, W_0. Fix $r, \rho > 0$, $\varepsilon \geq 0$, and $p_0 \in (0,1)$ satisfying Conditions 5.2 and 5.3. Then Lemma 5.5 and Lemma 2.4 imply that $W_0 = W_0(\Lambda_0)$ depends only on the potential for social learning in the economy. The key finding is the possibility of a saturation effect: For sufficiently large Λ_0, additional increases in opportunities for social learning are welfare-neutral.

5.3.1 Nature of Inefficiency

We first note that, as is to be expected, the equilibrium is in general inefficient relative to the socially optimal cooperative benchmark:

Proposition 5.8. *Fix $r, \rho > 0$, $\varepsilon \geq 0$, and $p_0 \in (0,1)$ satisfying Conditions 5.2 and 5.3. The unique equilibrium in Theorem 5.1 is socially optimal if and only if $\Lambda_0 < \Lambda^*(p_0)$.*

Proof. See Appendix Section E.1. ■

Note that if $\Lambda_0 < \Lambda^*(p_0)$, then in equilibrium all consumers adopt immediately upon first opportunity, which is exactly as prescribed by the cooperative benchmark in Proposition 3.3. Whenever $\Lambda_0 > \Lambda^*(p_0)$, then the proof of Proposition 5.8 demonstrates two sources of inefficiency relative to the cooperative benchmark. First, provided we also have $p_0 \leq \overline{p}$ (so that $t_1^* > 0$), then adoption begins *too late* in equilibrium. Second, provided we also have $\Lambda_0 > \Lambda^*(\overline{p})$ (so that $t_1^* < t_2^*$), then even once consumers begin to adopt, the initial rate of adoption is *too slow* due to partial adoption on (t_1^*, t_2^*). Note that both types of inefficiency rely on a sufficiently large potential for social learning. Moreover, for any $p_0 > 1/2$, the first type arises only if ε is sufficiently large or r is sufficiently small (in particular, if learning is purely social or if consumers are myopic, then $t_1^* = 0$). On the other hand, the second inefficiency relies on consumers being forward-looking, but can arise even if $\varepsilon = 0$.

5.3.2 Saturation Effect

The fact that the equilibrium can feature inefficiencies relative to the cooperative benchmark is to be expected. However, the second type of inefficiency discussed above, which arises when there is free-riding in the form of partial adoption, has the following more surprising implication:

Proposition 5.9. *Fix $r, \rho > 0$, $\varepsilon \geq 0$, and $p_0 \in (0,1)$ satisfying Conditions 5.2 and 5.3. Let $\overline{\Lambda}_0 := \max\{\Lambda^*(p_0), \Lambda^*(\overline{p})\}$. Then in the unique equilibrium of Theorem 5.1, $W_0(\Lambda_0)$ satisfies the following:*

(i). *$W_0(\Lambda_0)$ is strictly increasing in Λ_0 whenever $\Lambda_0 < \overline{\Lambda}_0$;*

(ii). *$W_0(\Lambda_0) = W_0(\overline{\Lambda}_0)$ is constant in Λ_0 for all $\Lambda_0 \geq \overline{\Lambda}_0$.*

Proof. See Appendix Section F.1. ■

When $p_0 < p^\sharp$ so that $\overline{\Lambda}_0$ is finite, Proposition 5.9 states that an economy's ability to harness its potential for social learning is subject to a *saturation effect*: If Λ_0 is small, increases in Λ_0 are strictly beneficial; however, once Λ_0 is sufficiently large, any additional increase in Λ_0 is completely welfare-neutral. This is in stark contrast to the cooperative benchmark: There increases in Λ_0 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 $\Lambda_0 \to \infty$. We illustrate this in Figure 6,[36] which for varying levels of Λ_0 plots the ratio of equilibrium and socially optimal welfare levels.

To see the intuition for Proposition 5.9, suppose that $p_0 > \overline{p}$, so that $t_1^* = 0$. Then as long as $\Lambda_0 < \overline{\Lambda}_0$, all consumers adopt immediately upon their first opportunity until there is a breakdown. In this case, a slight increase in Λ_0 does not change consumers' behavior in the absence of a breakdown; however, conditional on the innovation being *bad*, it does increase the probability of a breakdown occurring prior to any given time—this is clearly welfare-improving, as more consumers are able to avoid the bad product. On the other hand, whenever $\Lambda_0 > \overline{\Lambda}_0$, then $t_1^* < t_2^*$, so that consumers must initially be *indifferent* between adopting and delaying. Then irrespective of the value of $\Lambda_0 \geq \overline{\Lambda}_0$, equilibrium incentives immediately imply that $W_0(\Lambda_0) = 2p_0 - 1$. In the next section, we provide some more intuition for the source of the saturation effect by studying the impact of increases in Λ_0 on equilibrium learning and adoption dynamics.

5.4 The Effect of Increased Opportunities for Social Learning

To further elucidate the saturation effect, this section examines the impact of an increase in Λ_0 on equilibrium learning dynamics and adoption levels. We find that the saturation effect

[36] Parameters used to generate the figure are: $\varepsilon = 0$, $r = 1$, $p_0 = 0.6$, and $\rho = 1$.

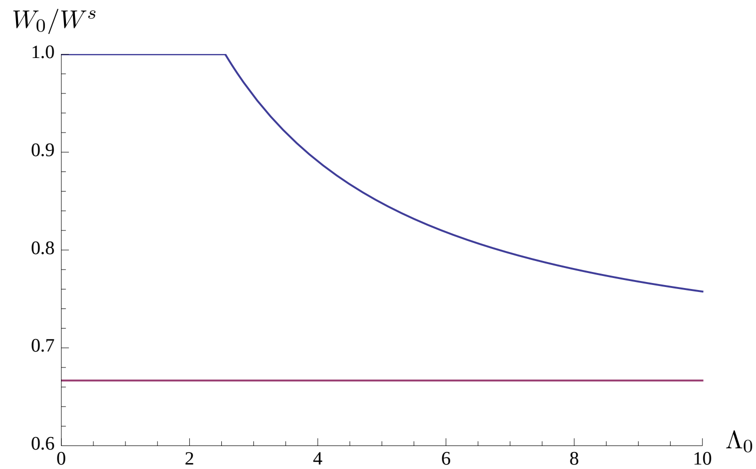

Figure 6: Ratio of equilibrium welfare to socially optimal welfare

corresponds to the following two surprising implications of partial adoption: Increased opportunities for social learning lead to strictly *less* learning over some periods of time and to a strict reduction in adoption of both good and bad innovations at all times.

Throughout this section we fix $r, \rho > 0$, $\varepsilon \geq 0$ and p_0 satisfying Conditions 5.2 and 5.3. To isolate the role of partial adoption, which is the inefficiency driving the saturation effect, we assume that $p^\sharp > p_0 > \overline{p}$, so that $t_1^* = 0$ and $\Lambda^*(p_0) < +\infty$. With these parameters fixed, Lemma 5.5 implies that Λ_0 is a sufficient statistic for all quantities we consider in this section. The following preliminary observation is central to the main results of this section:

Lemma 5.10. *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 $\hat{N}$ and N. 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. See Appendix Section F.2.1. ■

Point (ii) states that at all times during which there is partial adoption under both $\hat{\Lambda}_0$ and Λ_0, the rate of social learning is the same. Intuitively, this is because in order to maintain indifference between immediate adoption and an instantaneous delay, Equation (7) uniquely pins down the instantaneous arrival rate of breakdowns in the partial adoption region. The first bullet point states that increased opportunities for social learning prolong the initial period of partial adoption. To see the intuition, consider Figure 3: For any posterior p, $\Lambda^*(p)$, which represents the amount of future social information required to make consumers indifferent between adopting immediately at p and delaying, is the same under both $\hat{\Lambda}_0$ and Λ_0. However, since $\hat{\Lambda}_0 > \Lambda_0$ and since by (ii) the evolution of beliefs in region (4) is the same under $\hat{\Lambda}_0$ and Λ_0, it takes longer to reach the Λ^*-curve from the initial point $(p_0, \hat{\Lambda}_0)$.

5.4.1 Non-Monotonicity of Learning

In this section, we consider the effect of increased opportunities for social learning on the evolution of equilibrium beliefs. The following proposition states a non-monotonicity result: Increases in Λ_0 do not necessarily translate into increases in p_t at all times t. Specifically, if $\Lambda_0 \geq \Lambda^*(p_0)$, corresponding to the cutoff for the saturation effect, then upon an increase in Λ_0 there is a period of times at which p_t is strictly lower:

Proposition 5.11. *Fix r, ρ, ε, and p_0 satisfying Conditions 5.2 and 5.3 and such that $p_0 \in (\overline{p}, p^\sharp)$. Consider $\hat{\Lambda}_0 = \hat{\lambda}\hat{\bar{N}}_0$ and $\Lambda_0 = \lambda \bar{N}_0$ such that $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. Then there exists some $\overline{t} \in (t_2^*(\Lambda_0), +\infty)$ 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* $t \in (t_2^*(\Lambda_0), \overline{t})$,
- $p_t^{\Lambda_0} < p_t^{\hat{\Lambda}_0}$ *for all* $t > \overline{t}$.

However, when $\Lambda_0 < \Lambda^*(p_0)$, then $p_t^{\Lambda_0}$ is strictly increasing in Λ_0 for all t.

Proof. See Appendix Section F.2.2. ■

Note that by Equation (1), the probability of a breakdown occurring prior to time t conditional on the innovation being bad is given by

$$1 - e^{-\int_0^t (\varepsilon + \lambda N_s) ds} = 1 - \frac{p_0 (1 - p_t)}{p_t (1 - p_0)}, \tag{8}$$

which is increasing in p_t. Thus, Proposition 5.11 has the surprising implication that whenever Λ_0 is large enough, any additional increase in opportunities for social learning will result in consumers being strictly *less* likely to find out about a bad product over a period of times.

The intuition for Proposition 5.11 is closely related to whether or not there is free-riding in the form of partial adoption (and hence relies on consumers being forward-looking social learners). Whenever $\Lambda_0 < \Lambda^*(p_0)$, then $0 = t_1^* = t_2^*$, so that absent breakdowns all consumers adopt immediately upon their first opportunity. In this case, it is easy to see from Theorem 5.1 that the rate λN_t at which social learning occurs is strictly increasing in Λ_0: We have $\lambda N_t = \rho e^{-\rho t}\Lambda_0$ for all t. Thus, by Equation (8) increasing Λ_0 necessarily speeds up learning at all times.

On the other hand, if $\Lambda_0 > \Lambda^*(p_0)$, then $0 = t_1^* < t_2^*(\Lambda_0)$ and the equilibrium features an initial region of partial adoption. In this case, an increase to $\hat{\Lambda}_0 > \Lambda_0$ has the following effect. By Lemma 5.10, free-riding occurs over a longer period of time: $t_2^*(\hat{\Lambda}_0) > t_2^*(\Lambda_0)$; moreover, at all times $t \leq t_2^*(\Lambda_0)$ where there is free-riding under both Λ_0 and $\hat{\Lambda}_0$, the rate of social learning is the *same*: $\lambda N_t = \hat{\lambda}\hat{N}_t$. This explains the first bullet point in Proposition 5.11. The strict slowdown in learning at times just after $t_2^*(\Lambda_0)$ is due to the following: The proof of Theorem 5.1 shows that whenever $t_1^* < t_2^*$, then the flow of adopters N_t is continuous at all times *except* at exactly t_2^*, where there is a discontinuous increase. This is evident from the adoption curve in Figure 5 where a visible non-differentiability exists at the point of transition from partial adoption to immediate adoption. Since $t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$, this means that at $t_2^*(\Lambda_0)$ the difference between $\hat{\lambda}\hat{N}_t$ and λN_t jumps from 0 to a strictly negative value, resulting in the temporary slowdown in learning stated in the second bullet point.

Finally, learning under $\hat{\Lambda}_0$ must eventually overtake learning under Λ_0, because at time 0 the payoff to immediate adoption is the same under both Λ_0 and $\hat{\Lambda}_0$ and in both cases consumers are indifferent between adopting immediately and delaying. This relates back to the saturation effect for welfare observed in Proposition 5.9 as follows. By Lemma 2.4, ex-ante welfare W_0 under $\Lambda_0 > \Lambda^*(p_0)$ can be written as

$$W_0(\Lambda_0) = \int_0^\infty \rho e^{-(r+\rho)\tau} \frac{p_0}{p_\tau^{\Lambda_0}} \left(2p_\tau^{\Lambda_0} - 1\right) d\tau = p_0 \int_0^\infty \rho e^{-(r+\rho)\tau} \left(2 - \left(p_\tau^{\Lambda_0}\right)^{-1}\right) d\tau,$$

and similarly for $\hat{\Lambda}_0 > \Lambda_0$. The non-monotonicity result for beliefs then has the following implication. If a consumer obtains his first adoption opportunity prior to $t_2^*(\Lambda_0)$, his expected payoff is the same under Λ_0 and $\hat{\Lambda}_0$; if his first adoption opportunity is during $(t_2^*(\Lambda_0), \bar{t})$, he is strictly worse off under $\hat{\Lambda}_0$, because in case the innovation is bad he is less likely to have found out by then; finally, if his first opportunity is after $\bar{t}$, 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 $(t_2^*(\Lambda_0), \bar{t})$. This produces the saturation effect in Proposition 5.9.

Even more strongly, in Section 7, we exploit the non-monotonicity result for beliefs to construct an example involving consumers with heterogeneous discount rates in which an increase in Λ_0 is not only not beneficial, but in fact strictly *hurts* aggregate welfare.

5.4.2 Slowdown in Adoption

We now consider the effect of increased opportunities for social learning on observed adoption levels, analyzing separately the case of a good innovation and of a bad innovation.

Adoption Conditional on a Good Product: Recall that A_t denotes the percentage of consumers in the population who adopt the innovation by time t conditional on no breakdowns before t, which is the same as the percentage of adopters at t conditional on the innovation being good:

$$A_t(G) = A_t := \int_0^t \frac{N_s}{\bar{N}_0} ds.$$

Note also that by Lemma 5.5, Λ_0 is a sufficient statistic for the equilibrium levels of A_t holding fixed r, ρ, p_0 and ε, because $\frac{N_s}{\bar{N}_0} = \frac{\lambda N_s}{\Lambda_0}$ and Λ_0 is a sufficient statistic for λN_s.

For a good innovation, we show that when the potential for social learning Λ_0 is small, additional small increases in opportunities for social learning have no effect on adoption levels, but when Λ_0 is sufficiently large, increases strictly drive down adoption levels at all times. Once again, the cutoff is given by the level $\Lambda^*(p_0)$ above which partial adoption occurs.

Proposition 5.12. *Fix r, ρ, ε, and p_0 satisfying Conditions 5.2 and 5.3 and such that $p_0 \geq \overline{p}$. Then for all t, $A_t(\Lambda_0, G)$ is constant in Λ_0 for all $\Lambda_0 \leq \Lambda^*(p_0)$ and strictly decreasing in Λ_0 for all $\Lambda_0 > \Lambda^*(p_0)$.*

Proof. See Appendix Section F.3. ∎

The reason why $A_t(\Lambda_0, G)$ is constant for all $\Lambda_0 \leq \Lambda^*(p_0)$ is familiar: For all such Λ_0, consumers adopt upon their first opportunity and $A_t = 1 - e^{-\rho t}$. If $\Lambda_0 > \Lambda^*(p_0)$, then the strict slowdown in adoption is due to increased free-riding in the form of partial adoption. More precisely, an increase from $\Lambda_0 > \Lambda^*(p_0)$ to $\hat{\Lambda}_0$ has two effects, as summarized in Lemma 5.10: First, on the *extensive margin*, increased opportunities for social learning push out t_2^* and lead to a longer period of free-riding under $\hat{\Lambda}_0$. Second, on the *intensive margin*, the increase strictly drives down the growth rate of A_t at all times prior to $t_2^*(\Lambda_0)$:

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

Figure 7 illustrates these two effects and their implications for a strict slowdown in adoption. Finally, from $t_2^*(\Lambda_0)$ adoption occurs at a maximal rate under Λ_0, so that from then on $\hat{A}_t$ must remain below A_t by feasibility.

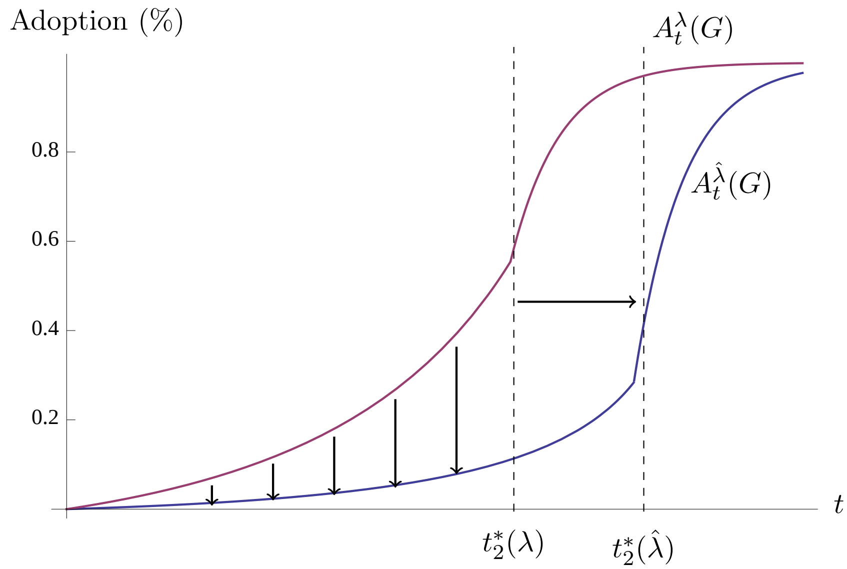

Figure 7: Changes in adoption levels of a good product under perfect bad news ($\hat{\lambda} > \lambda$)

Two remarks are in order. First, our prediction of a strict slowdown of adoption of the good product in response to increased opportunities for social learning once again relies crucially on consumers being forward-looking. If consumers are myopic, then by the first part of Proposition 5.12 adoption levels at all times remain unchanged following the increase. More interestingly, if consumers are myopic, it is not possible to generate this prediction under perfect bad news even if we allow for an *arbitrary* distribution of heterogeneity in tastes. Thus, while models of innovation adoption by myopic social learners, such as Young (2009), can generate S-shaped adoption curves by imposing suitable distributions of consumer heterogeneity, the prediction in our model of a strict reduction in initial adoption of a good innovation is novel.

Second, Proposition 5.12 implies that conditional on a good product, increased opportunities for social learning are welfare-neutral at best (if $\Lambda_0 < \Lambda^*(p_0)$) and potentially strictly harmful (if $\Lambda_0 \geq \Lambda^*(p_0)$), since adoption levels are unchanged in the former case and in the latter case adoption is strictly delayed. Therefore any potential welfare gains due to increased opportunities for social learning must result from more consumers being able to avoid the *bad* product. We now study this point by analyzing the effect of increases in Λ_0 on adoption levels of a bad product.

Adoption Conditional on a Bad Product: Conditional on a bad innovation, adoption is stochastic, following A_t until the first breakdown, which occurs at a random time, and remaining constant from then on. We therefore study the effect of increased opportunities for social learning on the *expected* percentage of adopters at time t conditional on a bad product, which is given by:

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

where the second line is obtained by integrating the first expression by parts. Again, Λ_0 is a sufficient statistic for $A_t(B)$ when r, ρ, p_0, and ε are fixed. For bad innovations, increased opportunities for social learning always produce strict decreases in the expected level of adoption at all times, irrespective of the original level of Λ_0:

Proposition 5.13. *Fix r, ρ, ε, and p_0 satisfying Conditions 5.2 and 5.3 and such that $p_0 \geq \overline{p}$. Then $A_t(\Lambda_0, B)$ is strictly decreasing in Λ_0 for all $t > 0$.*

Proof. See Appendix Section F.3. ∎

If $\Lambda_0 < \Lambda^*(p_0)$, this is immediate since by Proposition 5.11 and Proposition 5.12 adoption levels conditional on no breakdowns are the same, but breakdowns prior to any time are more likely for higher values of $\Lambda_0 < \Lambda^*(p_0)$. If $\Lambda_0 \geq \Lambda^*(p_0)$, then there is a tension: On the one hand, Proposition 5.12 implies that an increase in Λ_0 leads at all times to strictly lower adoption levels conditional on no breakdowns, but on the other hand, the non-monotonicity result for learning implies that there are times before which a breakdown is strictly more likely under lower Λ_0. We show that the former effect always strictly dominates.

Proposition 5.12 and Proposition 5.13 relate to the saturation effect observed in Proposition 5.9 as follows: If $\Lambda_0 < \Lambda^*(p_0)$, then small increases in opportunities for social learning do not affect adoption conditional on the good product, but strictly decrease the number of consumers adopting the bad product by any time, leading to an overall welfare gain. On the other hand, if $\Lambda_0 \geq \Lambda^*(p_0)$, then increased opportunities for social learning strictly decrease adoption both for good products (which is harmful) and for bad products (which is beneficial), making welfare predictions a priori ambiguous. However, the saturation effect illustrates that in welfare terms these two implications balance out exactly.

6 Perfect Good News

6.1 Equilibrium Characterization

We now turn to study equilibrium behavior when learning is via perfect good news. As under perfect bad news, the unique equilibrium is Markovian in the state variables (p_t, Λ_t). Surprisingly, however, regardless of the potential for social learning in the economy, the unique equilibrium under perfect good news does not exhibit any region of partial adoption and adoption at each time is *all-or-nothing*:

Theorem 6.1. *Let $r, \rho, \bar{N}_0 > 0$, $p_0 \in (0,1)$, and $\lambda, \varepsilon \geq 0$. There exists a unique equilibrium. In the unique equilibrium, N_t is Markovian in (p_t, Λ_t) (or equivalently $(p_t, \bar{N}_t)$) for all t and satisfies:*

$$N_t = \begin{cases} \rho \bar{N}_t & \text{if } p_t > p^* \\ 0 & \text{if } p_t \leq p^*. \end{cases} \tag{9}$$

where

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

To prove Theorem 6.1 we again invoke the quasi-single crossing property for equilibrium incentives established in Theorem 4.1. Suppose $N_{t\geq 0}$ is an equilibrium flow of adopters. Let $p_{t\geq 0}$ and $W_{t\geq 0}$ be the corresponding no-news posterior and value to waiting, and let $\Lambda_{t\geq 0} := \lambda \bar{N}_{t\geq 0}$ describe the evolution of the economy's potential for social learning. By Theorem 4.1, there are times[37]

$$\begin{aligned} t_1^* &:= \inf\{t : 2p_t - 1 \leq W_t\}, \\ t_2^* &:= \inf\{t : 2p_t - 1 < W_t\}, \end{aligned}$$

such that (appealing also to right-continuity) N must satisfy

$$\begin{cases} N_t = \rho \bar{N}_t & \text{if } t < t_1^*, \\ 2p_t - 1 = W_t & \text{if } t \in [t_1^*, t_2^*) \\ N_t = 0 & \text{if } t \geq t_2^*. \end{cases}$$

In the following, we build on this fact to establish the existence of a unique equilibrium as a function of the parameter values. The following lemma establishes the all-or-nothing nature of the perfect good news equilibrium:

[37] With the usual convention that $\inf \varnothing = +\infty$.

Lemma 6.2. *Suppose either $\varepsilon > 0$ or $p_0 \neq \frac{1}{2}$.*[38] *Let $N_{t\geq 0}$ be an equilibrium with associated threshold times t_1^* and t_2^*. Then $t_1^* = t_2^* =: t^*$.*

Proof. See Appendix Section C.2. ■

Thus, absent breakthroughs, *all* consumers adopt immediately if given the chance prior to t^*, and after t^*, consumers stop adopting altogether and rely solely on information generated by exogenous sources (if $\varepsilon = 0$, both adoption and learning come to a permanent standstill at this point). If a breakthrough occurs at any time (prior to or after t^*), then from then on all consumers adopt the innovation whenever given a chance.

To see the intuition for the all-or-nothing nature of the equilibrium, suppose we had $t_1^* < t_2^*$. Then consumers would be indifferent between adopting and delaying at each time $t \in (t_1^*, t_2^*)$. As with perfect bad news, we can compare a consumer's payoff to adopting at t with the payoff to delaying his decision by an instant and decompose the difference into two terms:

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

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 if given the chance and delaying. The second term represents the gain to immediate adoption if there is a breakthrough between t and $t + dt$, which happens with instantaneous probability $p_t(\lambda N_t + \varepsilon)dt$. 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, which contradicts 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

[38] If $\varepsilon = 0$ and $p_0 = \frac{1}{2}$, then it is easy to see that the unique equilibrium must be $N \equiv 0$, so that $t_1^* = 0 < t_2^* = +\infty$.

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: If there is a breakdown between t and $t + dt$, then the innovation is revealed to be bad and no one is willing to adopt from $t + dt$ on. 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.[39]

With Lemma 6.2, the derivation of Theorem 6.1 is straightforward. To this end, we show that any equilibrium can be characterized in terms of a cutoff posterior that only depends on ε, ρ, and r. Given any equilibrium $N_{t\geq 0}$ with associated no-news posteriors $p_{t\geq 0}$, value to waiting $W_{t\geq 0}$, and cutoff time t^*, 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. By optimality of W_t, we must have $H_t \leq W_t$ for all t. We can define a lower bound for H_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. Note that for all t, we have $H(p_t, 0) \leq H_t$: When breakthroughs are generated both exogenously and at rate λN_t, then the probability that a breakthrough is generated by any given time is (weakly) greater than if learning is purely exogenous; this benefits a consumer who only adopts once there has been a breakthrough. Moreover, for all $t \geq t^*$, we have $W_t = H(p_t, 0)$.

Recall the definition of p^* in Theorem 6.1, $p^* := \frac{(\varepsilon+r)(\rho+r)}{2(\varepsilon+\rho)(\varepsilon+r)-\varepsilon\rho}$, and note that p^* is the unique solution to $2p^* - 1 = H(p^*, 0)$. By definition of p^*, if $p_t \leq p^*$ at any time t, then

$$2p_t - 1 \leq H(p_t, 0) \leq H_t \leq W_t,$$

[39]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 the case.

so for all $t < t^*$, we must have $p_t > p^*$. Conversely, if $t^* < +\infty$ and $t \geq t^*$, then

$$2p_t - 1 \leq W_t = H(p_t, 0),$$

so $p_t \leq p^*$ for all $t \geq t^*$. We summarize this in the following lemma:

Lemma 6.3. *Let $N_{t\geq 0}$ be an equilibrium with corresponding cutoff time t^* and no-news posterior $p_{t\geq 0}$. Then*

$$p_t > p^* \Leftrightarrow t < t^*.$$

Given Lemma 6.2 and Lemma 6.3, the equilibrium characterization under perfect good news follows readily. Equation (9) is immediate from Lemma 6.3. For fixed parameters, we then obtain the unique equilibrium as follows: If $p_0 \leq p^*$, then $t^* = 0$ and $N_t = 0$ for all t. If $p_0 > p^*$, then we must have $t^* > 0$ and $N_t = \rho e^{-\rho t}\bar{N}_0$ for all $t < t^*$; if in addition $\varepsilon > 0$ or $p_0\left(1 + e^{-\lambda \bar{N}_0}\right) < 1$, then $t^* < +\infty$ is uniquely determined as the solution to

$$p_t = \frac{p_0}{p_0 + (1 - p_0)\, e^{\varepsilon t + (1 - e^{-\rho t})\bar{N}_0}} = p^*. \tag{10}$$

If instead $p_0 > p^*$ and $\varepsilon = 0$ and $p_0\left(1 + e^{-\lambda \bar{N}_0}\right) \geq 1$, then Equation (10) does not admit a solution, and we must have $t^* = +\infty$: In this case, the potential for social learning in the economy is so small that even a bad innovation is eventually adopted by all consumers, despite the fact that no breakthroughs are ever generated.

As highlighted at the beginning of the section, the equilibrium under perfect good news is Markovian in (p_t, Λ_t). However, in marked contrast to the bad news case, if $\varepsilon = 0$, then adoption behavior is independent of the discount rate r: Even very patient consumers will behave entirely myopically, adopting the innovation at all posteriors above $\frac{1}{2}$ and not adopting otherwise. 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 λ or $\bar{N}_0$: Social learning only affects the *time* t^* at which adoption ceases conditional on no breakthroughs. Moreover, as under perfect bad news, it is easy to see that holding fixed other parameters, $\Lambda_0 = \lambda \bar{N}_0$ is a sufficient statistic for equilibrium behavior:

Lemma 6.4. *Fix $r, \rho > 0$, $p_0 \in (0, 1)$, and $\varepsilon \geq 0$. Suppose that $\hat{\lambda}\hat{\bar{N}}_0 = \lambda_0 \bar{N}_0$. Let $\hat{N}_t$ and N_t denote the unique equilibrium adoption flows under $(\hat{\lambda}, \hat{\bar{N}}_0)$ and $(\lambda, \bar{N}_0)$, respectively, and let $\hat{p}_t$, $\hat{t}^*$ and p_t, t^* denote the corresponding equilibrium beliefs and cutoff times. Then*

(i). $\hat{t}^* = t^*$ *;*

(ii). $\hat{p}_t = p_t$ *for all* t

(iii). *and* $\hat{\lambda}\hat{N}_t = \lambda N_t$ *for all* t.

Proof. Immediate from the proof of Theorem 6.1. ■

6.2 Shape of Adoption Curve

Theorem 6.1 has the following implication for the shape of adoption curves in good news markets:

Corollary 6.5. *In the unique equilibrium of Theorem 6.1, the proportion of adopters in the population is strictly increasing and concave for all* $t < t^*$ *and given by*

$$A_t := \int_0^t \frac{N_s}{\bar{N}_0} ds = 1 - e^{-\rho t}.$$

If there is a breakthrough prior to t^**, then the proportion of adopters is given by* $1 - e^{-\rho t}$ *for all* t*; if the first breakthrough occurs at* $s > t^*$*,*[40] *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* $1 - e^{-\rho(t^*+t-s)}$.

Thus, as illustrated in Figure 8,[41] adoption proceeds in *concave "bursts"*: Up to time t^*, all consumers adopt the innovation upon their first opportunity, with the flow of new adopters declining at the rate ρ at which adoption opportunities arrive. Conditional on no breakthroughs, adoption comes to a standstill at time t^*, because by that point consumers are pessimistic enough about the product to prefer to delay adoption. If $\varepsilon > 0$, then exogenous news sources might generate a breakthrough after t^*, in which case a second concave burst in adoption occurs.

While less common than the S-shaped curves we predicted under bad news,[42] this type of adoption pattern also corresponds to recurrent empirical findings. For instance, the marketing literature[43] has coined the term "fast-break" product life cycle (PLC) to describe goods with large initial sales volumes accompanied by a gradual decline in new purchases (implying a concave adoption pattern), in contrast to S-shaped PLCs that initially feature low sales volumes accompanied by a gradual increase in the number of new purchases. The textbook example for

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

[41] The parameters used to generate the figure are: $\varepsilon = 1/2$, $r = 1$, $\rho = 1$, $\lambda = 0.5$, and $p_0 = 0.7$.

[42] Note that in our model purely concave adoption curves can also arise under bad news if the economy's potential for social learning is relatively limited or consumers are very optimistic (so that $t_1^* = t_2^*$). The key difference is that under perfect good news adoption curves are *necessarily* concave, even in economies with a large potential for social learning or with fairly pessimistic and forward-looking consumers.

[43] Cf. Keillor (2007) pp. 51-61.

fast-break PLCs is the movie industry,[44] as illustrated in Figure 9. Given that the movie industry is also sometimes cited as a typical example of a good news market[45] with learning occurring predominantly via positive events such as awards and recommendations in social media, this finding appears to be in line with our model.

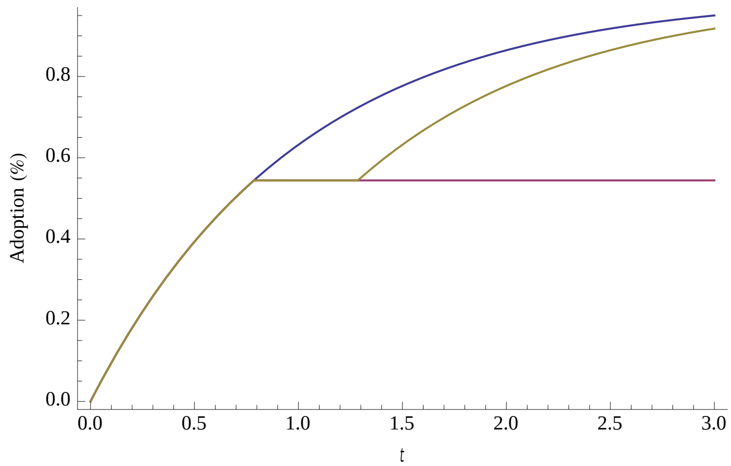

Figure 8: Adoption Curves under Perfect Good News (blue = breakthrough before t^*; yellow = breakthrough after t^*; pink = bad quality)

6.3 The Effect of Increased Opportunities for Social Learning

To further illustrate the distinction between good news and bad news markets, we now study the effect of increased opportunities for social learning under good news. In contrast to our results under perfect bad news, we find that increased opportunities for social learning (essentially) always speed up learning, leave initial adoption levels unaffected, and are strictly welfare-improving—all three results are due to the absence of partial adoption regions under good news. Throughout this section, we fix $r, \rho > 0$, $p_0 \in (0, 1)$, and $\varepsilon \geq 0$, and let p^* denote the equilibrium cutoff posterior:

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

[44] Additional evidence can be found in Davies (1979)'s study of the diffusion of 22 post-war process innovations among industries in the UK. In the context of his probit model of innovation diffusion, he finds that while S-shaped (logistic) diffusion paths are characteristic of complex and expensive innovations, they are less suited to fitting the diffusion paths of simpler and less expensive innovations, which typically feature rapid, essentially concave growth from the beginning and are better approximated by a lognormal model.

[45] Cf. Board and Meyer–ter–Vehn (2013)

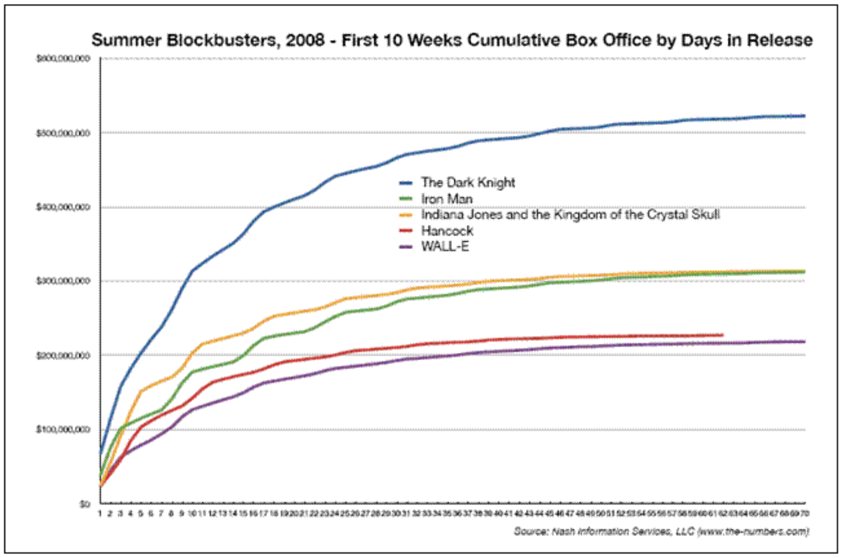

Figure 9: "Adoption" patterns for various blockbuster movies (Source: McLaren and DePaolo (2009))

which is independent of the potential for social learning. Given that all other parameters are fixed, Lemma 6.4 implies that $\Lambda_0 = \lambda \bar{N}_0$ is a sufficient statistic for all the quantities we consider in this section.

6.3.1 Learning Speeds Up

We first turn to the effect of increased opportunities for social learning on equilibrium beliefs. As a result of the all-or-nothing nature of the perfect good news equilibrium, we can see that learning necessarily speeds up—this is in contrast to the possibility of nonmonotonicities due to partial adoption under perfect bad news. More precisely:

Proposition 6.6. *Fix $\hat{\Lambda}_0 > \Lambda_0$[46] and let $t^*(\hat{\Lambda}_0)$, $p_t^{\hat{\Lambda}_0}$ and $t^*(\Lambda_0)$, $p_t^{\Lambda_0}$ denote the corresponding equilibrium cutoff times and posteriors conditional on no breakthrough.*

(i). *If $p_0 > p^*$, then*

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

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

(ii). *If* $p_0 \leq p^*$, *then*

- $t^*(\hat{\Lambda}_0) = t^*(\Lambda_0) = 0$
- $p_t^{\hat{\Lambda}_0} = p_t^{\Lambda_0}$ *for all* t.

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 6.1, there is never any partial adoption, so that an increase from Λ_0 to $\hat{\Lambda}_0$ directly translates into a faster rate of social learning at all times t prior to $\min\{t^*(\hat{\Lambda}_0), t^*(\Lambda_0)\}$: $\lambda N_t = \rho e^{-\rho t}\Lambda_0 < \rho e^{-\rho t}\hat{\Lambda}_0 = \hat{\lambda}\hat{N}_t$. Since the cutoff posterior p^* is independent of social learning, this implies that $t^*(\hat{\Lambda}_0) < t^*(\Lambda_0)$ and that learning is strictly faster under $\hat{\Lambda}_0$ at all times. However, once the cutoff posterior is reached, information is generated at the constant exogenous rate ε, which means that conditional on $t > t^*$, beliefs depend only on $t - t^*$, as summarized in the third bullet point under (i).

On the other hand, 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.

6.3.2 No Initial Slowdown of Adoption

The all-or-nothing nature of the perfect good news equilibrium also implies that increased opportunities for social learning do not affect initial adoption levels—this is again in contrast to the possibility of initial slowdowns due to partial adoption under perfect bad news. More precisely:

Proposition 6.7. *Suppose* $\hat{\Lambda}_0 > \Lambda_0$.

(i). *If* $p_0 > p^*$, *then:*

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

(ii). *If* $p_0 \leq p^*$, *then for all* t*:*

- $A_t(\Lambda_0; B) = A_t(\hat{\Lambda}_0; B) = 0$;
- $A_t(\Lambda_0; G) = A_t(\hat{\Lambda}_0; G) = \left(1 - \frac{\rho}{\rho-\varepsilon}e^{-\varepsilon t}\right) + \frac{\varepsilon}{\rho-\varepsilon}e^{-\rho t}$.

Until $t^*(\hat{\Lambda}_0)$ all consumers adopt immediately upon an opportunity under both Λ_0 and $\hat{\Lambda}_0$ regardless of the quality of the innovation. However, from $t^*(\hat{\Lambda}_0)$ on, expected adoption levels are strictly lower under $\hat{\Lambda}_0$ than under Λ_0: If the innovation is bad, this is because adoption comes to a permanent standstill under $\hat{\Lambda}_0$ (until a further breakthrough generated by the exogenous information ε), but continues until $t^*(\Lambda_0)$ under Λ_0. If the the innovation is good, the result is again immediate for all $t \leq t^*(\Lambda_0)$ since adoption occurs at the maximal rate under Λ_0. For $t > t^*(\Lambda_0)$, there are two opposing effects: On the one hand, the guaranteed lower bound on adoption is higher under Λ_0, but on the other hand the probability of a breakthrough occurring prior to time t is always higher under $\hat{\Lambda}_0$. We show in the Appendix Section G.1 that the former effect dominates.

On the other hand, if $p_0 \leq p^*$, then increased opportunities for social learning once again have no effect at all on adoption levels, because no consumers adopt until the exogenous news source generates a breakthrough.

6.3.3 No Saturation Effect

Proposition 6.7 showed that from time $t^*(\hat{\Lambda}_0)$ on, adoption levels for both good and bad quality products are strictly lower under $\hat{\Lambda}_0 > \Lambda_0$ than under Λ_0. In welfare terms, the former effect is harmful while the latter is beneficial. This raises the question whether welfare under perfect good news might be subject to a similar saturation effect as under bad news. Provided $p_0 > p^*$ and $\varepsilon > 0$, the answer is negative:

Proposition 6.8. *Suppose* $\hat{\Lambda}_0 > \Lambda_0$.

- *If* $p_0 > p^*$ *and* $\varepsilon > 0$, *then* $W_0(\hat{\Lambda}_0) > W_0(\Lambda_0)$.
- *If* $p_0 \leq p^*$ *or* $\varepsilon = 0$, *then* $W_0(\hat{\Lambda}_0) = W_0(\Lambda_0)$.

Thus, in contrast to the perfect bad news case, increased opportunities for social learning are always strictly beneficial, except in two cases: If consumers rely entirely on exogenous information ($p_0 \leq p^*$), or if there is no exogenous information ($\varepsilon = 0$). Welfare-neutrality in these two exceptional cases is clear: Increased opportunities for social learning can have an effect on welfare only if there are histories at which a consumer's decision whether to adopt or delay is affected by information generated as a result of social learning. If $p_0 \leq p^*$, then consumers' behavior depends only on information obtained exogenously (and no adoption ever takes place if $\varepsilon = 0$). If $\varepsilon = 0$ and $p_0 > p^* = \frac{1}{2}$, then consumers are willing to adopt at all histories, since no matter how large Λ_0, the equilibrium posterior always remains weakly above $\frac{1}{2}$.

On the other hand, if $p_0 > p^*$ and $\varepsilon > 0$, then under both Λ_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 π^* of a breakthrough occurring prior to p^* being reached is the same under both Λ_0 and $\hat{\Lambda}_0$: $\pi^* = \frac{1-p_0}{1-p^*}$. And because learning occurs at the same rate once p^* is reached, the continuation value W^* conditional on p^* being reached is also the same: $W^* = p^* \int_0^\infty \varepsilon e^{-(\varepsilon+r)t} \frac{\rho}{r+\rho} dt = 2p^* - 1$. So the only difference is that conditional on no breakthroughs, the time t^* at which p^* is reached occurs earlier under $\hat{\Lambda}_0$. To see that this is strictly beneficial, note that W_0 is composed of the following two terms:

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

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

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

Intuitively, in the second case the consumer adopts the innovation only once it has been revealed to be good while in the first case he adopts it regardless of its quality, and the resulting benefit from avoiding a bad innovation outweighs the cost of possibly having to delay adoption of a good innovation.

Nature of Inefficiency: Even though there is no saturation effect and consumers are able to always benefit from increased opportunities for social learning, equilibrium adoption behavior is not in general socially optimal. Let p^s denote the cutoff posterior for the cooperative benchmark derived in Proposition 3.2.

Proposition 6.9. *If $\varepsilon = 0$, equilibrium adoption behavior is socially optimal if and only if either $p_0(1 + e^{-\Lambda_0}) \geq 1$ or Inequality 5 holds. If $\varepsilon > 0$, then equilibrium adoption behavior is socially optimal if and only if $p^s \geq p_0$.*

Consider first the case where $\varepsilon = 0$. Then if $p_0(1 + e^{-\Lambda_0}) \geq 1$, we have that $t^* = t^s = +\infty$; and if Inequality 5 holds, then $t^* = t^s = 0$. For the converse and to deal with the case when $\varepsilon > 0$, it then suffices to show that $p^s < p^*$: This implies that whenever $p_0 > p^s$, then conditional on no breakthroughs *adoption ends too soon* in equilibrium (or doesn't take place at all if $p_0 \leq p^*$ even though the cooperative benchmark prescribes some initial adoption). On the other hand, if $p_0 \leq p^s$, then both the cooperative benchmark and the equilibrium prescribe no adoption until there has been an exogenously generated breakdown. Note that adoption ending too soon under the perfect good news equilibrium is the analog of adoption beginning inefficiently late under perfect bad news. However, since the perfect good news equilibrium does not feature regions of partial adoption, there is no analog of the second type of inefficiency that arose under perfect bad news: Whenever adoption does occur under perfect good news, it takes place at an optimal rate.

To see that $p^s < p^*$, note that

$$\left(2 - \frac{\varepsilon}{\varepsilon + r}\frac{\rho}{\rho + r}\right) p^* - 1 = 0.$$

Using the above equality and evaluating the derivative of the objective function of the cooperative problem in Equation 3 at p^*, we obtain:

$$(1 - \pi^*)p^*\Lambda_0\rho e^{-\rho t^*}\frac{r}{\varepsilon + r}\left(e^{-(r+\rho)t^*}\frac{\rho}{\rho + r}\bar{N}_0\right) > 0.$$

This shows that $t^s > t^*$ and so $p^s < p^*$ as the objective function of the cooperative problem is single-peaked.

7 More Social Learning Can Hurt: An Example

In Proposition 5.9 we established the saturation effect, whereby increased opportunities for social learning under perfect bad news are welfare-neutral when Λ_0 is sufficiently large relative to the other fundamentals. Nevertheless, under the assumption of completely homogeneous consumers in the previous sections, increases in Λ_0 never produced ex ante welfare losses. In this section, we establish the surprising result that when consumers are heterogeneous, increased opportunities for social learning can *strictly hurt* some consumers and 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 M_0^p of patient types with discount rate $r_p > 0$ and a mass M_0^i of impatient types with discount rate $r_i > r_p$. To simplify the analysis we assume that $\varepsilon = 0$ and $p_0 > 1/2$, although our arguments easily extend to the case where $\varepsilon > 0$. Because our purpose is simply to construct an example illustrating the possibility of welfare loss, we restrict attention to a perfect bad news setting.

To construct our example, we begin by examining equilibria in which only types with discount rate r_p exist in the economy. Recall from Section 5.1 that for any discount rate $r > 0$, we can define the function Λ_r^* implicitly for every $p \in (\frac{1}{2}, \frac{\rho+r}{\rho+2r})$ by

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

Then by Theorem 5.1, whenever $p_0 < \frac{\rho+r_p}{\rho+2r_p}$ and $\hat{\lambda}M_0^p > \lambda M_0^p > \Lambda^*_{r_p}(p_0)$, then in the game consisting solely of consumers of type r_p, the two equilibria corresponding to information structures λ and $\hat{\lambda}$ both feature initial regions of partial adoption.

The main argument in the construction of our example is to consider heterogeneous economies where the mass M_0^i of impatient types is very small, holding fixed the mass of patient types at M_0^p. More specifically, we show that when the mass of impatient types is sufficiently small, the equilibrium behavior of the patient types in both equilibria (under information process $\hat{\lambda}$ and λ) approximates the behavior in the corresponding equilibria when only patient types are present. Then using arguments about the properties of equilibria in the game with only patient types, in particular the non-monotonicity result for learning established in Proposition 5.11, we can obtain the following result:

Theorem 7.1. *Suppose* $0 < r_p < r_i < +\infty$. *Fix* $M_0^p > 0$ *and* $\hat{\lambda} > \lambda > 0$ *such that* $\hat{\lambda}M_0^p > \lambda M_0^p > \Lambda^*_{r_p}(p_0)$. *Then there exists* $\eta > 0$ *such that whenever* $M_0^i < \eta$, $W_0^i(\hat{\lambda}) < W_0^i(\lambda)$ *and* $W_0^p(\hat{\lambda}) = W_0^p(\lambda)$. *Thus, whenever* $M_0^i < \eta$, *the ex ante payoff profile in the* λ*-equilibrium Pareto-dominates the ex ante payoff profile in the* $\hat{\lambda}$*-equilibrium and*

$$M_0^i W_0^i(\hat{\lambda}) + M_0^p W_0^p(\hat{\lambda}) < M_0^i W_0^i(\lambda) + M_0^p W_0^p(\lambda).$$

Here we sketch the main arguments of the theorem. Consider first an economy consisting only of types with discount rate r_p: $M_0^i = 0$ and $M_0^p > 0$. If $\hat{\lambda}M_0^p > \lambda M_0^p > \Lambda^*_{r_p}(p_0)$, then the two equilibria corresponding to information structures λ and $\hat{\lambda}$ both feature initial regions of partial adoption. Thus $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$.

Now consider the payoffs that a hypothetical type r_i (even though such a type does not exist in this economy) would obtain if he were to behave optimally when faced with the flow of information generated in each of these equilibria. Because an optimal strategy (there will be a continuum of optimal strategies) of a consumer of type r_p is to adopt upon first opportunity absent breakdowns, it is straightforward to show that an optimal strategy of such a hypothetical type r_i would also be to adopt upon first opportunity.

Given this, the payoffs of the hypothetical type r_i in the two equilibria are given by the following two expressions:

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

Furthermore, patient types begin in a partial adoption phase in both equilibria:

$$2p_0 - 1 = W_0^p(\hat{\lambda}) = \int_0^\infty \rho e^{-(r_p+\rho)\tau} \frac{p_0}{p_\tau^{\hat{\lambda}}} \left(2p_\tau^{\hat{\lambda}} - 1\right) d\tau$$
$$2p_0 - 1 = W_0^p(\lambda) = \int_0^\infty \rho e^{-(r_p+\rho)\tau} \frac{p_0}{p_\tau^{\lambda}} \left(2p_\tau^{\lambda} - 1\right) d\tau.$$

Recall from Proposition 5.11 that there exists $\bar{t} > t^* := t_2^*(\lambda)$ such that $p_\tau^{\hat{\lambda}} = p_\tau^{\lambda}$ for all $\tau \le t^*$, $p_\tau^{\hat{\lambda}} < p_\tau^{\lambda}$ 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, because $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$, the deceleration in learning followed by a later acceleration must balance out exactly so that the patient type r_p obtains the same ex ante payoff under λ and $\hat{\lambda}$. But then these adjustments must strictly hurt the less patient hypothetical type r_i, because relative to type r_p, type r_i weights the losses due to the slow down of learning more heavily than the benefits that arrive at later times.[47]

To complete the proof, we can show that even when $M_0^i > 0$, as long as M_0^i 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 is the result of a simple continuity argument. The second equality comes from the fact that even upon perturbing M_0^i slightly, 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 ρ is finite, because of a natural delay in adoption, the impatient types may not receive any adoption opportunities for a long time. As a result, if an impatient type obtains his first adoption opportunity late in the game, then the information available at that point in time would be strictly lower under the equilibrium with information process $\hat{\lambda}$ than λ. This decrease in information (due to increased free-riding of the patient types) when

[47] A formal argument is provided in Appendix Section H.

impatient types receive adoption opportunities late in the game 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 they would adopt at exactly time 0 in both the λ and $\hat{\lambda}$-equilibria. Thus the example here illustrates an interesting interaction between heterogeneity and delays due to limited opportunities for adoption.

8 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, and highlights the way in which these incentives vary across different informational environments. The possibility of free-riding in the form of partial adoption is found to be particularly important, because it casts doubt on the the received wisdom that the recent internet-driven surge in opportunities for social learning should speed up learning and benefit consumers. Owing to the advantages of continuous time and Poisson learning, the model is very tractable, yielding closed-form expressions for key quantities and allowing us to compute numerous comparative statics.

We briefly discuss some questions for ongoing and future research that could build on the modeling framework and techniques developed in this paper. Current work in progress relaxes the assumption of perfect Poisson learning to allow for signals that while indicative of bad (respectively good) quality are *not* conclusive. Serving as a robustness check for our results obtained under perfect Poisson learning, preliminary results suggest that many key qualitative features are preserved, for example the possibility of partial adoption regions in bad news markets (which once again coincide with convex growth in adoption levels) as well as the absence of such regions under good news learning. In addition, the extension to imperfect Poisson learning introduces interesting new questions that cannot be studied when signals are conclusive. For example, when $\varepsilon = 0$, then under imperfect (but not under perfect) bad news it is possible for good innovations to fail, because even good products can generate strings of breakdowns that might permanently halt adoption. This suggests investigating the "fragility" of the adoption process as a function of parameters such as the initial market belief and the relative rates at which bad and good products generate breakdowns.

Further work in progress relaxes the assumption that signals are public. To see the idea, suppose that learning is social, but that signals derived from past adopters' experiences are observed privately and independently (at rate λN_t) by each remaining consumer, instead of publicly and simultaneously as in the model in this paper. This captures the intuition of *decen-*

tralized social learning, for example when consumers frequent different blogs and social media platforms. Assuming that at any time consumers make inferences based only on their own private signals and on the expected number of adopters in the population, another interesting difference between bad news and good news markets emerges: Under bad news, a consumer who privately observes a breakdown will never adopt the innovation in the future and hence will never generate any signals himself; this has a *dampening* effect on the production of information in the economy and reduces free-riding incentives. By contrast, consumers who privately observe breakthroughs under good news will adopt the innovation at their next opportunity, thus *amplifying* information production in the future and increasing free-riding incentives. This difference has important implications for aggregate adoption dynamics and for the impact of increased opportunities for social learning.

Finally, moving beyond our focus in this paper on the purely informational aspects of the problem, one could explore the implications of incorporating consumer heterogeneity and pricing motives into the model. As we saw in Section 7, heterogeneity can interact in interesting ways with informational free-riding incentives, sometimes rendering increases in the potential for social learning strictly harmful. A more general characterization of this interaction under more complex distributions of consumer heterogeneity appears challenging but desirable. As for pricing, assume that the innovation is sold by a forward-looking monopolist who does not have any influence on the quality of the innovation and has access to exactly the same public information as consumers, but can influence the endogenous production of information via the price. As a simple first step, we could restrict the monopolist to setting a single fixed price and compute comparative statics on this price and on welfare under increased opportunities for social learning. More challengingly, we could allow the monopolist to commit to a time path of prices, examining for instance how the fact that information is generated endogenously by consumer purchases affects the monopolist's incentives for intertemporal price-discrimination relative to the well-known complete information results of Stokey (1979). We leave these two topics as interesting avenues for future research.

References

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

Baptista, R. (1999): "The diffusion of process innovations: a selective review," *International Journal of the Economics of Business*, 6, 107–129.

Bergemann, D. and J. Välimäki (1997): "Market Diffusion with Two-Sided Learning," *RAND Journal of Economics*, 773–795.

——— (2000): "Experimentation in Markets," *Review of Economic Studies*, 67, 213–234.

Board, S. and M. Meyer–ter–Vehn (2013): "Reputation for Quality," *Econometrica*, forthcoming.

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

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 (2013): "Optimal Design for Social Learning," Mimeo.

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

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.

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.

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.

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

Jensen, R. (1982): "Adoption and Diffusion of an Innovation of Uncertain Profitability," *Journal of Economic Theory*, 27, 182–193.

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

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.

——— (2013): "Breakdowns," Mimeo.

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

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

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.

McLaren, C. H. and C. A. DePaolo (2009): "Movie Data," *Journal of Statistics Education*, 17.

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.

Narayanan, V. K. and G. C. O'Connor (2010): *Encyclopedia of Technology and Innovation Management*, Wiley-Blackwell.

Persons, J. C. and V. A. Warther (1997): "Boom and Bust Patterns in the Adoption of Financial Innovations," *Review of Financial Studies*, 10, 939–967.

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

Stokey, N. L. (1979): "Intertemporal Price Discrimination," *The Quarterly Journal of Economics*, 93, 355–371.

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

A Preliminary Mathematical Tools for Equilibrium Analysis

A.1 General Properties of the Value to Waiting

Throughout this section N denotes an equilibrium adoption flow, with associated value to waiting W_t^N and no-news posterior p^N. We we establish some basic mathematical properties of the value to waiting W^N corresponding to any equilibrium adoption flow N.

Lemma A.1. *Let N be an equilibrium flow of adopters. Then W_t^N is continuous in t.*

Proof. This is immediate from Lemma 2.4. Note that W_t^N can be written as:

$$W_t^N = \int_t^\infty h(\tau) d\tau$$

for some $h \in L^1[0,\infty) \cap L^\infty[0,\infty)$. Then it is immediate that W_t^N is continuous in t. ■

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

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

Proof. By Lemma A.1, W_t^N must be continuous in t. Because $2p_t^N - 1$ is continuous in t, there exists some $\nu > 0$ such that $W_\tau^N < 2p_\tau^N - 1$ for all $\tau \in (t-\nu, t+\nu)$. This means that $N_\tau = \rho \bar{N}_\tau$ for all $\tau \in (t-\nu, t+\nu)$ and so N_τ must be continuous at all $\tau \in (t-\nu, t+\nu)$. From Lemma 2.4, W_τ^N can be rewritten for all $\tau \in (t-\nu, t+\nu)$ as

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

From this it is easy to see that W_τ^N is continuously differentiable with respect to τ for all $\tau \in (t-\nu, t+\nu)$.

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

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

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

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

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

as claimed. ■

We can prove a similar lemma for the case in which the equilibrium value to waiting is strictly above the payoff to adopting today.

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

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

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

Hence, W_τ^N satisfies

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

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

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

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

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

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

A.2 Special Properties of the Value to Waiting under PBN

Here we focus on learning via perfect bad news. By Equation 1, an upper bound on the no-news posterior is given by:

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

We now show that absent breakdowns, this posterior is attained in the limit.

Lemma A.4. *Let N be an equilibrium under PBN. Suppose that $\varepsilon > 0$ or $p_0 > 1/2$. Then $p_t^N \to \mu(\varepsilon,\Lambda_0,p_0)$ and $W_t^N \to \frac{\rho}{\rho+r}(2\mu(\varepsilon,\Lambda_0,p_0)-1)$ as $t\to\infty$.*

Proof. Consider first the case in which $\varepsilon > 0$. Then trivially $p_t^N \to 1$ as $t \to \infty$. So for any $\nu > 0$, there exists some t^* such that whenever $t > t^*$, then $1 - p_t^N < \nu$.

Then we can produce upper and lower bounds on W_t^N:

$$\frac{\rho}{\rho+r}(1-\nu) - \frac{\rho}{\rho+r}\nu < \frac{\rho}{\rho+r}\left(2p_t^N - 1\right) \le W_t^N \le \frac{\rho}{\rho+r}.$$

Since this is true for any $\nu > 0$, it follows 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 $W_t^N \le 2p_t^N - 1$ 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 $s \ge t$, 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 $N^* := \lim_{t\to\infty}\int_0^t N_s ds = \sup_t \int_0^t N_s ds \le \bar{N}_0$. 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 $2p_t^N - 1 \ge W_t^N$

for all t, we can then rewrite the value to waiting at time t as:

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

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

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

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 $p^* = \mu(\varepsilon, \Lambda_0, p_0)$. ■

B Quasi-Single Crossing Property for Equilibrium Incentives

B.1 Proof of Theorem 4.1 under Perfect Good News

From now on we drop the superscript N from W and p.

Proof. The proof consists of two steps. In the first step, we show that whenever $W_t = 2p_t - 1$, then $W_\tau \geq 2p_\tau - 1$ for all $\tau \geq t$. In the second step, we show that whenever $W_t > 2p_t - 1$, then $W_\tau > 2p_\tau - 1$ for all $\tau > t$.

Step 1: 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:

$$\lim_{s\downarrow s^*} \dot{W}_s - 2\dot{p}_s = r(2p_{s^*} - 1) + p_{s^*} \left(\varepsilon_G + \lambda_G N_{s^*} \right) \frac{r}{\rho+r} > 0.$$

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: Assume $W_t > 2p_t - 1$ at some t and suppose for a 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 $s \in (t, s^*)$. Then by Lemma A.3 the left-hand derivative of $W_s - (2p_s - 1)$ at s^* exists and is given by:

$$\lim_{s \uparrow s^*} \dot{W}_s - 2\dot{p}_s = r(2p_{s^*} - 1) + p_{s^*} \frac{r}{\rho + r} \varepsilon_G > 0.$$

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

B.2 Proof of Theorem 4.1 under Perfect Bad News

Proof. The proof consists of two steps. In the first step, we show that whenever $W_t = 2p_t - 1$, then $W_\tau \leq 2p_\tau - 1$ for all $\tau \geq t$. In the second step, we show that whenever $W_t < 2p_t - 1$, then $W_\tau < 2p_\tau - 1$ for all $\tau > t$.

Step 1: Suppose $W_t = 2p_t - 1$ at some time t and suppose for a contradiction that at some time $s' > t$ we have $W_{s'} > 2p_{s'} - 1$. Then because $W_t \to \frac{\rho}{\rho + r}(2\mu(\varepsilon, \Lambda, p_0) - 1) < 2\mu(\varepsilon, \Lambda, p_0) - 1$ by Lemma A.4, there exists $\underline{s} < \overline{s}$ such that $W_{\underline{s}} = 2p_{\underline{s}} - 1$, $W_{\overline{s}} = 2p_{\overline{s}} - 1$, and $W_s > 2p_s - 1$ for all $s \in (\underline{s}, \overline{s})$. By Lemma A.3, we have the following two limits:

$$\lim_{s \downarrow \underline{s}} \dot{W}_s = (r + (1 - p_{\underline{s}})\varepsilon)(2p_{\underline{s}} - 1).$$
$$\lim_{s \uparrow \overline{s}} \dot{W}_s = (r + (1 - p_{\overline{s}})\varepsilon)(2p_{\overline{s}} - 1).$$

Also, as usual

$$\lim_{s \downarrow \underline{s}} \frac{d}{ds}(2p_s - 1) = 2p_{\underline{s}}(1 - p_{\underline{s}})\varepsilon$$
$$\lim_{s \uparrow \overline{s}} \frac{d}{ds}(2p_s - 1) = 2p_{\overline{s}}(1 - p_{\overline{s}})\varepsilon.$$

In order that $W_s > 2p_s - 1$ for all $s \in (\underline{s}, \overline{s})$, we need:

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

Rearranging we get:

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

But this is impossible given that $p_{\overline{s}} > p_{\underline{s}}$. This completes the proof of Step 1.

Step 2: Suppose that $W_t < 2p_t - 1$ and suppose for a contradiction that there exists some $s' > t$ such that $W_{s'} \geq 2p_{s'} - 1$. Define

$$\underline{s} = \inf\{s' > t : W_{s'} \geq 2p_{s'} - 1\}.$$

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 $s > \overline{s}$. By Lemma A.2, we have the following two limits:

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

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 $p_s \leq p_{s^*}$ 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 $W_s < 2p_s - 1$ and $N_s = \rho\bar{N}_s$ yields

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

Rearranging we obtain:

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

But by Lemma A.2, the right-hand side is precisely the derivative $\frac{d}{ds}(2p_s - 1) - \dot{W}_s$. This implies that for all $s \in [t, s^*)$, $2p_s - 1 - W_s$ is strictly increasing, contradicting continuity and the fact that $2p_{s^*} - 1 = W_{s^*}$. This concludes the proof of Step 2. ■

C Equilibrium Uniqueness and Characterization

C.1 Equilibrium Characterization under Perfect Bad News

In this section, we do not impose Conditions 5.2 or 5.3. Recall that $p^* := \min\{\overline{p}, p^\sharp\}$, where

$$\overline{p} := \frac{\varepsilon + r}{\varepsilon + 2r},$$
$$p^\sharp := \frac{\rho + r}{\rho + 2r}.$$

Recall the definition of $G : [0,1] \times \mathbb{R}_+ \to \mathbb{R}$:

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

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, recall the definition of $\Lambda^* : (0,1) \to \mathbb{R}_+ \cup \{+\infty\}$:

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

The proof of Theorem 5.1 proceeds in three steps. Assuming that N is an equilibrium, we show in Lemma C.1 that if $t_1^* < t_2^*$, then the evolution of adoption behavior on (t_1^*, t_2^*) is uniquely pinned down by an ODE. We next prove Lemma 5.4, which provides a characterization of t_1^* and t_2^* in terms of (p_t, Λ_t). Given these two steps uniqueness is clear. Finally, we check feasibility in Lemma C.4, proving equilibrium existence.

C.1.1 Characterization of Adoption between t_1^* and t_2^*

Lemma C.1. *Suppose N is an equilibrium with no-news posterior p_t. Suppose that $t_1^* < t_2^*$. Then at (almost) all times $t \in (t_1^*, t_2^*)$,*

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

Proof. Note that because $2p_t - 1 = W_t^N$ at all $t \in (t_1^*, t_2^*)$ and p_t is weakly increasing, W_t^N and p_t are differentiable almost everywhere (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 $t \in (t_1^*, t_2^*)$:

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

Then for all t at which W_t^N and p_t are differentiable, we obtain:

$$\begin{aligned} \dot{W}_t^N &= \left(r + (\varepsilon + \lambda N_t)(1-p_t)\right) W_t^N \\ 2\dot{p}_t &= 2p_t(1-p_t)(\varepsilon + \lambda N_t). \end{aligned}$$

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

$$\dot{W}_t^N = 2\dot{p}_t.$$

This means that for almost all $t \in (t_1^*, t_2^*)$:

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

■

A direct corollary of the above lemma is the following:

Corollary C.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:

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

C.1.2 Proof of Lemma 5.4

We now prove a more general version of Lemma 5.4 in which we replace $\overline{p}$ in Lemma 5.4 with p^*.

Lemma C.3. *Let N be an equilibrium with corresponding no-news posterior $p_{t\geq 0}$ and threshold times t_1^* and t_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). $t_2^* = \inf\{t : \Lambda_t < \Lambda^*(p_t)\}$*; and*

(ii). $t_1^* = \min\{t_2^*, \sup\{t : p_t < p^*\}\}$.[48]

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 $t_2^* < +\infty$ and we must also have that p_t is strictly increasing for all $t > 0$.

For the first bullet point, note that by definition of t_2^* and by Theorem 4.1, we have $2p_t - 1 > W_t = G(p_t, \Lambda_t)$ for all $t > t_2^*$. This implies that $\Lambda_t < \Lambda^*(p_t)$ for all $t > t_2^*$. Moreover, if $0 < t_2^*$, then by continuity we must have $2p_{t_2^*} - 1 = W_{t_2^*} = G(p_{t_2^*}, \Lambda_{t_2^*})$ and so $\Lambda_{t_2^*} = \Lambda^*(p_{t_2^*})$. In this case, because Λ_s is decreasing in s and p_s is strictly increasing in s and $\Lambda^*(p)$ is increasing in p, we must have $\Lambda_s \geq \Lambda^*(p_s)$ for all $s < t_2^*$. This establishes the first bullet point.

For the second bullet point, it suffices to prove the following three claims:

(a) If $t_2^* > 0$, then $p_{t_2^*} < p^\sharp$.

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

(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 $W_t > 2p_t - 1$. Then by Lemma A.3 and because $W_{t_1^*} = 2p_{t_1^*} - 1$, we must 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 C.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 it is easy to see that $p_t \equiv p_0$ for all t. Thus, $t_2^* = +\infty = \inf\{t : \Lambda_t < \Lambda^*(p_0) = 0\}$. Also, if $p_0 < \frac{1}{2}$, then $t_1^* = +\infty = \sup\{t : p_t < p^* = \frac{1}{2}\}$; and if $p_0 = \frac{1}{2}$, then $t_1^* = 0 =: \sup\{t \geq 0 : p_t < p^* = \frac{1}{2}\}$. ■

With these lemmas, it is immediate that if an equilibrium exists, then it must take the form of the adoption flow given by Equation 6 inTheorem 5.1. Moreover, it is easy to see that given initial parameters, Equation 6 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.

C.1.3 Feasibility

It remains to check feasibility of the adoption flow implied by Equation 6 inTheorem 5.1. Note that feasibility is non-trivial only at times $t \in (t_1^*, t_2^*)$.

Lemma C.4. *Suppose $N_{t\geq 0}$ is an adoption flow satisfying Equation 6 in Theorem 5.1 such that $t_1^* < t_2^*$. Then for all $t \in (t_1^*, t_2^*)$,*

$$N_t \leq \rho\Lambda_t.$$

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^*).[49]

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 $t \in (t_2^* - \nu, t_2^*)$. Then 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, $2p_t - 1 < G(p_t, \lambda\bar{N}_t)$ for all $t < t_2^*$.

Since at all times $t \in (t_2^* - \nu, t_2^*)$ we have $N_t > \rho\bar{N}_t$, this implies that

$$W_t^N > G(p_t, \lambda\bar{N}_t) > 2p_t - 1.$$

But this is a contradiction since we already checked that the described adoption flow satisfies the condition that $W_t^N = 2p_t - 1$ for all $t \in (t_1^*, t_2^*)$. ■

C.2 Equilibrium Characterization under Perfect Good News

Theorem 6.1 follows readily from Lemma 6.2 and Lemma 6.3. Lemma 6.3 was proved in the text. It remains to prove Lemma 6.2.

Proof of Lemma 6.2: Suppose for a contradiction that $t_1^* < t_2^*$. From the definition of these cutoffs, we must have $2p_t - 1 = W_t$ 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:

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

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

Note that we must have $p_t \geq \frac{1}{2}$ on (t_1^*, t_2^*), since W_t is bounded below by 0. Given that we assume that either $\varepsilon = 0$ or $p_0 \neq \frac{1}{2}$, this means that either $\varepsilon > 0$ or $p_t > \frac{1}{2}$ for t sufficiently close

[49] This is only true if either $\varepsilon > 0$ or $p_0 > \frac{1}{2}$. If $\varepsilon = 0$ and $p_0 = \frac{1}{2}$, then $N_t = 0$ for all t and $t_1^* = 0 < t_2^* = +\infty$. But in this case feasibility is immediate.

to t_1^*. Then it follows that for sufficiently small Δ

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

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

D Violation of Condition 5.2 under Perfect Bad News

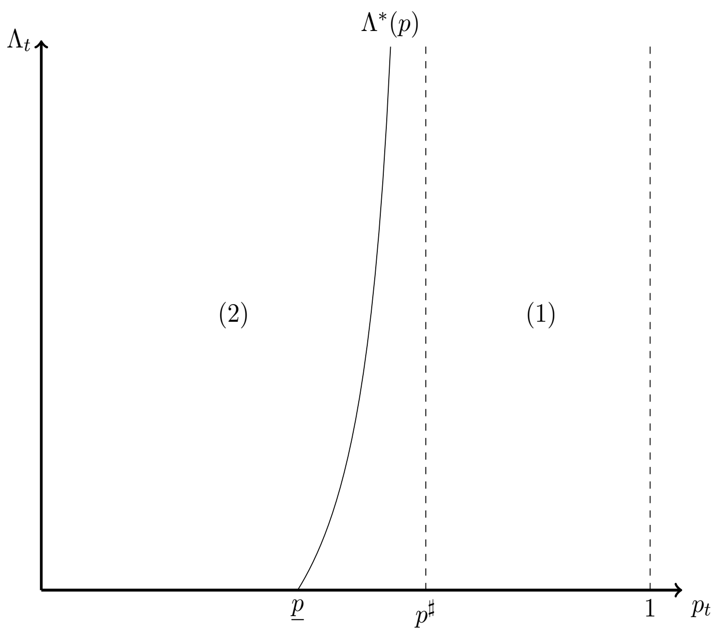

Figure 10: Partition of (p_t, Λ_t) when $\varepsilon \geq \rho$

In this section, we discuss the case in which $\rho \geq \varepsilon$. We saw in Theorem 5.1 that the characterization theorem holds even when Condition 5.2 is violated.

In this case because $\Lambda^*(p) = +\infty$ for all $p > p^*$, we have:

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

Note that now partial adoption never occurs and the unique equilibrium reduces to all-or-nothing adoption.

As a result the saturation effect discussed in Section 5 is no longer present and welfare always strictly increases in response to an increase in opportunities for social learning:

Proposition D.1. *Fix $r > 0$ and $p_0 \in (0,1)$ and suppose that $\varepsilon \geq \rho > 0$. Then W_0 is strictly increasing in Λ_0.*

E Inefficiency of Equilibria

E.1 Inefficiency under PBN

Proof of Proposition 5.8: From Proposition 3.3, recall that

$$p^s = \frac{K(\Lambda_0)}{K(\Lambda_0) + \frac{\rho}{r+\rho}\frac{r}{r+\varepsilon}},$$

where

$$K(\Lambda_0) = \int_0^\infty \rho e^{-(r+\rho)\tau} e^{-\varepsilon\tau - \Lambda_0(1-e^{-\rho\tau})} d\tau.$$

Note also that

$$K(\Lambda_0) < \frac{\rho}{r+\varepsilon+\rho}$$

which then implies that

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

Finally observe from the proof of Lemma 5.4 that $p_{t_1^*} \geq \underline{p}$.

If $\Lambda_0 > \Lambda^*(\overline{p})$, either $t_1^* > 0$ or $t_2^* > 0$. In the first case, adoption begins too late since $p_{t_1^*} \geq \underline{p} > p^s$ and therefore equilibrium is inefficient. If on the other hand, $t_1^* = 0 < t_2^*$, then again because $p^s < p_{t_1^*}$, adoption is too slow initially since consumers only partially adopt between t_1^* and t_2^*. Thus again equilibrium is inefficient.

On the other hand, if $\Lambda_0 \leq \Lambda^*(p_0)$, then equilibrium is efficient since both the cooperative benchmark and equilibrium prescribe that absent breakdowns all consumers adopt whenever given an opportunity. ■

F Comparative Statics under PBN

F.1 Saturation Effect: Proof of Proposition 5.9

Throughout Section F we impose Conditions 5.2 (so that $p^* = \overline{p}$) and 5.3 as in the text. We first prove Lemma 5.6.

Proof of Lemma 5.6: Let $\overline{\Lambda}_0 := \max\{\Lambda^*(p_0), \Lambda^*(\overline{p})\}$. 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 5.4, 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 5.4, 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 5.4, 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. ∎

Proof of Proposition 5.9: For the first bullet point, consider any $\Lambda_0^1 < \Lambda_0^2 \leq \overline{\Lambda}_0$ with corresponding threshold 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 5.6, 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 5.4 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 $t > 0$. Then by Lemma 2.4, $W_0^1 < W_0^2$.

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 $W_0^1 < W_0^2$. This proves the first bullet point.

To prove the second bullet point, suppose that $\Lambda_0^2 > \Lambda_0^1 > \overline{\Lambda}_0$. By Lemma 5.6, we then have $t_1^i < t_2^i$ for $i = 1, 2$. Moreover, by the proof of Lemma 5.4, we have $\max\{p_0, \overline{p}\} = p_{t_1^1}^1 = p_{t_1^2}^2$. Because $N_t^i = N_t^i = 0$ 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 once again,

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

for $i = 1, 2$, whence $W_0^1 = W_0^2$. ∎

F.2 Learning Dynamics

F.2.1 Proof of Lemma 5.10

Proof of Lemma 5.10. Suppose that $\hat{\Lambda}_0 > \Lambda_0 > \Lambda^*(p_0)$. Recall that we are assuming $p^\sharp > p_0 > \overline{p}$ so that $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}) = 0$ and $\overline{\Lambda}_0 = \Lambda^*(p_0)$. Then by Lemma 5.6, 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 C.2 implies that at all times $t < t_2^*$, we have $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0} = p_t$. By Lemma C.1, this implies that for all $t < t_2^*$,

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

To prove the first bullet point, note that Equation 11 implies that

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

By Lemma 5.4 and because $p_{t_2^*}^{\Lambda_0} = p_{t_2^*}^{\hat{\Lambda}_0}$, this 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 $t < t_2^* = t_2^*(\Lambda_0)$, proving the second bullet point. ■

F.2.2 Proof of Proposition 5.11

Proof. Clearly $p_t^{\Lambda_0}$ is strictly increasing 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 $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. By Lemma 5.10, $t^* := t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$, $\lambda N_t = \hat{\lambda}\hat{N}_t$, 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 Λ_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,

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

which by Lemma 5.4 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 $t > t^*$. By Lemma 2.4, this implies

$$W_{t^*}^{\hat{\Lambda}_0} > W_{t^*}^{\Lambda_0}.$$

But this is a contradiction, because we have that

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

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 of the proposition.

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 note that 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. This implies that

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

which as above implies that

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

Then it is again obvious that$\hat{\lambda}\hat{N}_t > \lambda N_t$ 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. ■

F.3 Adoption Behavior

Proof of Proposition 5.12: 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 5.3). Let $t^* = t_2^*(\Lambda_0)$. By Lemma 5.10, $\lambda N_t = \hat{\lambda}\hat{N}_t$ 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 $t > 0$. ■

Proof of Proposition 5.13. We first prove the proposition when we increase the information structure from λ to $\hat{\lambda} > \lambda$ holding fixed $\bar{N}_0$. Given this, proving the theorem for arbitrary changes from Λ_0 to $\hat{\Lambda}_0$ is straightforward.

Let N and $\hat{N}$ be the equilibrium under λ 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 and $t > 0$, consider the following map:

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

Note that for any $t > 0$, the above is strictly decreasing in λ. 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 Inequality (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 Proposition 5.12, $\int_0^\tau N_s ds \geq \int_0^\tau \hat{N}_s ds$ 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

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

This contradicts $\int_0^t N_s ds \geq \int_0^t \hat{N}_s ds$ as found in Proposition 5.12.

Having shown the above, consider any change from $\Lambda_0 = \lambda \bar{N}_0$ to $\hat{\Lambda}_0 = \hat{\lambda} \hat{\bar{N}}_0 > \Lambda_0$. Then there exists $\lambda^* > \lambda$ such that $\hat{\Lambda} = \lambda^* \bar{N}_0$. Let $N*$ be the equilibrium associated with the pair $(\lambda^*, \bar{N}_0)$. By Lemma 5.5, unique equilibrium for the pair $(\hat{\lambda}, \hat{\bar{N}}_0)$ satisfies $\hat{N}_t = (\lambda^*/\hat{\lambda}) N_t^*$. But then the above argument implies:

$$\begin{aligned} A_t(\Lambda, B) = \mathbb{E}\left[\int_0^t \frac{N_s}{\bar{N}_0} ds\right] > \mathbb{E}\left[\int_0^t \frac{N_s^*}{\bar{N}_0} ds\right] &= \mathbb{E}\left[\int_0^t \frac{\lambda^*}{\hat{\lambda}\hat{\bar{N}}_0} \frac{\hat{\lambda}\hat{\bar{N}}_0}{\lambda^* \bar{N}_0} N_s^* ds\right] \\ &= \mathbb{E}\left[\int_0^t \frac{\hat{N}_s}{\hat{\bar{N}}_0} ds\right] = A_t(\hat{\Lambda}, B). \end{aligned}$$

■

G Comparative Statics under PGN

G.1 Adoption Behavior

The only statement that was not proved in the text is: $A_t(\hat{\Lambda}_0, G) < A_t(\Lambda_0, G)$ for all $t > t^*(\hat{\Lambda}_0)$, as claimed in the first bullet of Proposition 6.7.

Proof. When $\varepsilon = 0$, the statement is trivial, so assume that $\varepsilon > 0$. The claim is also obvious for all $t \leq t^*(\Lambda_0)$ since adoption occurs at the maximal rate under Λ_0 whereas under $\hat{\Lambda}_0$, adoption ceases at times $t \in (t^*(\hat{\Lambda}_0), t^*(\Lambda_0))$ absent breakthroughs.

So assume that $t > t^*(\Lambda_0)$. Recall that the cutoff posterior p^* at which adoption ceases is unchanged upon a change from Λ_0 to $\hat{\Lambda}_0$. Then expected adoption up to time t for any $\Gamma \in [\Lambda_0, \hat{\Lambda}_0]$ can be expressed in the following manner:

$$A_t(\Gamma, G) = \pi^* \left(1 - e^{-\rho t}\right) + (1 - \pi^*) \left((1 - e^{-\rho t^*(\Gamma)}) + e^{-\rho t^*(\Gamma)} \int_{t^*(\Gamma)}^{t} \varepsilon e^{-\varepsilon(\tau - t^*(\Gamma))} \left(1 - e^{-\rho(t-\tau)}\right) \right)$$

where

$$(1 - \pi^*) = \frac{1 - p_0}{1 - p^*}.$$

Now for a fixed t, consider the function:

$$t^* \mapsto \pi^* \left(1 - e^{-\rho t}\right) + (1 - \pi^*) \left((1 - e^{-\rho t^*}) + e^{-\rho t^*} \int_{t^*}^{t} \varepsilon e^{-\varepsilon(\tau - t^*)} \left(1 - e^{-\rho(t-\tau)}\right) \right).$$

Then a straightforward computation yields that the derivative of the above map with respect to any $t^* < t$ is $\rho e^{-(\varepsilon-\rho)t^*} e^{-\varepsilon t} > 0$. Thus, the map is strictly increasing in t^* for all $t^* < t$. Because $t^*(\Gamma)$ is strictly decreasing in Γ, it follows that for all $t > t^*(\Gamma)$, $A_t(\Gamma, G)$ is strictly decreasing for all $\Gamma \in [\Lambda_0, \hat{\Lambda}_0]$. This proves the claim. ∎

H Heterogeneous Discount Rate Example

First we show the following basic mathematical fact.

Lemma H.1. *Let $\bar{t} > t^*$ and suppose that f and g are real-valued functions 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*

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

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. ∎

As in the main text, assume that $\hat{\lambda} M_0^p > \lambda M_0^p > \Lambda_{r_p}^*(p_0)$ and that $p_0 > 1/2$ and $\varepsilon = 0$. Then modifying the arguments from the proof of Theorem 5.1, it is easy to show that when M_0^i is sufficiently small, the unique equilibrium under both information processes λ, $\hat{\lambda}$ will be such that the impatient type adopts immediately upon opportunity at all times absent breakdowns and the patient type only partially adopts until some time $t^* > 0$ after which he switches to immediate adoption:[50]

$$\begin{aligned} \gamma N_t^i &= \rho M_t^i, \\ \gamma N_t^p &= \begin{cases} \frac{r_p(2p_t - 1)}{1 - p_t} - \gamma \rho M_t^i & \text{if } t < t^*(\gamma) \\ \gamma \rho M_t^p & \text{if } t \geq t^*(\gamma) \end{cases} \end{aligned}$$

for $\gamma \in \{\lambda, \hat{\lambda}\}$.

Then using arguments analogous to those in Lemma 5.10, we can show that $t^*(\lambda) < t^*(\hat{\lambda})$. Furthermore an analogue of Proposition 5.11 shows that there must exist some $\bar{t} > t^*(\lambda)$ such that

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

[50] The full proof of the modification is available upon request. Here we use a standard argument that shows that whenever the impatient type weakly prefers to wait, then the patient type must strictly prefer to wait. Similarly, if the patient type weakly prefers to adopt then the impatient type must strictly prefer to adopt.

Then using Lemma H.1, the proof follows along the lines illustrated in the main text. This proves Theorem 7.1.

I Cooperative Benchmark

I.1 Perfect Bad News

To prove the all-or-nothing property of the optimal policy, we write the Hamilton-Jacobi-Bellman (HJB) equation. Note that there are two state variables, p and $\bar{N}$.

$$rV(p,\bar{N}) = \max_{0\leq N\leq \rho\bar{N}} (2p-1)\,N + D_pV(p,\bar{N})p(1-p)(\varepsilon+\lambda N) - D_{\bar{N}}V(p,\bar{N})N - (1-p)(\varepsilon+\lambda N)V(p,\bar{N}).$$

Since the right hand side is linear in N, it is optimal to always choose either $N = 0$ or $N = \rho\bar{N}$.

To see that the optimal policy must be a cutoff strategy, define

$$\Pi(p,\bar{N}) := (2p-1) + D_pV(p,\bar{N})p(1-p)\lambda - D_{\bar{N}}V(p,\bar{N}) - \lambda(1-p)V(p,\bar{N})$$

and note that whenever $\Pi(p,\bar{N}) < 0$, then

$$rV(p,\bar{N}) = D_pV(p,\bar{N})\varepsilon p(1-p) - (1-p)\varepsilon V(p,\bar{N}) \tag{16}$$

so that this corresponds to the case where setting $N = 0$ is optimal. It then suffices to prove that

$$\Pi(p,\bar{N}) < 0 \Rightarrow \; \Pi(p',\bar{N}) < 0 \; \forall p' < p.$$

To prove this, note first that for every p such that $\Pi(p,\bar{N}) < 0$, there must exist some $p' > p$ such that $\Pi(p',\bar{N}) = 0$. (Otherwise $V(p',\bar{N}) = 0$ for all $p' > p$, which is clearly false.) So it suffices to prove that there cannot exist $\underline{p} < \overline{p}$ such that $\Pi(\underline{p},\bar{N}) = \Pi(\overline{p}) = 0$ and $\Pi(p,\bar{N}) < 0$ for all $p \in (\underline{p},\overline{p})$. Suppose for a contradiction that such an interval $(\underline{p},\overline{p})$ exists. Then ordinary differential equation (16) implies:

$$V(p,\bar{N}) = \left(\frac{p}{\bar{p}}\right)^{\frac{r+\varepsilon}{\varepsilon}} \left(\frac{1-p}{1-\bar{p}}\right)^{-\frac{r}{\varepsilon}} V(\bar{p},\bar{N})$$

for all $p \in (\underline{p},\overline{p})$. Then we can rewrite the expression for $\Pi(p,\bar{N})$ for $p \in (\underline{p},\overline{p})$:

$$\begin{aligned}
\Pi(p,\bar{N}) &= (2p-1) + \frac{r\lambda}{\varepsilon}V(p,\bar{N}) - D_{\bar{N}}V(p,\bar{N}) \\
&= (2p-1) + \frac{r\lambda}{\varepsilon}V(p,\bar{N}) - \left(\frac{p}{\bar{p}}\right)^{\frac{r+\varepsilon}{\varepsilon}} \left(\frac{1-p}{1-\bar{p}}\right)^{-\frac{r}{\varepsilon}} D_{\bar{N}}V(\bar{p},\bar{N}) \\
&= (2p-1) + \left(\frac{p}{\bar{p}}\right)^{\frac{r+\varepsilon}{\varepsilon}} \left(\frac{1-p}{1-\bar{p}}\right)^{-\frac{r}{\varepsilon}} \left(\frac{\lambda r}{\varepsilon}V(\bar{p},\bar{N}) - D_{\bar{N}}V(\bar{p},\bar{N})\right),
\end{aligned}$$

Note that if $\frac{\lambda r}{\varepsilon}V(\bar{p},\bar{N})-D_{\bar{N}}V(\bar{p},\bar{N}) \geq 0$, the last expression is increasing in p and so $\Pi(\underline{p},\bar{N}) < 0$ which is a contradiction.

If instead $\frac{\lambda r}{\varepsilon}V(\bar{p},\bar{N})-D_{\bar{N}}V(\bar{p},\bar{N}) < 0$, then the second term in the last expression is concave. Furthermore, the derivative of the last expression with respect to p at $\overline{p}$ must be weakly positive: If it were strictly negative, then because $\Pi(\overline{p},\bar{N}) = 0$, there would exist some $p \in (\underline{p},\overline{p})$ close to $\overline{p}$ such that $\Pi(p,\bar{N}) > 0$. But if the derivative of $\Pi(p,\bar{N})$ is weakly positive at $\bar{p}$, then by concavity it must be positive throughout $(\underline{p},\overline{p})$. But this again yields the contradiction that $\Pi(\underline{p},\bar{N}) < 0$. This completes the proof.

I.2 Perfect Good News

As in the perfect bad news case, we again write the Hamilton-Jacobi-Bellman equation:

$$\begin{aligned} rV(p,\bar{N}) = &\max_{0\leq N\leq\rho\bar{N}}(2p-1)N + p(\varepsilon+\lambda N)\left(\frac{\rho}{r+\rho}\bar{N} - V(p,\bar{N})\right) \\ &- D_pV(p,\bar{N})p(1-p)\left(\varepsilon+\lambda N\right) - D_{\bar{N}}V(p,\bar{N})N. \end{aligned}$$

Again the right hand side is linear in N and thus the optimal policy always chooses either $N = 0$ or $N = \rho\bar{N}$.

The easiest way to check that an optimal policy exists in cutoff strategies is to simply guess and check that the HJB equation is satisfied by such a strategy. This is straightforward from the social planner policy constructed in Section 3.1.