Essays on Learning, Uncertainty, and Choice

Citation

Permanent link

Terms of Use

Share Your Story

Essays on Learning, Uncertainty, and Choice

A dissertation presented

by

Mira Anna Phyllis Frick

to

The Committee on Degrees in Business Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

April 2015

Dissertation Advisors: *Author:*

Professor Drew Fudenberg   **Mira Anna Phyllis Frick**

Professor Tomasz Strzalecki

Essays on Learning, Uncertainty, and Choice

Abstract

This dissertation presents three independent essays in microeconomic theory.

Motivated by the rise of social media, Chapter 1 (co-authored with Yuhta Ishii) builds a model studying the effect of an economy's potential for social learning on the adoption of innovations of uncertain quality. Provided consumers are forward-looking (i.e. recognize the value of waiting for information), equilibrium dynamics depend non-trivially on qualitative and quantitative features of the informational environment. We identify informational environments that are subject to a *saturation effect*, whereby increased opportunities for social learning slow down adoption and learning and do not increase consumer welfare (possibly even being harmful). We also suggest a novel, purely informational explanation for different commonly observed adoption patterns (S-shaped vs. concave curves).

Chapter 2 (co-authored with Assaf Romm) studies the solution concept $S^\infty W$ (one round of elimination of weakly dominated strategies followed by iterated elimination of strongly dominated strategies) in incomplete-information games. Under complete information, Dekel and Fudenberg (1990) and Börgers (1994) motivate $S^\infty W$ via its connection with "approximate common certainty" (ACC) of admissibility. Under incomplete information, we cast doubt on this connection: $S^\infty W$ corresponds to ACC of admissibility only when this is not accompanied by even the slightest

changes to players' beliefs about states of nature. If we allow for vanishingly small perturbations to beliefs, then $S^{\infty}W$ is a (generally strict) subset of the predicted behavior, which we characterize in terms of a generalization of Hu's (2007) perfect p-rationalizable set.

Motivated by the literature on "choice overload", Chapter 3 studies a boundedly rational agent whose choice behavior admits a *monotone threshold representation*: There is an underlying rational benchmark, corresponding to maximization of a utility function v, from which the agent departs in a menu-dependent manner. The severity of the departure is quantified by a threshold map δ, which is monotone with respect to set inclusion. I axiomatically characterize the model, extending familiar characterizations of rational choice. I classify monotone threshold representations as a special case of Simon's theory of "satisficing", but as strictly more general than both Tyson's (2008) "expansive satisficing" model as well as Fishburn (1975) and Luce's (1956) model of choice behavior generated by a semiorder. I axiomatically characterize the difference, providing novel foundations for these models.

Contents

List of Figures

Acknowledgments

In an illustration of the broad applicability of economic theory, the title of this dissertation could double as a fairly accurate word cloud of my graduate-school experience, were it not for one glaring (if economically under-explored) omission: gratitude.

First and foremost, I am profoundly grateful to my primary advisors Drew Fudenberg and Tomasz Strzalecki. From both I have learned an incredible amount, beginning with their inspiring game theory and decision theory classes that sparked my interest in economic theory in my first year and planted the seeds for Chapters 2 and 3 of this dissertation, and continuing ever since in countless thought-provoking and illuminating meetings. In moments of uncertainty, they have gone above and beyond with generous pep talks and decisive support. And at every important choice on my research path, I couldn't have wished for a better duo of mentors to turn to for (oftentimes near-instantaneous) advice. My debt of gratitude to Jerry Green goes back even further, to when he convinced me to join the Business Economics program in Spring 2010. Since then, he has been very generous with his time and expertise, providing enlightening feedback and counsel whenever I sought it and, most recently, readily agreeing to serve on my dissertation committee.

I am also extremely grateful to have had the good fortune of...

...collaborating with Yuhta Ishii and Assaf Romm (on Chapters 1 and 2 of this dissertation) and Ryota Iijima (on ongoing work), with whom I experienced some of the most rewarding research moments of my graduate-school career.

...spending five years among the vibrant community of theorists and theory-philes in (or passing through) the Cambridge area: I benefited immeasurably from interactions with Nageeb Ali, Attila Ambrus, Pol Antràs, Glenn Ellison, Ben Golub,

To my family

Introduction

This dissertation comprises three independent chapters that span topics in microeconomic theory, including social learning and informational free-riding, epistemic game theory, and choice theory.

Motivated by the rise of social media, Chapter 1 (co-authored with Yuhta Ishii) builds a model studying the effect of an economy's potential for social learning on the adoption of innovations of uncertain quality. In our model, a large population of long-lived consumers faces stochastic opportunities to adopt a new product. Capturing *social learning*, news about the product's quality is generated endogenously, based on the consumption experiences of past adopters. Provided consumers are *forward-looking* (i.e. recognize the value of waiting for information), equilibrium adoption dynamics must resolve the following tension: If too many consumers adopt at any given time, then the expected amount of future information might be so great that all consumers would strictly prefer to wait; conversely, if too few consumers adopt, it might not be worthwhile for anyone to wait. Focusing on a class of Poisson learning processes, we show that this tension depends non-trivially on qualitative and quantitative features of the informational environment. We identify informational environments that are subject to a novel *saturation effect*, whereby increased opportunities for social learning slow down adoption and learning and do not increase consumer welfare (possibly even being harmful). We also suggest a

new, purely informational explanation for different commonly observed adoption patterns (S-shaped vs. concave curves).

Chapter 2 (co-authored with Assaf Romm) studies the solution concept $S^\infty W$ (one round of elimination of weakly dominated strategies followed by iterated elimination of strongly dominated strategies). Dekel and Fudenberg (1990) and Börgers (1994) have proposed this solution concept in the context of complete information games, motivating it by a characterization in terms of "approximate common certainty" of admissibility. We examine the validity of this characterization of $S^\infty W$ in an *incomplete* information setting. We argue that in Bayesian games with a nontrivial state space, the characterization is very sensitive to the way in which approximate common certainty of admissibility is taken to interact with the uncertainty already captured by players' beliefs about the states of nature: We show that $S^\infty W$ corresponds to approximate common certainty of admissibility when this is not allowed to coincide with any changes to players' beliefs about states. If approximate common certainty of admissibility is accompanied by vanishingly small perturbations to beliefs, then $S^\infty W$ is a (generally strict) subset of the predicted behavior, which we characterize in terms of a generalization of Hu's (2007) perfect p-rationalizable set.

Motivated by the literature on "choice overload", Chapter 3 studies a boundedly rational agent whose choice behavior admits a *monotone threshold representation*: There is an underlying rational benchmark, corresponding to maximization of a utility function v, from which the agent departs in a menu-dependent manner. The severity of the departure is quantified by a threshold map δ, which is monotone with respect to set inclusion. This formalizes the intuition that large menus of options may impose a cognitive strain on the agent, adversely impacting his ability to discriminate between the available alternatives. I axiomatically characterize the

2

model, extending familiar characterizations of rational choice. I classify monotone threshold representations as a special case of Simon's theory of "satisficing", but as strictly more general than both Tyson's (2008) "expansive satisficing" model as well as Fishburn (1975) and Luce's (1956) model of choice behavior generated by a semiorder. Finally, I axiomatically characterize the difference, providing novel foundations for these models.

Chapter 1

Innovation Adoption by Forward-Looking Social Learners[1]

1.1 Introduction

Suppose a new product of uncertain quality, such as a novel medical procedure or a new movie, is released into the market. In recent years, the rise of internet-based review sites, retail platforms, search engines, video-sharing websites, and social networking sites (such as Yelp, Amazon, Google, YouTube, and Facebook) has greatly increased the *potential for social learning* in the economy: If a patient suffers a serious complication or a movie-goer has a positive viewing experience, this is more likely than ever to find its way into the public domain; and there are more people than ever who have access to this common pool of consumer-generated information.

This paper builds a model studying the effect of an economy's potential for social learning on the adoption of innovations of uncertain quality. Our key contribution is a careful analysis of consumers' informational incentives and their dependence on quantitative and qualitative features of the news environment through which social learning occurs. Our analysis has two main implications. First, quantitatively, we suggest caution in evaluating the impact of *increases* in the potential for social learning: We identify news environments that are subject to a novel *saturation effect*, whereby beyond a certain level, increased opportunities for social learning can slow down adoption and learning and do not increase consumer welfare (possibly even being harmful). Second, at a qualitative level, we show that different news environments give rise to observable differences in aggregate adoption dynamics: This implies a new, *purely informational* explanation for two of the most commonly observed adoption patterns (S-shaped vs. concave curves), which we support with some suggestive evidence.

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

Forward-looking social learning is well documented empirically, notably in the

development economics literature studying the adoption of agricultural innova-
tions.[2] However, its informational ramifications have largely remained unexplored
theoretically: Existing learning-based models of innovation adoption typically as-
sume either that learning is social but consumers are myopic (e.g. Ellison and
Fudenberg, 1993; Young, 2009), or that consumers are forward-looking but informa-
tion arrives purely exogenously (e.g. Jensen, 1982). In either case, the dependence
on the informational environment is trivial, both quantitatively (a greater ease of
information transmission is always beneficial) and qualitatively (absent other forces
such as consumer heterogeneity, different news environments alone do not give rise
to interestingly different adoption dynamics[3]).

Summary of Model and Results: In our model (Section 1.2), an innovation of
fixed, but uncertain quality (better or worse than the status quo) is introduced to a
large population of forward-looking consumers. In the baseline setting, consumers
are (ex ante) identical, sharing the same prior about the quality of the innovation,
the same discount rate, and the same tastes for good and bad quality. At each
instant in continuous time, consumers receive stochastic opportunities to adopt
the innovation. A consumer who receives an opportunity must choose whether to
irreversibly adopt the innovation or to delay his decision until the next opportunity.
In equilibrium, consumers optimally trade off the opportunity cost of delays against

---

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

[3]See the discussion in Footnote 13 under Related Literature.

the benefit to learning more about the quality of the innovation.

Learning about the innovation is summarized by a public signal process, representing news that is obtained endogenously—based on the experiences of previous adopters; and possibly also from exogenous sources, such as professional critics or government watchdog agencies. To study the importance of quantitative and qualitative features of the news environment, we employ a variation of the Poisson models of strategic experimentation pioneered by Keller *et al.* (2005); Keller and Rady (2010, 2014).[4] Individual adopters' experiences generate public signals at a fixed Poisson rate which we use to quantify the potential for social learning. Qualitatively, there is a natural distinction (see also MacLeod, 2007; Board and Meyer-ter Vehn, 2013; Che and Hörner, 2014) between *bad news* markets, where signal arrivals (*breakdowns*) indicate bad quality and the absence of signals makes consumers more optimistic about the innovation; and *good news* markets, where signals (*breakthroughs*) suggest good quality and the absence of signals makes consumers more pessimistic.

For examples of innovations featuring learning via bad news (or the absence thereof), recall the extensive social media coverage of of a battery fire in a Tesla Model S electric car in October 2013, or the gradual increase of consumers' confidence in microwave ovens in the 1970s (following widespread initial concerns over possible "radiation leaks") or in risky new medical procedures such as gastric bypass surgery.[5] By contrast, learning via good news events (or their absence) is common in award-focused industries (e.g. movies or books); or for (essentially side-effect free) herbal remedies, beauty or fitness products. The news environment may also be determined by limitations or usage practices of the available social

---

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

[5]The latter two examples are discussed in more detail in Section 1.4.1.

learning systems, e.g. the fact that Facebook allows users to "Like" a product's site, but has no "Dislike" button; or that the overwhelming majority of book reviews on Amazon.com and BarnesandNoble.com appear to be positive.[6]

Section 1.3 analyzes and contrasts equilibrium adoption behavior in bad and good news markets. For tractability, we focus on *perfect* bad (respectively good) news environments, in which a *single* signal arrival *conclusively* indicates bad (respectively good) quality, so that equilibrium dynamics are non-trivial only in the absence of signals. A key insight facilitating our analysis is that consumers' equilibrium incentives across time must satisfy a quasi-single crossing property (Section 1.3.1): Absent signals, there can be at most one transition from strict preference for adoption to strict preference for waiting, or vice versa, with a possible period of indifference in between. This enables us to establish the existence of unique[7] equilibria. Equilibrium adoption dynamics admit simple closed-form descriptions, which are Markovian in current beliefs and in the mass of consumers who have not yet adopted.

Under perfect bad news (Section 1.3.2), the unique equilibrium is characterized by two times $0 \leq t_1^* \leq t_2^*$, which depend on the fundamentals: Until time t_1^*, no adoption takes place and consumers acquire information only from exogenous sources; from time t_2^* on, all consumers adopt immediately when given a chance, unless a breakdown occurs, in which case adoption comes to a permanent standstill. If $t_1^* < t_2^*$, then throughout (t_1^*, t_2^*) there is inefficiency in the form of *partial adoption*: Only *some* consumers adopt when given a chance, with others free-riding on the information generated by the adopters. The flow of new adopters on (t_1^*, t_2^*) is uniquely determined by an ODE that guarantees consumers' indifference between adopting and delaying throughout this interval. Given that consumers are forward-

---

[6]Cf. Chevalier and Mayzlin (2006), which we discuss in greater detail in Section 1.4.1.

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

8

looking, $t_1^* < t_2^*$ occurs in economies with a sufficiently large potential for social learning and not too optimistic consumers (on the other hand, if consumers are myopic or if there are no possibilities for social learning, then necessarily $t_1^* = t_2^*$).

By contrast, the perfect good news equilibrium (Section 1.3.3) features 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). The key difference with perfect bad news is that equilibrium adoption behavior is *all-or-nothing*: Regardless of the potential for social learning, there are no periods during which only some consumers adopt when given a chance. This highlights a fundamental way in which the nature of the news environment affects consumers' adoption incentives. During a period of time when, absent signals, a consumer is prepared to adopt the innovation, he will be willing to delay his decision only if he expects to acquire *decision-relevant information* in the meantime: Since originally he is prepared to adopt the innovation, such information must make him strictly prefer *not* to adopt. When learning is via bad news, breakdowns have this effect, since they reveal the innovation to be bad. By contrast, under perfect good news breakthroughs conclusively reveal the innovation to be good and hence cannot be decision-relevant to a consumer who is already willing to adopt.

Turning to implications of the equilibrium analysis, Section 1.4.1 shows that bad news and good news environments give rise to observably different adoption patterns. Under perfect bad news, adoption curves (which plot the percentage of adopters in the population against time) are S-shaped: Up to time t_1^* adoption is flat, on (t_1^*, t_2^*) adoption levels increase convexly (absent breakdowns), and from time t_2^* there is a concave increase. Convex growth throughout (t_1^*, t_2^*) is tied to consumer indifference during this region: As consumers grow increasingly optimistic absent breakdowns, their opportunity cost to delaying goes up. To maintain indifference,

9

this increase is offset by an increase in the flow of new adopters, which raises the odds that waiting will produce information allowing consumers to avoid a bad innovation. By contrast, adoption under perfect good news occurs in concave "bursts": Up to time t^* adoption levels increase concavely, then adoption flattens out, possibly followed by another region of concave growth if a breakthrough occurs. S-shaped and concave curves are arguably the two most widely documented empirical adoption patterns, with the typical marketing textbook devoting a chapter to this "stylized fact".[8] But as we discuss below under Related Literature, our model appears to be the first to point to different market learning environments as a possible source. Focusing on the aforementioned examples of good and bad news markets, we present some suggestive evidence for our predictions, pointing to an opportunity for more systematic empirical work.

Section 1.4.2 establishes the possibility of a saturation effect: If learning is via perfect bad news and the potential for social learning is great enough that $t_1^* < t_2^*$, then holding fixed other fundamentals, any additional increase in opportunities for social learning has no impact at all on (ex ante) equilibrium welfare levels. This is because any benefits from increasing the potential for social learning are balanced out by an expansion of the period (t_1^*, t_2^*) of informational free-riding. As a result, greater opportunities for social learning strictly slow down the adoption of good products and do not translate into uniformly faster learning about the quality of the innovation. In Section 1.4.3, we further build on this non-monotonicity in the speed of learning to construct an example with heterogeneous consumers, where increased opportunities for social learning are not only not beneficial, but in fact

---

[8] Cf. Hoyer *et al.* (2012), Ch. 15, p. 425ff. and Keillor (2007) p. 46–61. The former type of curve is sometimes referred to as "logistic" and the latter as "exponential" or "fast-break". In economics, S-curves are studied by Griliches (1957), Mansfield (1961, 1968), Gort and Klepper (1982), among many others; for (essentially) concave curves see some of the "group A innovations" in Davies (1979).

give rise to *Pareto-decreases* in ex ante welfare. By contrast, under perfect good news, increasing the potential for social learning is (essentially) always strictly beneficial and speeds up learning at all times.

Related Literature: We contribute to a large literature (spanning economics, marketing, and sociology)[9] that seeks to explain why the adoption of innovations is typically a drawn-out process and why different innovations follow different characteristic adoption patterns, notably the widely-documented S-shaped and concave adoption curves.[10] First, we identify a novel, *purely informational* source of these regularities: Forward-looking social learners may delay adoption to gather information about others' experiences, but delay incentives (and hence adoption patterns) are sensitive to the market learning environment.[11] Existing models appear to have overlooked this channel, appealing instead to (a combination of): (i) an assumed heterogeneity of potential adopters, with a distribution of characteristics that is imposed exogenously to fit the desired adoption pattern—as in "probit" models[12] or existing learning-based models;[13] (ii) non-informational "spillover" effects which, *independently of the quality* of the innovation, increase current adoption as a function of past adoption—e.g. by a process of contagion as in "epidemic"

---

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

[10]See footnote 8.

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

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

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

models,[14] or due to pure payoff externalities resulting from learning-by-doing (Jovanovic and Lach, 1989) or network effects (Farrell and Saloner, 1985, 1986); (iii) supply-side factors such as pricing (e.g. Bergemann and Välimäki, 1997; Cabral, 2012). To highlight the explanatory power of informational incentives alone we abstract away from (i)–(iii), but we do not wish to deny that a combination of these factors is likely often at play as well. Second, however, we investigate the effect of increased opportunities for social learning and obtain predictions (notably the saturation effect) that are outside the scope of existing models.

Our Poisson learning framework borrows from the strategic experimentation literature (Keller *et al.*, 2005; Keller and Rady, 2010, 2014),[15] but we depart in two key respects: First, since our focus is on large market applications, we assume that any individual's influence on public aggregate information is negligible. Second, we assume that adoption is irreversible rather than allowing for continuous back-and-forth switching; this is natural for innovations such as medical procedures or movies, for which "consumption" is usually a one-time event, or for technologies with large switching costs. An important theoretical implication is the absence from our model of the *encouragement effect*, which is central to the strategic experimentation literature.[16] This makes our analysis more tractable—e.g., in contrast with the

---

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

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

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

aforementioned papers, our equilibria are unique (at the aggregate level).[17] More substantively, we obtain differences between bad and good news environments that do not arise under strategic experimentation, as well as novel comparative statics with respect to the potential for social learning.[18]

In independent and contemporaneous work, Che and Hörner (2014) employ a similar variation of Keller *et al.* (2005) to model learning about a new product by a large population of consumers. However, they perform a *normative* analysis: Signals about past adopters' experiences are only visible to a benevolent mediator, who based on his information makes adoption recommendations that maximize social welfare subject to a credibility constraint. To counterbalance informational free-riding, the optimal mechanism under both good and bad news generally features regions of selective over-recommendation.[19]

Finally, informational externalities in social learning are also studied by the observational learning literature (e.g. Banerjee, 1992; Bikhchandani *et al.*, 1992; Smith and Sørensen, 2000, where the timing of players' moves is exogenous; and Chamley and Gale (1994); Rosenberg *et al.* (2007); Murto and Välimäki (2011), which like our paper feature endogenous timing). The key difference is that in this literature players hold *private* information about a payoff-relevant state variable and make

---

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

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

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

inferences by observing others' *actions*, whereas all relevant news in our model is *public* and derived from previous adopters' *experiences* (which better captures social learning via centralized internet-based review platforms). Given the assumption of private information, a particular focus of this literature is on the possibility of herding and informational cascades. By contrast, none of the cited papers derive adoption curves or study the way in which they are shaped by qualitative features of the news environment.[20]

1.2 Model

1.2.1 The Game

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

At each time t, consumers receive stochastic opportunities to adopt the innovation. Adoption opportunities are generated independently across consumers and across histories according to a Poisson process with exogenous arrival rate $\rho > 0$.[21] 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

---

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

[21]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.

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.

1.2.2 Learning

Over time, consumers observe public signals that convey information about the quality of the innovation. To highlight the importance of qualitative and quantitative features of the informational environment, we employ a variation of the Poisson learning models of Keller *et al.* (2005) and Keller and Rady (2010, 2014): Let N_t denote the flow of of consumers newly adopting the innovation at time t, which we define more precisely in Section 1.2.3. Then, conditional on quality θ, public signals arrive according to an inhomogeneous Poisson process with arrival rate $(\varepsilon_\theta + \lambda_\theta N_t)dt$, where $\lambda_\theta > 0$ and $\varepsilon_\theta \geq 0$ are exogenous parameters that depend on the quality θ of the innovation. The signal process summarizes news events that are generated from two sources:

First, the *social learning* term λN_t represents news generated endogenously, based on the experiences of other consumers: It captures the idea of a flow N_t of new adopters each generating signals at rate $\lambda \, dt$.[22] 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,

---

[22]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 adopter can generate a signal only once, namely at the time of adoption. This is appropriate for innovations such as new movies or medical procedures, for which "consumption" is a one-time event and quality is revealed upon consumption. For durable goods (e.g. cars or consumer electronics), it might be more natural to allow adopters to generate signals repeatedly over time, which can be captured by replacing 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.

we also allow for (but do not require) signals to arrive at a fixed *exogenous* rate $\varepsilon\,dt$, representing information generated independently of consumers' behavior, e.g. by professional critics or government watchdog agencies.

For tractability, we focus on learning via *perfect* Poisson processes, where a single signal provides *conclusive* evidence of the quality of the innovation. Qualitatively, we can then distinguish between learning via **perfect bad news**, where $\varepsilon_G = \lambda_G = 0$ and $\varepsilon_B = \varepsilon \geq 0$, $\lambda_B = \lambda > 0$, so that the arrival of a signal (called a *breakdown*) is conclusive evidence that the innovation is bad; and learning via **perfect good news**, where $\varepsilon_B = \lambda_B = 0$ and $\varepsilon_G = \varepsilon \geq 0$, $\lambda_G = \lambda > 0$, so that a signal arrival (called a *breakthrough*) is conclusive evidence that the innovation is good. As motivated in the Introduction and Section 1.4.1, the distinction between bad news and good news can be seen to reflect the nature of news production in different markets.

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

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

$$p_t = \frac{p_0 e^{-\int_0^t (\varepsilon_G + \lambda_G N_s)\,ds}}{p_0 e^{-\int_0^t (\varepsilon_G + \lambda_G N_s)\,ds} + (1 - p_0) e^{-\int_0^t (\varepsilon_B + \lambda_B N_s)\,ds}}.^{23} \tag{1.1}$$

---

[23]Definition 1.2.1 imposes measurability on N, so this expression is well-defined.

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

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

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

1.2.3 Equilibrium

Since our main interest is in the aggregate adoption dynamics of the population, we take as the primitive of our equilibrium concept the aggregate flow $N_{t \geq 0}$ of consumers newly adopting the innovation over time and do not explicitly model individual consumers' behavior. Given our focus on *perfect* news processes, consumers' incentives are non-trivial only in the absence of signals: Under perfect bad news, no new consumers adopt after a breakdown, while under perfect good news all remaining consumers adopt at their first opportunity after there has been a breakthrough. Therefore, we henceforth let N_t denote the flow of new adopters at t *conditional on no signals up to time t* and define equilibrium in terms of this quantity. Reflecting the assumption that aggregate adoption behavior is predictable with respect to the news process of the economy, we require that N_t be a *deterministic* function of time. We consider all such functions which are feasible in the following sense:

Definition 1.2.1. A feasible flow of adopters is a right-continuous function $N \colon [0, +\infty) \rightarrow \mathbb{R}$ such that $N_t := N(t) \in [0, \rho \bar{N}_t]$ for all $t \in [0, +\infty)$, where $\bar{N}_t := \bar{N}_0 - \int_0^t N_s ds$.

Here $\bar{N}_t$ denotes the mass of consumers remaining in the game at time t. We

require that $N_t \leq \rho \bar{N}_t$ so that N_t is consistent with the remaining $\bar{N}_t$ consumers independently receiving adoption opportunities at Poisson rate ρ. Any feasible adoption flow $N_{t \geq 0}$ defines an associated no-news posterior p_t^N as given by Equation (1.1).

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

Let 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.[24]

Let $W_t^N(\sigma)$ denote the expected payoff to waiting at t and following σ in the future, given the aggregate adoption flow $N_{s \geq 0}$. Specifically, 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.

---

[24]Formally, let $(X_s)_{s \geq t}$ denote the stochastic process representing the number of arrivals generated on $[t, s]$ by a Poisson process with arrival rate ρ, and let $(X_{s-})_{s > t}$ denote the number of arrivals on $[t, s)$. Then $\tau^\sigma := \inf\{s \geq t : \sigma_s \times (X_s - X_{s-}) > 0\}$, with the usual convention that $\inf \varnothing := +\infty$. 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.

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

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

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

We are now ready to formally define equilibrium:

Definition 1.2.3. An equilibrium is a feasible adoption flow $N_{t \geq 0}$ such that

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

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

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

19

consumers optimally trading off the expected payoff to immediate adoption against the value to waiting.[25]

1.3 Equilibrium Analysis

1.3.1 Quasi-Single Crossing Property for Equilibrium Incentives

We now proceed to equilibrium analysis. As a preliminary step, we first establish a useful property of equilibrium incentives under both perfect bad news and perfect good news. Suppose that $N_{t \geq 0}$ is an *arbitrary* feasible flow of adopters, with associated no-news posterior $p_{t \geq 0}^N$ and value to waiting $W_{t \geq 0}^N$ as defined in Definition 1.2.2. In general, the dynamics of the trade-off between immediate adoption at time t (yielding expected payoff $2p_t^N - 1$) and delaying and behaving optimally in the future (yielding expected payoff W_t^N) can be quite difficult to characterize, with $(2p_t^N - 1) - W_t^N$ changing sign many times. However, when $N_{t \geq 0}$ is an *equilibrium* flow, then for any t,

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

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

---

[25]Note that Definition 1.2.3 is essentially Nash equilibrium, i.e. we do not impose subgame perfection. The motivation is that in a continuum population any individual consumer's behavior has a negligible impact on aggregate adoption levels, so that any off-path history in which the flow of adopters differs from the equilibrium flow is more than a unilateral deviation from the equilibrium path. Thus, off-path histories do not affect individual consumers' incentives on the equilibrium path and are unimportant for equilibrium analysis.

Theorem 1.3.1. *Suppose that learning is either via perfect bad news ($\lambda_B > 0 = \lambda_G$) or via perfect good news ($\lambda_G > 0 = \lambda_B$). Let $N_{t \geq 0}$ be an equilibrium, with corresponding no-news posteriors $p^N_{t \geq 0}$ 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:*

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

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

The proof is in Appendix A.1. We briefly illustrate the intuition for the first bullet point when learning is via perfect bad news. Suppose that immediate adoption is strictly better than waiting today (and hence also in the near future provided there are no breakdowns).[26] Then in the near future all consumers adopt upon their first opportunity, so the no-news posterior strictly increases while the number of remaining consumers strictly decreases. Because information is generated endogenously, this means that the flow of information must be decreasing over time. As a result, immediate adoption becomes even more attractive relative to waiting, and consequently immediate adoption continues to be strictly preferable at all times in the future.

With any equilibrium $N_{t \geq 0}$, we associate two cutoff times $0 \leq t_1^* \leq t_2^* \leq +\infty$:[27] If learning is via perfect bad news, set

$$t_1^* := \inf\{t \geq 0 : N_t > 0\} \text{ and } t_2^* := \sup\{t \geq 0 : N_t < \rho \bar{N}_t\};^{28} \tag{1.2}$$

---

[26]This follows from the continuity of the equilibrium value to waiting, which is established in Lemma A.1.1 in the Appendix.

[27]With the convention that $\inf \varnothing = +\infty$ and $\sup \varnothing = 0$.
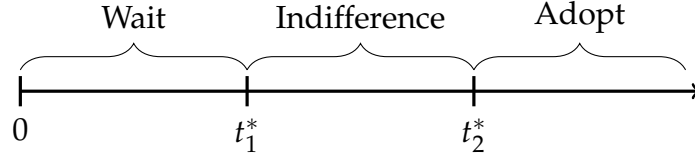
Figure 1.1: *Perfect Bad News*
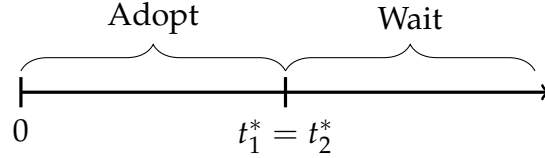
Figure 1.2: *Perfect Good News*

if learning is via perfect good news, set

$$t_1^* := \inf\{t \geq 0 : N_t < \rho \bar{N}_t\} \text{ and } t_2^* := \sup\{t \geq 0 : N_t > 0\}. \qquad (1.3)$$

Thus, if $N_{t \geq 0}$ is a perfect bad news equilibrium it features no adoption ($N_t = 0$) for all $t < t_1^*$ and immediate adoption ($N_t = \rho \bar{N}_t$) for all $t > t_2^*$ absent breakdowns; while under perfect good news $N_{t \geq 0}$ features immediate adoption prior to t_1^* and no adoption after t_2^* absent breakthroughs. Moreover, under both perfect bad and good news, Theorem 1.3.1 implies that at all times $t \in (t_1^*, t_2^*)$, consumers are indifferent ($2p_t^N - 1 = W_t^N$) between adopting and delaying.[29] This is illustrated in Figures 1 and 2. In Sections 1.3.2 and 1.3.3 we will build on this observation to establish the existence of unique equilibria under both perfect bad and good news. The cutoff times, as well as the flow of adopters between t_1^* and t_2^*, are fully pinned down by

---

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

[29]Suppose learning is via perfect good news. Consider $t \in (t_1^*, t_2^*)$. By the definition of t_1^* and t_2^*, there exist $k \in (t_1^*, t)$ and $l \in (t, t_2^*)$ such that $N_k < \rho \bar{N}_k$ and $N_l > 0$. Since N is an equilibrium, this implies $2p_k - 1 \leq W_k$ and $2p_l - 1 \geq W_l$, whence by Theorem 1.3.1 $2p_t - 1 = W_t$. The argument for perfect bad news is analogous.

22

the parameters. Looking ahead to Section 1.3.3, we will see that the perfect good news equilibrium satisfies $t_1^* = t_2^* = t^*$; thus, adoption behavior is **all-or-nothing**, with all consumers adopting upon first opportunity up to time t^* and adoption ceasing from then on absent breakthroughs. By contrast, for suitable parameters the perfect bad news equilibrium in Section 1.3.2 features a non-empty region (t_1^*, t_2^*). Maintaining indifference throughout (t_1^*, t_2^*) requires a form of informational free-riding, which we term **partial adoption**, whereby only *some* consumers adopt when given the chance (i.e. $N_t \in (0, \rho \bar{N}_t)$ at each $t \in (t_1^*, t_2^*)$). We will see that partial adoption has important implications for the shape of the adoption curve and for the impact of increased opportunities for social learning on welfare, learning, and adoption dynamics.

1.3.2 Equilibrium under Perfect Bad News

Assume that learning is via perfect bad news. The following theorem builds on the analysis of the previous section to establish the existence of an equilibrium $N_{t \geq 0}$, which is uniquely pinned down by the parameters. At all t, N_t is Markovian in the associated no-news posterior p_t and the time-t potential for social learning $\Lambda_t := \lambda \bar{N}_t$:[30]

Theorem 1.3.2 (Equilibrium under PBN). *Fix $r, \rho, \lambda, \bar{N}_0 > 0$, $\varepsilon \geq 0$, and $p_0 \in (0, 1)$. There exists a unique equilibrium. Furthermore, in the unique equilibrium, N_t is Markovian in (p_t, Λ_t) for all t: There exists a non-decreasing function $\Lambda^* : [0, 1] \to \mathbb{R} \cup \{\infty\}$ and*

---

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

some $p^* \in [\frac{1}{2}, 1)$ such that

$$
N_t = \begin{cases} 0 & \text{if } p_t < p^* \text{ and } \Lambda_t > \Lambda^*(p_t) \\[2mm] \frac{r(2p_t - 1)}{\lambda(1 - p_t)} - \frac{\varepsilon}{\lambda} \in (0, \rho \bar{N}_t) & \text{if } p_t \geq p^* \text{ and } \Lambda_t > \Lambda^*(p_t) \\[2mm] \rho \bar{N}_t & \text{if } \Lambda_t \leq \Lambda^*(p_t). \end{cases} \tag{1.4}
$$

The proof of Theorem 1.3.2 is in Appendix A.2.2. Here we sketch the basic argument. Fix parameters r, ρ, $\bar{N}_0 > 0$, ε, $\lambda \geq 0$, and $p_0 \in (0, 1)$, and suppose that $N_{t \geq 0}$ is an equilibrium. By the previous section, Equation (1.2) defines cutoff times $0 \leq t_1^* \leq t_2^* \leq +\infty$ such that $N_t = 0$ if $t < t_1^*$, $N_t = \rho \bar{N}_t$ if $t > t_2^*$, and at all $t \in [t_1^*, t_2^*)$, consumers are indifferent between adopting immediately and waiting for more information.

Partial adoption during (t_1^*, t_2^*)**:** Lemma A.2.1 in Appendix A.2.2 shows that the flow of adopters at all times $t \in (t_1^*, t_2^*)$ must satisfy $N_t = \frac{r(2p_t - 1)}{\lambda(1 - p_t)} - \frac{\varepsilon}{\lambda} \in (0, \rho \bar{N}_t)$—thus, adoption throughout (t_1^*, t_2^*) is *partial*, with only some consumers adopting when given a chance and others free-riding on the information generated by the adopters. Heuristically, maintaining consumer indifference requires that the cost and benefit of delaying be equal:

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

$$\tag{1.5}$$

Delaying one's decision by an instant is beneficial if a breakdown occurs at that instant, allowing a consumer to permanently avoid the bad product. The gain in this case is $(0 - (-1)) = 1$, and this possibility arises with an instantaneous probability of $(\varepsilon + \lambda N_t)(1 - p_t)dt$. On the other hand, if no breakdown occurs, which happens

with instantaneous probability $1 - (\varepsilon + \lambda N_t)\,(1 - p_t)dt$, then consumers incur an opportunity cost of $(2p_{t+dt} - 1)rdt$, reflecting the time cost of delayed adoption.[31] Ignoring terms of order dt^2 and rearranging yields $N_t = \frac{r(2p_t - 1)}{\lambda(1 - p_t)} - \frac{\varepsilon}{\lambda}$.[32]

Determining the cutoff times: Next, we derive an alternative description of t_1^* and t_2^* in terms of the evolution of the no-news posterior p_t and the potential for social learning Λ_t. To state this description, we define the following notation. For any $p \in (0,1)$ and $\Lambda \geq 0$, let

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

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

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

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

Define $\overline{p} := \lim_{\rho \to \infty} \underline{p} = \frac{\varepsilon + r}{\varepsilon + 2r}$ to be the lowest belief at which a hypothetical consumer to whom adoption opportunities arrive *continuously* would be willing

---

[31]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.

[32]A bit more precisely, ignoring terms of order dt^2, the right hand side of Equation 1.5 is given by $(1 - (\varepsilon + \lambda N_t)\,(1 - p_t)dt)(2(p_t + \dot{p}_t dt) - 1)rdt = r(2p_t - 1)dt$. Further rearrangement yields the desired expression.

to adopt immediately if all information in the future arrives exclusively through the exogenous new source. Define $p^\sharp := \lim_{\varepsilon \to \infty} \underline{p} = \frac{\rho+r}{\rho+2r}$ to be the lowest belief at which a consumer to whom adoption opportunities arrive at rate ρ would be willing to adopt immediately even if all uncertainty were to be *completely resolved* by the next adoption opportunity.[33]

Finally, define the function $\Lambda^* : [0,1] \to \mathbb{R}_+ \cup \{+\infty\}$ as follows. Let $\Lambda^*(p) = 0$ for all $p \leq \underline{p}$, $\Lambda^*(p) = +\infty$ for all $p \geq p^\sharp$, and for all $p \in (\underline{p}, p^\sharp)$, let $\Lambda^*(p) \in \mathbb{R}_+$ be the unique value such that $2p - 1 = G(p, \Lambda^*(p))$.[34] 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 $\Lambda^*(p)/\lambda$ consumers also adopt at their first opportunity in the future.

Then, letting $p^* := \min\{\overline{p}, p^\sharp\}$, Lemma A.2.3 in Appendix A.2.2 shows that $t_2^* = \inf\{t \geq 0 : \Lambda_t < \Lambda^*(p_t)\}$ and $t_1^* = \min\{t_2^*, \sup\{t \geq 0 : p_t < p^*\}\}$.[35]

Equilibrium dynamics given initial parameters: From the previous two steps, it is clear that any equilibrium must take the Markovian form in Equation (1.4), with Λ^* and p^* as defined above. It remains to show how Equation (1.4) uniquely pins down the evolution of N_t as a function of the initial parameters; and to verify that $N_{t \geq 0}$ thus obtained does indeed constitute an equilibrium (in particular, is feasible). Here we sketch the former argument, relegating the latter to Appendix A.2.2. Note first the following two special cases: If $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$, then Equation (1.4) implies that $N_t = 0$ for all t. Second, if $\varepsilon \geq \rho$ (so that $p^* := \min\{\overline{p}, p^\sharp\} = p^\sharp$), then because

---

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

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

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

$\Lambda^*(p) = +\infty$ for all $p \geq p^\sharp$, $N_t = 0$ as long as $\Lambda_t > \Lambda^*(p_t)$ and $N_t = \rho \bar{N}_t$ as soon as $\Lambda_t \leq \Lambda^*(p_t)$. Throughout the rest of the paper, we will be particularly interested in equilibria that feature a non-empty partial adoption region (t_1^*, t_2^*). Since the two cases above preclude this regardless of other parameters, we henceforth rule them out (Appendix A.2.2 discusses the second case in more detail):

Condition 1.3.3. The rate at which exogenous information arrives is smaller than the rate at which consumers obtain adoption opportunities: $\varepsilon < \rho$.

Condition 1.3.4. Either $\varepsilon > 0$ or $p_0 \in (\frac{1}{2}, 1)$.

Given these conditions, Figure 1.3 illustrates how the unique equilibrium is obtained as a function of the parameters. Regions (2) and (3) represent values of (p_t, Λ_t) corresponding to the first line of Equation (1.4), so that no adoption takes place in these regions. Region (4) corresponds to partial adoption as given by the second line of Equation (1.4). Finally, region (1) corresponds to the third line of Equation (1.4) and thus to immediate adoption.

If (p_0, Λ_0) is in region (2), then initially no adoption occurs and the no-news posterior drifts upward according to the law of motion $\dot{p}_t = p_t(1 - p_t)\varepsilon$, while 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); subsequently consumers adopt immediately upon an opportunity so that $N_t = \rho e^{-\rho(t - t_2^*)} \bar{N}_{t_2^*}$ uniquely pins down the evolution of (p_t, 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 $\Lambda_{t_1^*} = \Lambda_0 > \Lambda(p_{t_1^*}) = \Lambda(\bar{p})$, 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
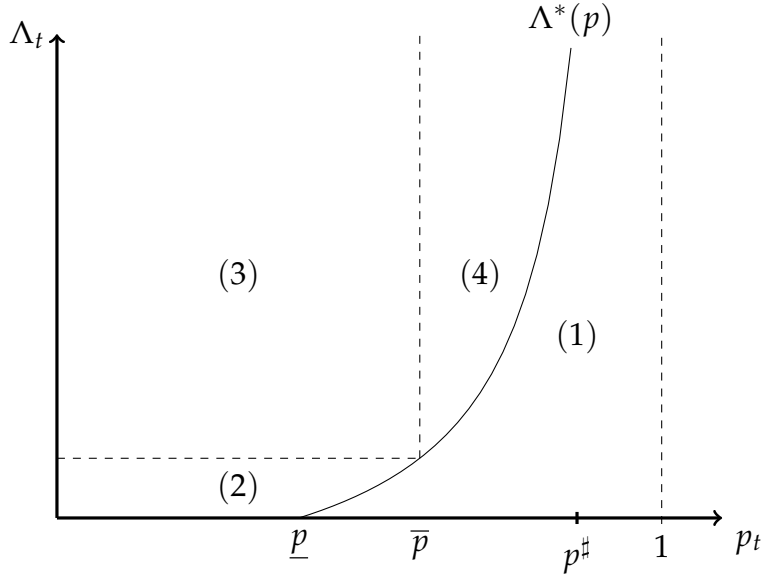
Figure 1.3: *Partition of (p_t, Λ_t) when $\varepsilon < \rho$*

Equation (1.4).[36] 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 which (p_t, Λ_t), evolving according to the second line of Equation (1.4), 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.

Conditions for partial adoption: As seen above, whether or not the equilibrium features a period of partial adoption depends on the fundamentals. More specifically, Figure 1.3 shows that if consumers are forward-looking and not too optimistic

---

[36]Specifically, combining the second line of Equation (1.4) with Equation (1.1) 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 $t_2^* = \inf\{t : \Lambda_t < \Lambda^*(p_t)\} < +\infty$.

28

$(p_0 < p^\sharp)$, then $t_1^* < t_2^*$ holds whenever the potential for social learning Λ_0 is sufficiently large. The following lemma states this precisely:

Lemma 1.3.5. *Fix ρ, ε and p_0 satisfying Conditions 1.3.3 and 1.3.4. Assume $p_0 < p^\sharp$. Then for all $r > 0$, there exists $\bar{\Lambda}_0(r) > 0$ such that $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$[37] if and only if $\Lambda_0 > \bar{\Lambda}_0(r)$.*

Proof. Set $\bar{\Lambda}_0(r) := \max\{\Lambda^*(p_0), \Lambda^*(\bar{p})\}$ and see Appendix A.2.4 . ∎

On the other hand, if as in existing learning-based models of innovation adoption, learning is *purely exogenous* ($\lambda = 0$ and $\varepsilon > 0$) or consumers are *myopic* ("$r = +\infty$"), then there is *never* any partial adoption, regardless of other parameters. In the former case, $0 = \Lambda_t < \Lambda^*(p)$ for all $p > \underline{p}$, so by Theorem 1.3.2 no consumers adopt until the no-news posterior hits $\underline{p}$ (at $t_1^* = t_2^*$) and from then on all consumers adopt immediately when given a chance. The latter case corresponds to $\underline{p} = \bar{p} = \frac{1}{2}$ and $\Lambda^*(p) = +\infty$ for all $p > \frac{1}{2}$, so $t_1^* = t_2^* = \inf\{t : p_t > \frac{1}{2}\}$. Thus, the possibility of partial adoption in equilibrium hinges crucially both on consumers being forward-looking and on there being opportunities for social learning.

1.3.3 Equilibrium under Perfect Good News

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

---

[37]Note that by the Markovian description of equilibrium dynamics, Λ_0 is a sufficient statistic for equilibrium; i.e., holding all other fundamentals fixed, Λ_0 fully pins down the corresponding no-news equilibrium adoption flow, beliefs and cutoff times $t_1^*(\Lambda_0)$ and $t_2^*(\Lambda_0)$.

Theorem 1.3.6 (Equilibrium under PGN). *Let $r, \rho, \lambda, \bar{N}_0 > 0$, $p_0 \in (0,1)$, and $\varepsilon \geq 0$. There exists a unique equilibrium. Moreover, in the unique equilibrium, N_t is Markovian in (p_t, Λ_t) (or equivalently $(p_t, \bar{N}_t)$) for all t and satisfies:*

$$N_t = \begin{cases} \rho \bar{N}_t & \text{if } p_t > p^* \\ 0 & \text{if } p_t \leq p^*, \end{cases} \tag{1.6}$$

where

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

To prove Theorem 1.3.6 we again invoke the quasi-single crossing property for equilibrium incentives established in Theorem 1.3.1. As we saw in Section 1.3.1, this implies that in any equilibrium, there are times $0 \leq t_1^* \leq t_2^* \leq +\infty$ defined by Equation (1.3) such that absent breakthroughs, $N_t = \rho \bar{N}_t$ if $t < t_1^*$, $N_t = 0$ if $t > t_2^*$, and throughout (t_1^*, t_2^*) consumers are indifferent between adopting immediately and waiting for more information.

The key observation (Lemma A.2.6 in Appendix A.2.3) is that we must in fact have $t_1^* = t_2^* =: t^*$. To see the intuition, suppose $t_1^* < t_2^*$. Then consumers would be indifferent between adopting and delaying at each time $t \in (t_1^*, t_2^*)$. Moreover, there is $t \in (t_1^*, t_2^*)$ and $\Delta \in (0, t_2^* - t)$ such that $N_\tau > 0$ throughout $[t, t + \Delta)$.[38] 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:

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

The first term represents the gain to immediate adoption if no breakthrough occurs between t and $t + dt$, which happens with instantaneous probability $(1 -$

---

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

$p_t(\lambda N_t + \varepsilon)dt$). Just as with perfect bad news, the gain to adopting immediately in this case is $r(2p_{t+dt} - 1)dt$, representing time discounting at rate r and the fact that at $t + dt$ the consumer remains indifferent between adopting given an opportunity and delaying. Ignoring terms of order dt^2 yields $r(2p_t - 1)dt$. The second term represents the gain to immediate adoption if there is a breakthrough between t and $t + dt$, which happens with instantaneous probability $p_t(\lambda N_t + \varepsilon)dt > 0$. Now the situation is very different from the perfect bad news setting: A breakthrough conclusively signals *good* quality, so a consumer who delays his decision by an instant will adopt immediately at his next opportunity. This results in a discounted payoff of $\frac{\rho}{r+\rho}$, reflecting the stochasticity of adoption opportunities. On the other hand, by adopting at t, the consumer receives a payoff of $1 > \frac{\rho}{r+\rho}$ immediately. Thus, regardless of whether or not there is a breakthrough between t and $t + dt$, there is a strictly positive gain to adopting immediately at t, contradicting indifference at t.

The above argument illustrates a fundamental difference between the bad news and good news setting. In order to maintain indifference over a period of time between immediate adoption and waiting, it must be possible to acquire *decision-relevant information* by waiting: Consumers who are prepared to adopt at t will be willing to delay their decision by an instant only if there is a possibility that at the next instant they will no longer be willing to adopt. In the bad news setting, this is indeed possible, because a breakdown might occur. On the other hand, if learning is via good news, this cannot happen: A breakthrough between t and $t + dt$ reveals the innovation to be good, so consumers strictly prefer to adopt from $t + dt$ on; if there is no breakthrough, then consumers remain indifferent at $t + dt$, so in either case the information obtained is not decision-relevant.[39]

---

[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.

Given that $t_1^* = t_2^* = t^*$, Theorem 1.3.6 follows from the observation that $p_t \leq p^*$ if and only if $t \geq t^*$ (Lemma A.2.7 in Appendix A.2.3). It is worth noting that if $\varepsilon = 0$, then $p^* = \frac{1}{2}$, so regardless of the discount rate r, consumers behave *entirely myopically*. If $\varepsilon > 0$, then consumers' forward-looking nature is reflected by the fact that the cutoff posterior p^* below which consumers are unwilling to adopt is $\frac{(r+\rho)(r+\varepsilon)}{2(r+\rho)(r+\varepsilon)-\rho\varepsilon} > \frac{1}{2}$. In both cases, the cutoff posterior does not depend on λ or $\bar{N}_0$: Social learning only affects the *time t^** at which adoption ceases conditional on no breakthroughs.

1.4 Implications

1.4.1 Adoption Curves: S-Shaped vs. Concave

The differing informational incentives of bad and good news environments have observable implications. Consider the *adoption curve* of the innovation, which plots the percentage of adopters in the population against time. Conditional on no news up to time t, this is given by $A_t := \int_0^t N_s / \bar{N}_0 \, ds$.

Theorems 1.3.2 and 1.3.6 translate directly into different predictions for the shape of the adoption curve, as summarized by the following corollary: Under perfect bad news, A_t exhibits an *S-shaped* (i.e. convex-concave) growth pattern, where the region of convex growth coincides precisely with the partial adoption region (t_1^*, t_2^*). By contrast, under perfect good news, adoption proceeds in *concave "bursts"*:

Corollary 1.4.1. *Perfect Bad News: In the unique equilibrium of Theorem 1.3.2, A_t has the following shape: For $0 \leq t < t_1^*$, $A_t = 0$; for $t_1^* \leq t < t_2^*$, A_t is strictly increasing and convex in t; for $t \geq t_2^*$, A_t is strictly increasing and concave in t. If the first breakdown occurs at time t, then adoption comes to a standstill from then on.*

Perfect Good News: *In the unique equilibrium of Theorem 1.3.6, $A_t = 1 - e^{-\rho t}$ for all $t < t^*$, which is strictly increasing and concave. If there is a breakthrough prior to t^*, then the proportion of adopters is given by $1 - e^{-\rho t}$ 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 $1 - e^{-\rho(t^* + t - s)}$.*
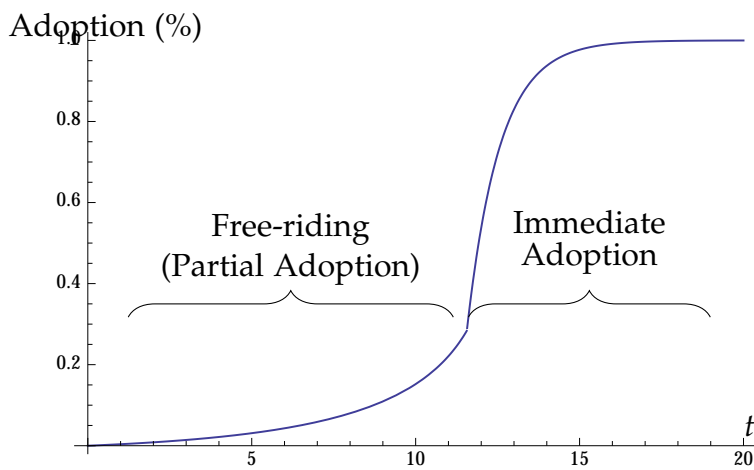
Figure 1.4: *Adoption curve under PBN conditional on no breakdowns ($\varepsilon = 0$)*

Figures 1.4 and 1.5[41] illustrate the differing adoption patterns. As we discussed in the Introduction, both patterns have been widely documented empirically, but our model differs from existing explanations in identifying a purely informational source of this regularity: We predict S-shaped adoption curves in bad news markets with a sufficiently large potential for social learning and sufficiently forward-looking and not too optimistic consumers (so that $t_1^* < t_2^*$ by Lemma 1.3.5), and concave adoption patterns in good news markets (or in bad news markets with little potential for social learning and with very optimistic and impatient consumers).

---

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

[41]Associated parameter values: $\varepsilon = 1/2$, $r = 1$, $\rho = 1$, $\lambda = 0.5$, and $p_0 = 0.7$.
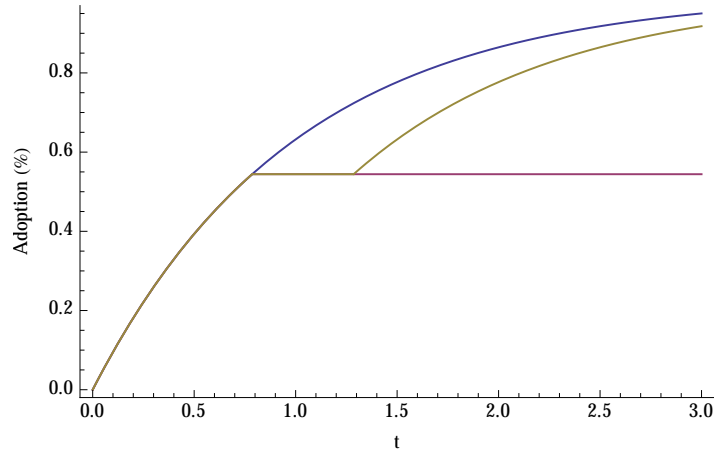
Figure 1.5: *Adoption curves under PGN (blue = breakthrough before t^*; yellow = breakthrough after t^*; pink = bad quality)*

The convex[42] growth region of A_t under perfect bad news coincides precisely with the partial adoption region (t_1^*, t_2^*) and is tied to consumer indifference in this region: Conditional on no breakdowns during this period, consumers grow increasingly optimistic about the quality of the innovation, which increases their opportunity cost of delaying adoption. To maintain indifference, the benefit to delaying adoption must then also increase over time: This is achieved by increasing the arrival rate of future breakdowns, which improves the odds that waiting will allow consumers to avoid the bad product. But since the arrival rate of information is increasing in the flow N_t of new adopters, this means that N_t must be strictly increasing throughout (t_1^*, t_2^*). Since N_t represents the rate of change of A_t, this is equivalent to A_t being convex.[43] As we discussed following Lemma 1.3.5, partial adoption depends on the *joint* assumption of forward-looking consumers and social

---

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

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

learning. This is why we are able to generate S-shaped adoption curves even when consumers are ex ante identical, whereas existing learning-based models with myopic consumers (Young, 2009) or purely exogenous learning (Jensen, 1982) must appeal to specific distributions of consumer heterogeneity.[44]

Our predictions suggest the need for empirical work that would systematically investigate the qualitative and quantitative features of consumer learning about different innovations and compare the associated adoption patterns. Here we provide some suggestive evidence:

Microwave Oven Ownership

Source: U.S. Department of Commerce
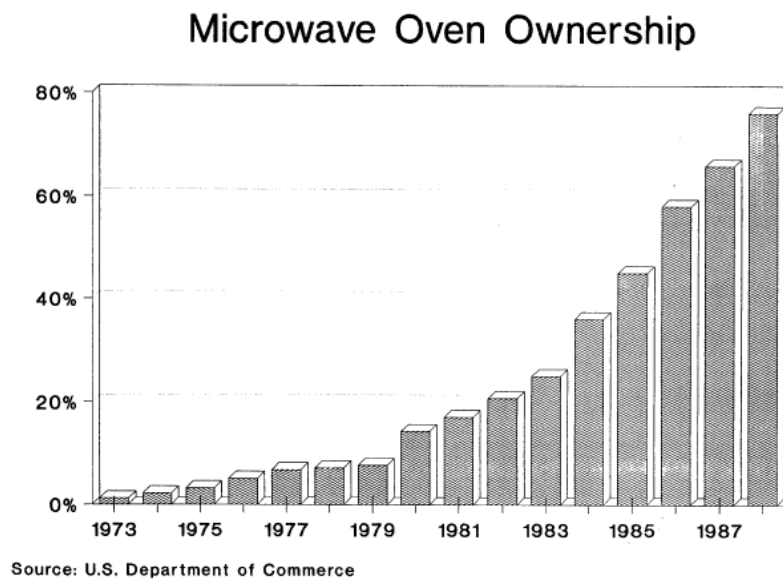
Figure 1.6: *Adoption of microwaves by US households (Source: Guenthner et al. (1991)).*

Learning via bad news events (or their absence) seems especially plausible in the case of new technologies or medical procedures whose introduction was accom-

---

[44]See footnote 13. One exception is Kapur (1995), where a finite number of identical firms engage in a sequence of waiting contests to adopt a new technology and more information is revealed when more firms adopt during a given waiting contest. This can be viewed as a form of forward-looking social learning. He shows that the mean duration of waiting contests shrinks over time, suggesting a crude approximation of convex diffusion.

panied by initial safety concerns: For example, following Raytheon's introduction to the US market of the first countertop household microwave oven in 1967, the 1970s were characterized by widespread concerns about possible "radiation leaks", stirred up for instance by a Consumers' Union (1973) report which concluded that "we are not convinced that they are completely safe to use"[45] and by Paul Brodeur's 1977 bestseller *The Zapping of America*.[46] Thus, it seems plausible that some consumers would have delayed their purchase in the hope of learning whether previous adopters experienced any adverse effects, as suggested for instance by Wiersema and Buzzell (1979).[47] Consistent with our predictions for bad news markets, the microwave is a textbook example of an innovation with an S-shaped adoption pattern: Figure 1.6 shows the convex growth in US adoption levels through the late 1980s, with later growth slowing to reach ownership levels of around 97% in 2011.[48] A second example is bariatric surgery, a collection of surgical weight loss procedures (including gastric bypass and gastric band surgery) which began gaining momentum in the mid-1990s. As with any major surgery, complications are possible, with typical health advice websites containing statements such as "a small degree of risk, including death, is inherent to all types of surgery" and "Because bariatric surgery is a relatively new surgical specialty, there are not yet enough medical

---

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

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

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

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

data to predict with certainty which patients will have better outcomes."[49] Again, consistent with some patients deciding to delay the procedure to learn whether previous adopters suffered serious complications, the available data suggests an S-shaped growth pattern.[50]
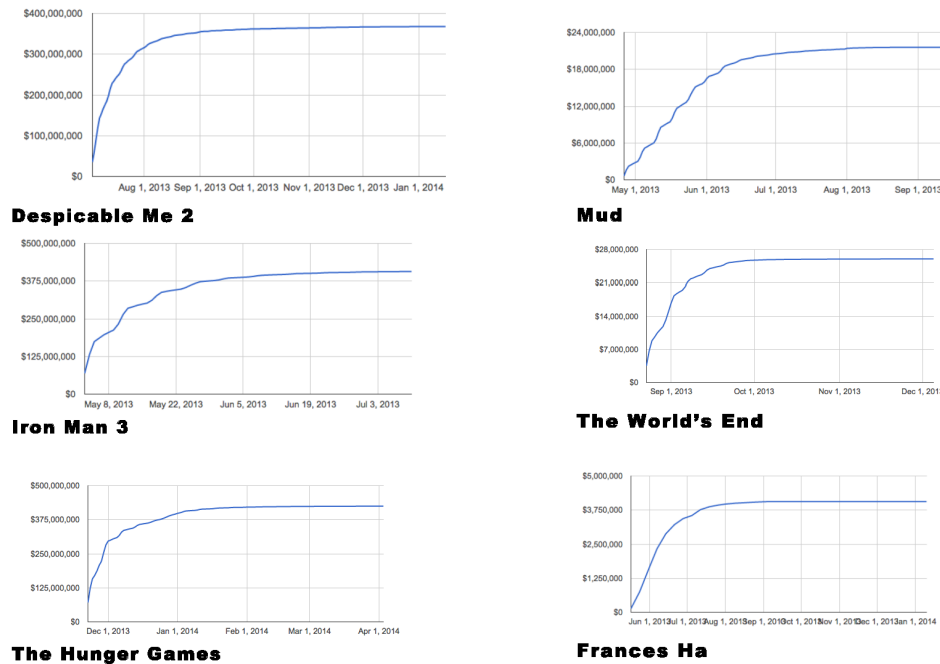
Figure 1.7: *2013 cumulative box office sales for various blockbuster (left) and independent (right) movies (Source:* `http://www.the-numbers.com`*).*

Concave adoption patterns have been studied in the marketing literature under the name "fast-break product life cycles", with movies (Figure 1.7), books, music

---

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

[50]According to Buchwald and Oien (2009) p. 1609 and Buchwald and Oien (2013) p. 428, the annual number of procedures performed worldwide (i.e. the number of *new* adoptions) increased from 40,000 in 1998 to 146,301 in 2003 and to 344,221 in 2008, and then plateaued at 340,768 in 2011. We note that an explanation in terms of reduced costs does not seem possible: For example, in the US the number of annual procedures increased from 13,386 to 121,055 between 1998 and 2004, while the average cost per procedure saw only a limited decrease, from $10,970 to $10,395; cf. Zhao and Encinosa (2007) Table 1, p. 6.

and similar leisure-enhancing products as canonical examples.[51] Consistent with our predictions, these domains appear to better fit the good news than the bad news model. For example, in a 2003–2004 study of consumer reviews of a representative sample of 6405 books on Amazon.com and BarnesandNoble.com, Chevalier and Mayzlin (2006) find that reviews "are overwhelmingly positive overall at both sites,"[52] suggesting that social learning in this domain proceeds via good news signals (or their absence) rather than via bad news signals: On a scale from one (worst) to five (best) stars, the modal review in the study is 5 stars, the mean star rating exceeds 4, and the fraction of 1-star ratings is in the range of 0.03–0.08.[53] As far as exogenously generated news is concerned, it would again appear that positive events, such as Academy Award, Grammy Award, or Booker Prize wins, receive far greater coverage than the occasional damning review by a critic (for this reason, Board and Meyer-ter Vehn (2013)[54] also cite the movie industry as an example of a good news market). Based on our model, we would also conjecture concave adoption patterns for (essentially side-effect free) herbal remedies and other alternative medical treatments, and for many beauty and fitness products, for which anecdotal evidence suggests that consumer learning is primarily about "whether they actually work" (i.e. good news events or their absence).

---

[51]Cf. Keillor (2007), pp. 51–61.

[52]Chevalier and Mayzlin (2006), p. 347.

[53]Ibid., Table 1, p. 347.

[54]Board and Meyer-ter Vehn (2013) footnote 2, p. 2382.

1.4.2 The Effect of Increased Opportunities for Social Learning

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

Under perfect bad news, an economy's ability to harness its potential for social learning is subject to a surprising *saturation effect*: Up to a certain cutoff level, increasing Λ_0 strictly increases ex-ante welfare, speeds up learning, and decreases expected adoption levels of bad products while leaving adoption levels of good products unaffected; but beyond this cutoff level, further increases in Λ_0 are ex-ante welfare-neutral, cause learning to slow down over certain periods, and strictly slow down the adoption of good products. By contrast, there is no such saturation effect under perfect good news.

Throughout this section we fix r, ρ, ε, and p_0 and study the effect of increasing Λ_0 on ex-ante equilibrium welfare $W_0(\Lambda_0)$; equilibrium cutoff times $t_1^*(\Lambda_0)$, $t_2^*(\Lambda_0)$; no-news posteriors $p_t^{\Lambda_0}$; and expected adoption levels $A_t(\Lambda_0, G)$ and $A_t(\Lambda_0, B)$ conditional on good and bad quality, respectively.[55]

Perfect Bad News: The following proposition, which we prove in Appendix A.2.4, summarizes the saturation effect.

Proposition 1.4.2. *Consider learning via perfect bad news. Fix $r, \rho > 0$, $\varepsilon \geq 0$, and p_0 satisfying Conditions 1.3.3 and 1.3.4 and such that $p_0 \in (\bar{p}, p^\sharp)$.[56] Consider $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. Then:*

---

[55]Note that because of the Markovian description of the equilibrium in Theorem 1.3.2 and Theorem 1.3.6, Λ_0 is a sufficient statistic for these quantities when all other parameters are fixed.

[56]We assume $p_0 \in (\bar{p}, p^\sharp)$, so that $t_1^* = 0$, to focus on the inefficiency due to partial adoption without having to take into account the effect of Λ_0 on t_1^*. As we show in Appendix A.2.4, the welfare-neutrality result remains valid if $p_0 \in (0, \bar{p}]$, but now the cutoff-level above which it holds is $\Lambda^*(\bar{p})$ rather than $\Lambda^*(p_0)$.

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

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

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

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

- $p_t^{\Lambda_0} < p_t^{\hat{\Lambda}_0}$ *for all* $t > \bar{t}$.

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

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

The saturation effect obtains once Λ_0 exceeds $\Lambda^*(p_0)$. This is precisely the level above which the equilibrium features an initial partial adoption region ($0 = t_1^* < t_2^*(\Lambda_0)$), so that consumers at time 0 are indifferent between delaying and adopting. This immediately implies welfare-neutrality, because $W_0(\Lambda_0) = 2p_0 - 1$ irrespective of the value of $\Lambda_0 \geq \Lambda^*(p_0)$.[57] This result is in stark contrast to the cooperative benchmark in which consumers coordinate on socially optimal adoption levels: Here increased opportunities for social learning are always strictly beneficial and for any $p_0 > \frac{1}{2}$ the first-best (complete information) payoff of $\frac{\rho}{r+\rho} p_0$ can be approximated in the limit as $\Lambda_0 \to \infty$.[58]

(ii) and (iii) further illuminate the forces behind welfare-neutrality: Because an increase in Λ_0 affects learning dynamics in a non-monotonic manner, the impact

---

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

[58]The cooperative benchmark is derived in Section 3.2 of an earlier version of this paper, Frick and Ishii (2014): It takes an all-or-nothing form, with no adoption below a cutoff belief p^s and immediate adoption above p^s. Relative to this, equilibrium displays two types of inefficiency: First, because $p^s < \underline{p}$, adoption generally begins *too late*. Second, whenever $t_1^* < t_2^*$, then once consumers begin to adopt, the initial rate of adoption is *too low*. Cf. Frick and Ishii (2014), section 5.3.

on a consumer's expected payoff varies with the time t at which he obtains his first adoption opportunity: If $t \leq t_2^*(\Lambda_0)$, his expected payoff is the same under Λ_0 and $\hat{\Lambda}_0$; if $t \in (t_2^*(\Lambda_0), \bar{t})$, he is strictly worse off under $\hat{\Lambda}_0$, because in case the innovation is bad he is *less* likely to have found out by then than under Λ_0; finally, if $t > \bar{t}$, he is strictly better off under $\hat{\Lambda}_0$. Depending on $\hat{\Lambda}_0$, $\bar{t}$ adjusts endogenously to balance out the benefits, which arrive at times after $\bar{t}$, with the costs incurred at times $(t_2^*(\Lambda_0), \bar{t})$.

Similarly, by (iii), an increase in Λ_0 strictly decreases $A_t(\Lambda_0, G)$ (which is harmful), but also decreases $A_t(\Lambda_0, B)$ (which is beneficial), and welfare-neutrality is achieved because these forces balance out in equilibrium. Figure 1.8 illustrates that the strict slow-down in the adoption of good products is due to two effects: On the *extensive margin*, the increase in Λ_0 pushes out t_2^* (i.e. prolongs free-riding in the form of partial adoption); on the *intensive margin*, the increase strictly drives down the growth rate of A_t at all $t < t_2^*(\Lambda_0)$.
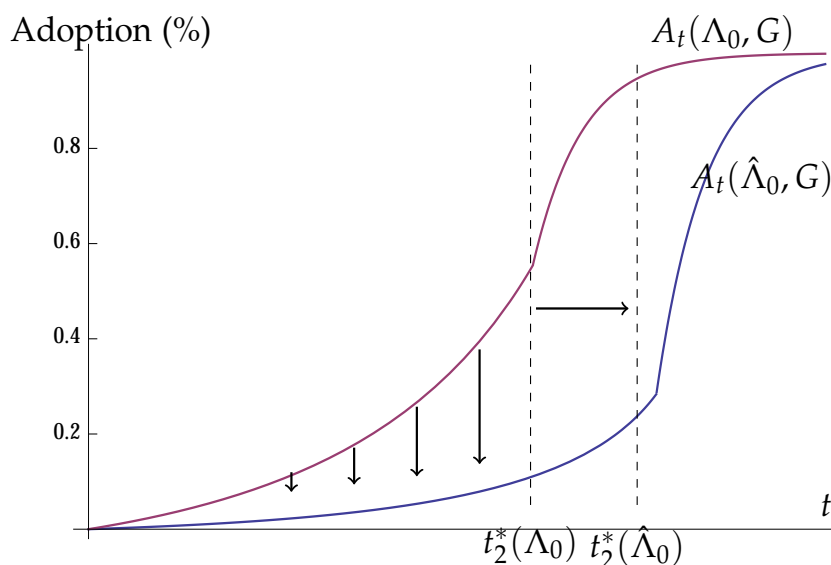
Figure 1.8: *Changes in adoption levels of a good product as a result of increased opportunities for social learning under PBN ($\hat{\Lambda}_0 > \Lambda_0$)*

Since it only arises in the presence of partial adoption, the saturation effect once again relies crucially on the interaction between forward-looking consumers and social learning, setting us apart from models of myopic social learning or forward-looking exogenous learning in which ex-ante welfare necessarily increases in response to more informative signals (*even if* consumers are heterogeneous).[59]

Perfect Good News: Under perfect good news, there is no partial adoption. Correspondingly, there is no saturation effect:[60]

Proposition 1.4.3. *Consider learning via perfect good news. Fix $r, \rho > 0$, $\varepsilon \geq 0$, and $p_0 \in (p^*, 1)$.[61] Suppose $\hat{\Lambda}_0 > \Lambda_0 \geq 0$.[62] Then:*

(i). **Strict Welfare Gains:** *Provided $\varepsilon > 0$,[63] we have $W_0(\hat{\Lambda}_0) > W_0(\Lambda_0)$.*

(ii). *Learning Speeds Up:*

- $0 < t^*(\hat{\Lambda}_0) < t^*(\Lambda_0)$

- $p_t^{\hat{\Lambda}_0} < p_t^{\Lambda_0}$ *for all $t > 0$*

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

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

---

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

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

[61]Recall that $p^* := \frac{(\varepsilon+r)(\rho+r)}{2(\varepsilon+\rho)(\varepsilon+r)-\varepsilon\rho}$ is the equilibrium cutoff posterior under perfect good news. If $p_0 \leq p^*$, then all consumers rely entirely on the exogenous news source from the beginning, so the potential for social learning is irrelevant.

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

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

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

1.4.3 More Social Learning Can Hurt: An Example

Assuming ex-ante identical consumers, Proposition 1.4.2 established a saturation effect under perfect bad news: Beyond a certain level of Λ_0, further increases in the potential for social learning are welfare-neutral. Perhaps even more surprisingly, we show in this section that when consumers are *heterogeneous*, increased opportunities for social learning can bring about *Pareto-decreases* in ex-ante welfare. To illustrate this, we introduce some heterogeneity in consumers' patience levels.

Consider a population consisting of two types of consumers: There is a mass $\bar{N}_0^p$ of patient types with discount rate $r_p > 0$ and a mass $\bar{N}_0^i$ of impatient types with discount rate $r_i > r_p$. Because our aim is simply to construct an example exhibiting welfare loss, we restrict attention to the perfect bad news setting. To simplify the analysis we assume that $\varepsilon = 0$ and $p_0 > 1/2$, but our arguments extend easily to the case where $\varepsilon > 0$.

Recall from Section 1.3.2 that for any discount rate $r > 0$, we can define the function Λ_r^* implicitly for every $p \in (\frac{1}{2}, \frac{\rho + r}{\rho + 2r})$:

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

Suppose $p_0 < \frac{\rho + r_p}{\rho + 2r_p}$ and $\hat{\lambda}\bar{N}_0^p > \lambda\bar{N}_0^p > \Lambda_{r_p}^*(p_0)$ and consider first the game consisting only of mass $\bar{N}_0^p$ consumers of type r_p (and no consumers of type r_i). Then Theorem 1.3.2 implies that the two equilibria corresponding to information structures λ and $\hat{\lambda}$ both feature initial regions of partial adoption, so that $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$.

The following theorem states that provided the mass of impatient types is small,

then in the game consisting of *both* types of consumers, the patient types' ex-ante payoffs continue to be $2p_0 - 1$ under both λ and $\hat{\lambda}$; however, the impatient types' ex-ante payoffs are *strictly* lower under $\hat{\lambda}$ than under λ:

Theorem 1.4.4. *Suppose $0 < r_p < r_i < +\infty$ and $p_0 \in (\frac{1}{2}, \frac{\rho + r_p}{\rho + 2r_p})$. Fix $\bar{N}_0^p > 0$ and $\hat{\lambda} > \lambda > 0$ such that $\hat{\lambda}\bar{N}_0^p > \lambda\bar{N}_0^p > \Lambda_{r_p}^*(p_0)$. Then there exists $\eta > 0$ such that whenever $\bar{N}_0^i < \eta$, then $W_0^i(\hat{\lambda}) < W_0^i(\lambda)$ and $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$. Thus, whenever $\bar{N}_0^i < \eta$, the ex-ante payoff profile $(W_0^i(\lambda), W_0^p(\lambda))$ in the λ-equilibrium Pareto-dominates the ex-ante payoff profile $(W_0^i(\hat{\lambda}), W_0^p(\hat{\lambda}))$ in the $\hat{\lambda}$-equilibrium.*

The proof is in Appendix A.2.6. The basic idea is as follows. Consider first the equilibrium adoption flows that are generated under each of λ and $\hat{\lambda}$ in the game consisting solely of mass $\bar{N}_0^p$ of patient consumers of type r_p. What are the payoffs that a hypothetical impatient type r_i (which does not exist in this game) would obtain if he were to behave optimally when faced with these adoption flows (and the expected future information they imply)? Since the patient types are initially indifferent between adopting or delaying in both equilibria, a monotonicity argument in types shows that in both cases the optimal strategy of the hypothetical impatient type r_i is to adopt upon first opportunity. Given this, the ex-ante payoff of the hypothetical type r_i under signal arrival rate $\gamma \in \{\lambda, \hat{\lambda}\}$ satisfies:

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

By the non-monotonicity result for learning established in Proposition 1.4.2, 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, since $W_0^p(\hat{\lambda}) = W_0^p(\lambda) = 2p_0 - 1$, the cost of the deceleration

in learning on $(t^*, \bar{t})$ and the benefit of the acceleration in learning at times after $\bar{t}$ must balance out in such a way that the patient type r_p obtains the same ex-ante payoff under λ and $\hat{\lambda}$. But as a result, these adjustments must strictly hurt the less patient hypothetical type r_i, because relative to type r_p, type r_i weights the early losses due to the slow-down in learning more heavily than the later benefits due to the acceleration.

To complete the proof, we show that as long as $\bar{N}_0^i > 0$ is sufficiently small, we must still have $W_0^i(\hat{\lambda}) < W_0^i(\lambda)$ and $W_0^p(\hat{\lambda}) = W_0^p(\lambda)$. The first inequality follows from a simple continuity argument. The second equality reflects the fact that provided $\bar{N}_0^i$ is sufficiently small, the patient type must continue to partially adopt initially in both equilibria.

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

1.5 Conclusion

This paper develops a model of innovation adoption when consumers are forward-looking and learning is social. Our analysis isolates the effect of purely informational incentives on aggregate adoption dynamics, learning, and welfare. We highlight the role of the news environment in shaping these incentives; most importantly, in determining whether or not there is informational free-riding in the form of partial adoption. The presence or absence of partial adoption has observable implications, suggesting a novel explanation for why adoption curves are S-shaped for some innovations and concave for others. Moreover, partial adoption has important welfare implications, entailing that increased opportunities for social learning need not benefit consumers and can be strictly harmful.

To illustrate these points in the simplest possible framework, we have restricted attention to perfect bad and good news Poisson learning. This made our equilibrium analysis very tractable, yielding closed-form expressions for all key quantities and allowing us to compute numerous comparative statics. Nevertheless, many of our conclusions extend to more general information structures: Especially worth noting is the fact that partial adoption relies crucially on the possibility of news events that trigger discrete downward jumps in beliefs (although such events need not *conclusively* signal bad quality as was the case under *perfect* bad news). Without such events (e.g. when learning is based on imperfect good news Poisson signals or Brownian motion), a similar logic as in Section 1.3.3 shows that there cannot be continuous regions of partial adoption, because a consumer who is willing to adopt cannot acquire decision-relevant information by delaying his decision by an instant.[64]

---

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

To highlight the implications of purely informational considerations, we have abstracted away from forces emphasized by existing models of innovation adoption, notably consumer heterogeneity and supply-side factors such as pricing. Nevertheless, exploring the way in which these forces *interact* with informational incentives represents an interesting avenue for future theoretical work. To give a taste, Section 1.4.3 shows that heterogeneity can further exacerbate the welfare implications of informational free-riding.

Finally, our predictions lend themselves to empirical investigation. Section 1.4.1 provides some suggestive evidence for the prediction that S-shaped (respectively concave) adoption curves are typical of bad (respectively good) news markets, but a more systematic analysis is called for. The saturation effect implies that the proportion of adopters of an innovation may grow more slowly in communities with more potential consumers or with a greater ease of information transmission. The former could be tested by contrasting the adoption paths of new agricultural technologies across villages with different population sizes,[65] while for the latter one might exploit the staggered introduction of certain social media platforms across different US cities or differences across states in legislation mandating the disclosure of adverse medical events.

---

[65]This is related to Bandiera and Rasul's (2006) finding which we discussed in footnote 2: They find that an individual farmer's likelihood of adoption is (from a certain point on) decreasing in the number of adopters in his network. But then, in equilibrium, larger networks of farmers should feature lower percentages of adoption.

Chapter 2

Rational Behavior under Correlated Uncertainty[1]

2.1 Introduction

In complete information games, Dekel and Fudenberg (1990) and Börgers (1994) have proposed the solution concept $S^\infty W$ (one round of elimination of weakly dominated strategies followed by iterated elimination of strongly dominated strategies), motivating it via its connection with "approximate common certainty" of admissibility. Admissibility (expected utility maximization with respect to some full-support conjecture about opponents' behavior) and iterated admissibility are commonly used refinements of Bayesian rationality (e.g. Luce and Raiffa, 1957; Kohlberg and Mertens, 1986).

Börgers' interest in approximate common certainty of admissibility is driven by epistemic considerations, namely the aim to establish an analog of Tan and

da Costa Werlang's (1988) well-known result that the behavioral implications of common certainty of rationality are given by S^∞ (iterated strong dominance). There is a logical tension between admissibility (holding full-support beliefs about opponents' behavior) and common certainty of admissibility (which in general rules out some opponent strategies). This tension disappears when common certainty is relaxed to approximate common certainty in the sense of common p-belief for p close to 1. Börgers (formalized by Hu (2007)) shows that $S^\infty W$ encapsulates the behavioral implications of the latter notion. Dekel and Fudenberg are motivated by robustness considerations, focusing on the special case where approximate common certainty of admissibility is the result of small amounts of payoff uncertainty. They ask which strategies can arise if players behave according to iterated admissibility, but there is vanishingly small payoff uncertainty, which they model via sequences of elaborations converging to a game. Once again the answer to this question is given by $S^\infty W$.

This paper examines the connection between approximate common certainty of admissibility and $S^\infty W$ in an *incomplete* information setting. Consider a Bayesian game G with state space Θ in which each player i has first-order belief ϕ_i over Θ. We obtain extensions of Börger's and Dekel and Fudenberg's characterizations of $S^\infty W$, but show that these are very sensitive to the way in which uncertainty in the form of approximate common certainty of admissibility is taken to interact with the uncertainty (represented by each player i's belief ϕ_i on Θ) that is already present in G.

Interpreting $S^\infty W$ in the interim-correlated sense of Dekel *et al.* (2007), Section 2.3.1 extends Börgers' characterization: We show that if there is common p-belief of admissibility *and* of the fact that each player i's first-order belief over Θ is *exactly* ϕ_i, then for p close enough to 1, $S^\infty W$ once again emerges as the set of

behavioral implications (Theorem 2.3.1). In Appendix B.1, we provide an analogous extension of Dekel and Fudenberg's result: Proposition B.1.2 shows that $S^\infty W$ is the robust extension of W^∞ under elaborations in which "sane" types may assign vanishingly small probability to opponents being "crazy" (i.e. having very different payoffs and beliefs as in the original game), but must themselves have *exactly* the same beliefs (and payoffs) as in the original game.

However, these results break down when approximate common certainty of admissibility is accompanied by vanishingly small perturbations to players' beliefs about states. In Section 2.3.2 we show that if there is common p-belief of admissibility and of the fact that each player i's first-order belief about Θ is "approximately" ϕ_i, then the behavioral implications are given by a generalization of Hu's (2007) perfect p-rationalizable set (Theorem 2.3.5). But even in the limit as p goes to 1 and as the uncertainty about ϕ_i becomes vanishingly small, this set is in general a strict superset of $S^\infty W$.[2]

In addition to Börgers (1994), Dekel and Fudenberg (1990), and Hu (2007), our paper connects more broadly with the growing literatures on epistemic conditions related to admissibility and on the robustness of solution concepts to small amounts of uncertainty. Within the former literature, an alternative epistemic characterization of $S^\infty W$ is obtained by Schuhmacher (1999): Instead of relaxing common certainty to common p-belief, he relaxes the rationality requirement implied by admissibility to "ε-rationality". Brandenburger *et al.* (2008) replace probabilistic beliefs with lexicographic probability systems to obtain epistemic foundations for iterated admissibility.[3] Both these papers are set in a complete information en-

---

[2]Similarly, under elaborations in which sane types' beliefs over Θ are allowed to be vanishingly small perturbations of the beliefs in G, $S^\infty W$ is in general a strict subset of the robust closure of W^∞.

[3]See Dekel and Siniscalchi (2014) for a more comprehensive discussion of this literature.

vironment. Within the latter literature, which was pioneered by Fudenberg *et al.* (1988) and Kajii and Morris (1997), our paper relates most closely to Weinstein and Yildiz (2007): They show that in incomplete information games satisfying a richness assumption, any action in S^∞ can be made *uniquely* rationalizable if we allow for *arbitrary* perturbations to players' entire hierarchy of beliefs. Put differently, this means that any refinement of S^∞ must be derived from restrictions on the class of perturbations that we are considering. One of the contributions of our paper lies in clarifying the nature of restrictions needed to obtain $S^\infty W$.

The rest of the paper is organized as follows. Section 2.2 introduces notation and definitions. Section 2.3 contains our main results. Section 2.4 concludes. Appendix B.1 extends Dekel and Fudenberg's characterization to our setting, and Appendix B.2 contains all proofs.

2.2 Preliminaries

Throughout the paper we fix an incomplete information normal-form game $G = (I, \Theta, (A_i, u_i, \phi_i)_{i \in I})$, where Θ is a finite set of states of nature, I is a finite set of players, and for every $i \in I$, A_i is a finite set of strategies, $u_i \colon A_i \times A_{-i} \times \Theta \to \mathbb{R}$ denotes i's state-dependent payoffs, and $\phi_i \in \Delta^\circ(\Theta)$ is i's full-support belief about Θ.[4] We work with the following notions of admissibility, iterated dominance, and approximate common certainty of admissibility:

---

[4]For expositional clarity, we assume that in G there is a single type for each player. With slight adjustments, our results extend readily to games with finite type spaces.

2.2.1 Interim Correlated Dominance

We interpret strong and weak dominance in the "interim correlated" sense of Dekel *et al.* (2007).[5,6] This allows any player's conjectures about her opponents' behavior to be correlated with her beliefs about the state of nature. As argued by Dekel *et al.* (2007), allowing for this kind of correlation is the natural approach in Bayesian games, because imposing independence on conjectures about states and opponents' behavior produces solution concepts which are very sensitive to "redundant" aspects of the type space.[7] Nevertheless, as we discuss in Section 2.4, imposing the latter kind of independence on our definition of iterated dominance would not fundamentally alter our results in Section 2.3.

Definition 2.2.1. Consider non-empty subsets $\hat{A}_j \subseteq A_j$ for every j, and let $\hat{A} := \prod_j \hat{A}$ and $\hat{A}_{-j} := \prod_{i \neq j} \hat{A}_i$. We say that $\alpha_i \in \Delta(\hat{A}_j)$ **strongly dominates** $a_i \in \hat{A}_i$ on $\hat{A}$ if for all beliefs $\mu_i \colon \Theta \to \Delta(\hat{A}_{-i})$,

$$\sum_{\theta, a_{-i}} \phi_i(\theta) \mu_i(\theta)[a_{-i}] (u_i(\alpha_i, a_{-i}, \theta) - u_i(a_i, a_{-i}, \theta)) > 0.$$

$\alpha_i \in \Delta(\hat{A}_i)$ **weakly dominates** $a_i \in \hat{A}_i$ on $\hat{A}$ if for all beliefs $\mu_i \colon \Theta \to \Delta(\hat{A}_{-i})$,

$$\sum_{\theta, a_{-i}} \phi_i(\theta) \mu_i(\theta)[a_{-i}] (u_i(\alpha_i, a_{-i}, \theta) - u_i(a_i, a_{-i}, \theta)) \geq 0,$$

with strict inequality for at least one such μ_i.

Denote by $S(\hat{A})_i$ the set of all strategies $a_i \in \hat{A}_i$ which are not strongly dominated on $\hat{A}$ by any $\alpha_i \in \Delta(\hat{A}_i)$, and let $S(\hat{A}) := \prod_j S(\hat{A})_j$. Let $S^1(\hat{A}) := S(\hat{A})$,

---

[5]See also Battigalli and Siniscalchi (2003).

[6]Our notion of interim correlated weak dominance, which does not appear in Dekel *et al.* (2007), is derived from their concept of strong dominance in the natural way.

[7]See Dekel *et al.* (2007) for details and examples.

$S^{k+1}(\hat{A}) := S(S^k(\hat{A}))$ and $S^\infty(\hat{A}) := \cap_k S^k(\hat{A})$. Similarly, denote by $W(\hat{A})_i$ the set of all **admissible** strategies for player i on $\hat{A}$, i.e. all $a_i \in \hat{A}_i$ which are not weakly dominated on $\hat{A}$ by any $\alpha_i \in \Delta(\hat{A}_i)$, and let $W(\hat{A}) := \prod_j W(\hat{A})_j$. Let $W^1(\hat{A}) := W(\hat{A})$, $W^{k+1}(\hat{A}) := W(W^k(\hat{A}))$ and $W^\infty(\hat{A}) := \cap_k W^k(\hat{A})$. If $\hat{A}_j = A_j$ for all j, we write $S^k(G)$, $W^k(G)$ and $S^l W^k(G)$ for $S^k(\hat{A})$, $W^k(\hat{A})$ and $S^l W^k(\hat{A})$, respectively.

The following lemma is an incomplete-information extension of Pearce's well-known result (Pearce, 1984): A strategy is not strongly dominated if and only if it is a best response to some state-dependent belief about opponent strategies, and it is admissible if and only if it is a best response to a state-dependent belief having *full support* on the relevant set of opponent strategies in every state:

Lemma 2.2.2 (Equivalence of undominance and best response formulations). *Let $\hat{A}_i$ and $\hat{A}_{-i}$ be nonempty subsets of A_i and A_{-i}, respectively. Suppose $a_i \in \hat{A}_i$. Then:*

(i). *$a_i \in S(\hat{A})_i$ if and only if a_i is a best response in $\hat{A}_i$ to some $\lambda_i \colon \Theta \to \Delta(\hat{A}_{-i})$.*[8]

(ii). *$a_i \in W(\hat{A})_i$ if and only if a_i is a best response in $\hat{A}_i$ to some $\lambda_i \colon \Theta \to \Delta^\circ(\hat{A}_{-i})$.*

2.2.2 Approximate Common Certainty of Admissibility

To model approximate common certainty of admissibility, we associate with G an epistemic type structure: We let $\mathcal{T} = \big(I, (\Theta \times A_{-i}, T_i, \beta_i)_{i \in I}\big)$ be the universal type space in which each player i's basic space of uncertainty is $\Theta \times A_{-i}$.[9] Then each set of types T_i is a compact metric space and for each i, the belief $\beta_i \colon T_i \to \Delta(\Theta \times A_{-i} \times T_{-i})$ is a homeomorphism.[10] We will use $\mathcal{T}$ to model small amounts

---

[8]I.e., $a_i \in \operatorname{argmax}_{a_i' \in \hat{A}_i} \sum_{\theta, a_{-i}} \phi_i(\theta) \lambda_i(\theta)[a_{-i}] u_i(a_i', a_{-i}, \theta)$.

[9]Cf. Brandenburger and Dekel (1993) for details regarding the construction of this space.

[10]For any compact metric space M, we equip the set $\Delta(M)$ of probability measures on the Borel σ-algebra of M with the weak* topology, under which $\Delta(M)$ is itself compact metric.

of uncertainty about the fact that players choose admissible strategies. In our incomplete information setting, it is natural to allow this uncertainty to interact with the uncertainty already present in G: Player i's uncertainty about opponents' behavior might be accompanied by small doubts about opponents' beliefs about Θ and might be correlated with i's own beliefs about Θ. This is captured in $\mathcal{T}$ by the fact that each $\beta_i(t_i)$ is a joint probability distribution on $\Theta \times A_{-i} \times T_{-i}$ and that $\mathrm{marg}_\Theta \beta_j(t_j)$ is not necessarily equal to ϕ_j. Let $f_i := \mathrm{marg}_{\Theta \times A_{-i}} \beta_i$, let $\Omega := \Theta \times \Pi_{i \in I}(A_i \times T_i)$, and denote by θ, a_i and t_i the projections from Ω onto Θ, A_i and T_i, respectively.

Common p-belief: For any event $E \subseteq \Omega$,[11] and any $\omega \in \Omega$, let $E^\omega := \{(\theta, a_{-i}, t_{-i}) \in \Theta \times A_{-i} \times T_{-i} : (\theta, a_i(\omega), t_i(\omega), a_{-i}, t_{-i}) \in E\}$. For any $p \in (0,1]$, we define the following events:

- i p-believes E: $B_i^p(E) := \{\omega \in \Omega : \beta_i(t_i(\omega))(E^\omega) \geq p\}$

- mutual p-belief of E: $B^p(E) := \bigcap_{i \in I} B_i^p(E)$

- $B^{p,1}(E) := B^p(E)$ and inductively $B^{p,n+1}(E) := B^p(B^{p,n}(E))$ for $n \geq 1$

- common p-belief of E: $\mathrm{CB}^p(E) := \bigcap_{n \in \mathbb{N}} B^{p,n}(E)$.

Approximate common certainty (ACC) of admissibility: Define the event that **player i is rational** given utility u_i[12] by

$$
\mathcal{R}_i := \left\{ \omega \in \Omega : a_i(\omega) \in \mathrm{argmax}_{a_i' \in A_i} \sum_{\theta \in \Theta, a_{-i} \in A_{-i}} f_i(t_i(\omega))(\theta, a_{-i}) u_i(a_i', a_{-i}, \theta) \right\},
$$

and the event that all players are rational by $\mathcal{R} := \bigcap_{i \in I} \mathcal{R}_i$. Define the event that **player i's first-order belief has full support** by $\mathcal{P}_i :=$

---

[11] An event is a measurable set w.r.t. the Borel σ-algebra on Ω.

[12] In Section 2.4 we briefly discuss the implications of allowing for perturbations of the utility.

$\{\omega \in \Omega : f_i(t_i(\omega)) \in \Delta^\circ(\Theta \times A_{-i})\}$, and that all players have full-support first-order beliefs by $\mathcal{P} := \bigcap_{i \in I} \mathcal{P}_i$. Then $\mathcal{R} \cap \mathcal{P}$ is the event that all players play admissible strategies.

In the complete information setting where $|\Theta| = 1$, Börgers (1994) and Hu (2007) model approximate common certainty (henceforth ACC) of admissibility by $CB^p(\mathcal{R} \cap \mathcal{P})$ as $p \to 1$, capturing the idea that players' uncertainty (and higher-order uncertainty) about opponents' choice of admissible strategies becomes vanishingly small. In the case where $|\Theta| > 1$, we allow player's doubts about opponents' behavior to be accompanied by doubts about opponents' beliefs about Θ (in the sense that we do not impose that $\mathrm{marg}_\Theta \beta_j(t_j) = \phi_j$). Correspondingly, a natural definition of ACC should require both types of doubts to vanish.

We consider two ways of modeling this: Define the event that **player i's first-order beliefs on Θ are ϕ_i** by $[\phi_i] := \{\omega \in \Omega : \mathrm{marg}_\Theta \beta_i(t_i(\omega)) = \phi_i\}$, and let $[\phi] := \bigcap_{i \in I}[\phi_i]$. For any $\varepsilon > 0$, define the event that **player i's first-order beliefs on Θ are ε-close to ϕ_i** by $[\phi_i, \varepsilon] := \{\omega \in \Omega : ||\mathrm{marg}_\Theta \beta_i(t_i(\omega)) - \phi_i||_\infty \leq \varepsilon\}$, and let $[\phi, \varepsilon] := \bigcap_{i \in I}[\phi_i, \varepsilon]$.

In Section 2.3.1 we consider the behavioral implications of **strong ACC of admissibility**, defined as $\mathrm{SACCA}(G) := \bigcap_{p \in (0,1)} \mathrm{Proj}_A CB^p([\phi] \cap \mathcal{R} \cap \mathcal{P})$, which imposes common p-belief of the *exact* profile of priors ϕ. Section 2.3.2 considers the behavioral implications of **ACC of admissibility with perturbed priors**, defined as

$$\mathrm{PACCA}(G) := \lim_{\bar{p} \to 1^-, \bar{\varepsilon} \to 0^+} \bigcap_{p \in [\bar{p},1), \varepsilon \in (0,\bar{\varepsilon}]} \mathrm{Proj}_A CB^p([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P}),^{13}$$

which only imposes common p-belief of the fact that priors on Θ are perturbations

---

[13]Note that whenever $0 < p \leq p' < 1$ and $\varepsilon \geq \varepsilon' > 0$, then $\mathrm{Proj}_A CB^p([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P}) \supseteq \mathrm{Proj}_A CB^{p'}([\phi, \varepsilon'] \cap \mathcal{R} \cap \mathcal{P})$. Thus, by finiteness of A there is $\bar{p} \in (0,1)$ and $\bar{\varepsilon} > 0$ such that $\mathrm{PACCA}(G) = \mathrm{Proj}_A CB^{\bar{p}}([\phi, \bar{\varepsilon}] \cap \mathcal{R} \cap \mathcal{P}) = \mathrm{Proj}_A CB^p([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P})$ for all $(p, \varepsilon) \in [\bar{p}, 1) \times (0, \bar{\varepsilon}]$.

of ϕ, where we require these perturbations to vanish in the limit as $p \to 1$.

Conditional p-belief: In the following, we will also need to consider the event that *in every state θ* there is common p-belief in admissibility. For this we introduce the notion of *conditional p-belief:* For any family $\mathcal{F} := \{F_1, \ldots, F_N\}$ of events, define the following events:

- i p-believes E conditional on F_k:

$$B_i^p(E \mid F_k) := \left\{ \omega \in \Omega \;\middle|\; \beta_i(t_i(\omega))(F_k^\omega) > 0 \text{ and } \frac{\beta_i(t_i(\omega))(E^\omega \cap F_k^\omega)}{\beta_i(t_i(\omega))(F_k^\omega)} \geq p \right\}$$

- i p-believes E conditional on each F_k: $B_i^p(E \mid \mathcal{F}) := \bigcap_{n \in \{1,\ldots,N\}} B_i^p(E \mid F_k)$

- mutual p-belief conditional on each F_k: $B^p(E \mid \mathcal{F}) := \bigcap_{i \in I} B_i^p(E \mid \mathcal{F})$

- $B^{p,1}(E \mid \mathcal{F}) := B^p(E \mid \mathcal{F})$ and $B^{p,n+1}(E \mid \mathcal{F}) := B^p(B^{p,n}(E \mid \mathcal{F}) \mid \mathcal{F})$ for all $n \geq 1$

- common p-belief conditional on each F_k: $CB^p(E \mid \mathcal{F}) := \bigcap_{n \in \mathbb{N}} B^{p,n}(E \mid \mathcal{F})$.

Letting $E(\theta) := \{\omega \in \Omega \mid \theta(\omega) = \theta\}$ and $E(\Theta) := \{E(\theta)\}_{\theta \in \Theta}$, the event that i p-believes E in every state θ is then given by $B_i^p(E \mid E(\Theta))$. We will make use of the following relationship between conditional and unconditional p-beliefs:

Lemma 2.2.3. *For any $i \in I$, $p \in (0,1]$, partition $\mathcal{F} = \{F_1, \ldots, F_N\}$ of Ω, and $\theta \in \Theta$:*

- $B_i^p(E \mid \mathcal{F}) \subseteq B_i^p(E)$*; and*

- $B_i^p(E) \cap [\phi_i] \subseteq B_i^{\max\left\{1 - \frac{1-p}{\phi_i(\theta)}, 0\right\}}(E \mid E(\theta))$.

2.3 $S^\infty W$ and ACC of Admissibility

In a complete information setting, Börgers (1994) (formalized by Hu (2007)) shows that for large enough p, we have that $S^\infty W(G) = \text{Proj}_A CB^p(\mathcal{R} \cap \mathcal{P})$. In this section

we examine the validity of this result under the two incomplete-information notions
of ACC of admissibility defined in Section 2.2.2.

2.3.1 Strong ACC of Admissibility

Consider first the behavioral implications of *strong* ACC of admissibility, as defined
by the set SACCA(G) in Section 2.2.2. In this case, we obtain an incomplete
information extension of Börgers's (1994) characterization:

Theorem 2.3.1. *There exists $\bar{p} \in (0,1)$ such that for all $p \in [\bar{p},1)$ we have*

$$S^{\infty}W(G) = \mathrm{Proj}_A \, \mathrm{CB}^p \left([\phi] \cap \mathcal{R} \cap \mathcal{P}\right) = \mathrm{SACCA}(G). \tag{2.1}$$

The proof of Theorem 2.3.1 proceeds in two steps, both of which make use of an
incomplete information analog of Hu's (2007) perfect p-rationalizable set: For any
$A' := \prod_{i \in i} A'_i \subseteq A$ and $i \in I$, let

$$D_i^p(A') := \left\{ a_i \in A'_i : \exists \mu \colon \Theta \to \Delta^{\circ}(A_{-i}) \text{ s.t. } a_i \in \mathrm{BR}_{A'_i}^{\phi_i}(\mu) \text{ and } \mu(\theta)(A'_{-i}) \geq p \, \forall \theta \in \Theta \right\},$$

where $BR_{A'_i}^{\phi_i}(\mu) := \mathrm{argmax}_{a'_i \in A'_i} \sum_{\theta \in \Theta, a_{-i} \in A_{-i}} \phi_i(\theta)\mu(\theta)(a_{-i})u_i(a'_i, a_{-i}, \theta)$. Inductively
define $\tilde{\Lambda}^{p,0}(A) = D^p(A) := \prod_{i \in I} D_i^p(A)$ and $\tilde{\Lambda}^{p,n+1} := D^p(A_i, \tilde{\Lambda}_{-i}^{p,n}(A))$ for all $n \geq$
0. Then the **perfect p-rationalizable set** of G is given by $\tilde{R}^p(G) := \bigcap_{n \in \mathbb{N}} \tilde{\Lambda}^{p,n}(A)$.

The first step in the proof of Theorem 2.3.1 shows that for large enough p,
the perfect p-rationalizable set $\tilde{R}^p(G)$ is precisely $S^{\infty}W(G)$. This follows from an
extension of the key lemma from Börgers (1994) to our incomplete information
setting:

Lemma 2.3.2. *There exists $\pi \in (0,1)$ such that for all $i \in I$, $p \in [\pi,1)$, and non-empty
$\hat{A}_i \subseteq A_i$ and $\hat{A}_{-i} \subseteq A_{-i}$, we have that $a_i \in \hat{A}_i$ is in $D_i^p(\hat{A}_i, \hat{A}_{-i})$ if and only if the
following two conditions hold:*

(i). $a_i \in W(\hat{A}_i \times A_{-i})_i$, *i.e. there is no $\alpha_i \in \Delta(\hat{A}_i)$ which weakly dominates a_i on A_{-i}.*

(ii). $a_i \in S(\hat{A}_i \times \hat{A}_{-i})_i$, *i.e. there is no $\alpha_i \in \Delta(\hat{A}_i)$ which strongly dominates a_i on $\hat{A}_{-i}$.*

Corollary 2.3.3. *There exists $\pi \in (0,1)$ such that for all $i \in I$ and $p \in [\pi, 1)$, we have $\tilde{R}_i^p(G) = S^\infty W(G)_i$.*

The second step in proving Theorem 2.3.1 is an extension of Theorem 5.1 from Hu (2007) to our incomplete information setting. We show that for any p, $\tilde{R}^p(G)$ is the set of strategy profiles which are played when *in every state* there is common p-belief of the event $[\phi] \cap \mathcal{R} \cap \mathcal{P}$:

Proposition 2.3.4. *For all $p \in (0,1)$, $\tilde{R}^p(G) = \mathrm{Proj}_A \mathrm{CB}^p \left([\phi] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta) \right)$.*

For any $p \in (0,1)$, Proposition 2.3.4 combined with the first part of Lemma 2.2.3 implies $\tilde{R}^p(G) \subseteq \mathrm{Proj}_A \mathrm{CB}^p \left([\phi] \cap \mathcal{R} \cap \mathcal{P} \right)$. And for any $p \in (0,1)$ such that $1 - \frac{1-p}{\min_{i \in I, \theta \in \Theta} \phi_i(\theta)} > 0$, Proposition 2.3.4 along with the second part of Lemma 2.2.3 implies

$$\mathrm{Proj}_A \mathrm{CB}^p \left([\phi] \cap \mathcal{R} \cap \mathcal{P} \right) \subseteq \tilde{R}^{1 - \frac{1-p}{\min_{i \in I, \theta \in \Theta} \phi_i(\theta)}}(G).$$

Letting $\pi \in (0,1)$ be as in Corollary 2.3.3, there exists $\bar{p} \in (\pi, 1)$ such that for all $p \geq \bar{p}$ we have $1 - \frac{1-p}{\min_{i \in I, \theta \in \Theta} \phi_i(\theta)} \geq \pi$. Then for all $p \geq \bar{p}$, $\tilde{R}^{1 - \frac{1-p}{\min_{i \in I, \theta \in \Theta} \phi_i(\theta)}}(G) = \tilde{R}^p(G) = S^\infty W(G)$. Hence, $\mathrm{Proj}_A \mathrm{CB}^p \left([\phi] \cap \mathcal{R} \cap \mathcal{P} \right) = \tilde{R}^p(G) = S^\infty W(G)$, completing the proof of Theorem 2.3.1.

2.3.2 ACC of Admissibility with Perturbed Priors

We now study the behavioral implications of ACC of admissibility with *perturbed priors*, as defined by the set PACCA(G) in Section 2.2.2. In this case, Theorem 2.3.1 breaks down: $S^\infty W(G)$ continues to be a subset of PACCA(G), but in general the containment is strict.

To see this, we first provide a characterization of $\text{PACCA}(G)$ in terms of a generalization of the perfect p-rationalizable set. Define the **ε-perturbed perfect p-rationalizable set** to be $\tilde{R}^{\varepsilon,p}(G) := \bigcap_{n \in \mathbb{N}} \tilde{\Lambda}^{\varepsilon,p,n}(A)$, where for any $A' := \prod_{i \in i} A_i' \subseteq A$ and $i \in I$, we let

$$\tilde{\Lambda}_i^{\varepsilon,p}(A') := \left\{ a_i \in A_i : \begin{array}{l} \exists \phi_i' \in \Delta(\Theta) \ \& \ \mu \colon \Theta \to \Delta^\circ(A_{-i}) \text{ s.t.} \\ a_i \in \text{BR}_{A_i}^{\phi_i'}(\mu) \ \& \ ||\phi_i - \phi_i'||_\infty \leq \varepsilon \ \& \ \mu(\theta)(A_{-i}') \geq p \ \forall \theta \in \Theta \end{array} \right\},$$

and we inductively define $\tilde{\Lambda}^{\varepsilon,p,0}(A) = \tilde{\Lambda}^{\varepsilon,p}(A) := \prod_{i \in I} \tilde{\Lambda}_i^{\varepsilon,p}(A)$ and $\tilde{\Lambda}^{\varepsilon,p,n+1} := \tilde{\Lambda}^{\varepsilon,p}(\tilde{\Lambda}^{p,n}(A))$ for all $n \geq 0$. We have the following analog of Proposition 2.3.4:

Theorem 2.3.5. *For all* $p \in (0,1)$ *and* $\varepsilon > 0$, $\text{Proj}_A \text{CB}^p([\phi,\varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) = \tilde{R}^{\varepsilon,p}(G)$.

By the same logic as in the paragraph following Proposition 2.3.4, Theorem 2.3.5 implies that there exists $\bar{p} \in (0,1)$ and $\bar{\varepsilon} > 0$ such that for all $p \in [\bar{p},1)$ and $\varepsilon \in (0,\bar{\varepsilon}]$, $\tilde{R}^{\varepsilon,p}(G) = \tilde{R}^{\bar{\varepsilon},\bar{p}}(G) = \text{PACCA}(G)$. Since $\tilde{R}^p(G) \subseteq \tilde{R}^{\varepsilon,p}(G)$ for all $p \in (0,1)$ and $\varepsilon > 0$, it follows from Theorem 2.3.1 that $S^\infty W(G) \subseteq \text{PACCA}(G)$. However, as the following example shows, this inclusion is generally strict:

Example 2.3.6. Consider the 2-player game G where $A_1 = \{U,D\}$, $A_2 = \{L,R\}$, $\Theta = \{\theta, \theta'\}$, player 2 is indifferent across all outcomes and states, and 1's payoffs are given by:

θ	L	R
U	1	1
D	0	0

θ'	L	R
U	0	1
D	1	0

Suppose that $\phi_1 = \phi_2 =: \phi$ assigns equal probability to θ and θ'. Then $D \notin S^\infty W(G)_1$ because it is weakly dominated by U under ϕ: Indeed, writing XY (for

any $X, Y \in A_2$) to denote the belief $\mu \colon \Theta \to A_2$ such that $\mu(\theta) = X$ and $\mu(\theta') = Y$, player 1's expected payoffs from U and D given ϕ are summarized by

	LL	RL	LR	RR
U	$\frac{1}{2}$	$\frac{1}{2}$	1	1
D	$\frac{1}{2}$	$\frac{1}{2}$	0	0

On the other hand, consider the sequence of perturbed beliefs $\phi^n \to \phi$ given by $\phi^n(\theta) = \frac{1}{2} - \frac{1}{n}$, $\phi^n(\theta') = \frac{1}{2} + \frac{1}{n}$ for $n \geq 3$, and consider the sequence of conjectures $\mu^n \colon \Theta \to \Delta^\circ(A_2)$ given by $\mu^n(\theta)[L] = \mu^n(\theta')[L] = 1 - \frac{1}{n^2}$, $\mu^n(\theta)[R] = \mu^n(\theta')[R] = \frac{1}{n^2}$. Given ϕ^n, 1's expected payoffs against μ^n from U and D are $\frac{1}{2} - \frac{1}{n} + (\frac{1}{2} + \frac{1}{n})\frac{1}{n^2}$ and $(\frac{1}{2} + \frac{1}{n})(1 - \frac{1}{n^2})$, respectively. For large enough n, the latter is strictly greater than the former, so $D \in BR^{\phi^n}(\mu^n)$. Thus, $\tilde{\Lambda}_1^{\frac{1}{n}, p, 0}(A) = \{U, D\} = A_1$. Hence, for all $p \in (0, 1)$ and $\varepsilon > 0$, $\tilde{R}^{\varepsilon, p}(G) = A \supsetneq S^\infty W(G)$.

The preceding example has the special feature that $\tilde{R}^{\frac{1}{n}, p}(G) = \mathrm{Proj}_{A_1} CB^p([\phi^n] \cap \mathcal{R} \cap \mathcal{P})$ for all $p \in (0, 1)$ and large enough n. However, in general $\tilde{R}^{\varepsilon, p}(G)$ does not impose common p-belief in any single profile (ϕ'_1, ϕ'_2) of ε-perturbed priors, instead allowing each player to be uncertain about the way in which his opponents' priors are perturbed, about the way in which his opponents think his priors might be perturbed etc. This additional uncertainty can generate additional behavior outside $S^\infty W(G)$: A given strategy might be played by player 1 only if player 1's belief about Θ is perturbed in a particular way ϕ'_1 and if player 1 believes that player 2 plays a particular strategy—but the latter choice of strategy by player 2 might in turn require player 2 to believe that player 1 is willing to play some other strategy which player 1 would only play under a different belief perturbation ϕ''_1. The following example illustrates this:

60

Example 2.3.7. Consider the 2-player game G where $A_1 = \{U, I, D\}$, $A_2 = \{L, M, R\}$, $\Theta = \{\theta, \theta'\}$, and payoffs are summarized by the following two tables:

θ	L	M	R
U	$1,0$	$0,\frac{1}{2}$	$0,1$
I	$\frac{1}{2},0$	$\frac{1}{2},\frac{1}{2}$	$\frac{1}{2},0$
D	$0,0$	$0,\frac{1}{2}$	$0,0$

θ'	L	M	R
U	$0,1$	$0,\frac{1}{2}$	$0,0$
I	$\frac{1}{2},0$	$\frac{1}{2},\frac{1}{2}$	$\frac{1}{2},0$
D	$1,0$	$0,\frac{1}{2}$	$0,0$

Suppose that $\phi_1 = \phi_2 =: \phi$ assigns equal probability to θ and θ'. Then for any $p \in (0,1)$ and $\varepsilon > 0$, we have $\tilde{R}^{\varepsilon,p}(G) = A$: Indeed, note first that under the belief $\phi_1' \in \Delta(\Theta)$ with $\phi_1'(\theta) = \alpha$, player 1's expected payoffs from U, I and D are summarized by the following table (again XY denotes the belief that X is played in θ and Y is played in θ'):

	LL	LR	LM	RR	RL	RM	MM	MR	ML
U	α	α	α	0	0	0	0	0	0
I	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
D	$1-\alpha$	0	0	0	$1-\alpha$	0	0	0	$1-\alpha$

From this it is easy to see that $\tilde{\Lambda}_1^{\varepsilon,p,0}(A) = \{U, I, D\} = A_1$, because U is rationalizable under any ϕ_1' with $\alpha > \frac{1}{2}$ and D is rationalizable under any ϕ_1' with $\alpha < \frac{1}{2}$ (and I is a weakly dominant strategy). Similarly, under the belief $\phi_2' \in \Delta(\Theta)$ with $\phi_2'(\theta) = \alpha$, player 2's expected payoffs from L, R, and M are summarized by:

	UU	UI	UD	IU	II	ID	DU	DI	DD
L	$1-\alpha$	0	0	$1-\alpha$	0	0	$1-\alpha$	0	0
M	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
R	α	α	α	0	0	0	0	0	0

Thus, by the same reasoning as above $\tilde{\Lambda}_2^{\varepsilon,p,0}(A) = \{L, M, R\} = A_2$, whence $\tilde{R}^{\varepsilon,p}(G) = A$.

On the other hand, fix $\varepsilon > 0$ sufficiently small and choose $p = p(\varepsilon)$ sufficiently close to 1. We claim that there is no $\phi' := (\phi'_1, \phi'_2) \in \Delta^\circ(\Theta)$ with $||\phi' - \phi|| \leq \varepsilon$ such that $D \in \mathrm{Proj}_{A_1} CB^p([\phi'] \cap \mathcal{R} \cap \mathcal{P})$. The basic intuition is the following: An easy adaptation of the arguments in Section 2.3.1 shows that if ε is sufficiently small and p sufficiently close to 1, then for all ϕ' with $||\phi' - \phi|| \leq \varepsilon$, we have $\mathrm{Proj}_{A_1} CB^p([\phi'] \cap \mathcal{R} \cap \mathcal{P}) = \tilde{R}^p(G_{\phi'})$, where $G_{\phi'}$ is the incomplete information game with the same states of nature, actions and payoffs as G, but with belief profile ϕ' rather than ϕ. Now note that for any ϕ' such that $D \in \tilde{\Lambda}_1^{p,0}(G_{\phi'})$, we must have $\phi'_1(\theta) < \frac{1}{2}$ and hence $U \notin \tilde{\Lambda}_1^{p,0}(G_{\phi'})$. But for $p > \frac{1}{2}$, $U \notin \tilde{\Lambda}_1^{p,0}(G_{\phi'})$ implies $L \notin \tilde{\Lambda}_2^{p,1}(G_{\phi'})$: Indeed, for any conjecture $\mu : \Theta \to \Delta^\circ(A_1)$ which in each state puts probability greater than $\frac{1}{2}$ on $\tilde{\Lambda}_1^{p,0}(G_{\phi'})$ (and hence on player 1 not playing U), player 2 is strictly better off playing M than L. But then, at the next step of the iteration, this means that $D \notin \tilde{\Lambda}_1^{p,2}(G_{\phi'})$: If player 1 puts probability greater than $\frac{1}{2}$ on opponent strategies in $\tilde{\Lambda}_2^{p,1}(G_{\phi'})$ (and hence on L not being played), then he is strictly better off playing I than D.

2.4 Discussion

Our positive result in Section 2.3.1 shows that in an incomplete information setting, $S^\infty W(G)$ continues to admit clean epistemic foundations. However, from a modeler's perspective, the discussion in Section 2.3.2 suggests caution in making predictions based on $S^\infty W$, except in situations where common p-belief of an *exact* profile of priors seems a reasonable assumption—for example, this might be the case if the game is played in an experimental lab setting and prior to the start of the game the experimenter publicly announces an "objective" probability distribution ϕ on Θ to all players. Additional interpretations of ACC of admissibility naturally spring to mind: For instance, if we further weaken ACC of admissibility to allow players to entertain vanishingly small doubts about opponents' state-dependent payoffs,[14] then it is easy to see that any action profile in S^∞ can be predicted.[15] Once again, one might argue that in many situations a modeler would be hard-pressed to rule out uncertainty in the form of such vanishingly small perturbations. If so, this suggests caution in applying any refinement of rationality short of S^∞.[16]

Finally, we note that while we have interpreted $S^\infty W$ in the interim-correlated sense of Dekel *et al.* (2007), working with an alternative notion of $S^\infty W$ which requires each player's beliefs about opponents' behavior and about the states of

---

[14] That is, we replace the event $\mathcal{R}_i$ for each player with the event $\mathcal{R}_i^\varepsilon$ which requires $a_i(\omega)$ to be a best-response given beliefs $f_i(t_i(\omega))$ and given a state-dependent utility function u_i' such that $||u_i' - u_i||_\infty \leq \varepsilon$.

[15] This notion can be regarded as as an incomplete information analog of the notion of "weak convergence" also considered in Dekel and Fudenberg (1990), but since they study normal forms derived from extensive form games they restrict to payoff perturbations which respect the associated extensive form.

[16] In this sense, our conclusion would be similar in spirit to Weinstein and Yildiz (2007). However, our point is simply that any given action a in S^∞ can arise under ACC of admissibility with perturbed payoffs, not that a will be *uniquely* rationalizable. To make this point, we can work with a more limited subset of perturbations of the type space than Weinstein and Yildiz (2007) and we do not need to impose any richness assumption on the game.

nature to be *independent* would not fundamentally alter our results: Let $S^\infty W_I$ denote this independent version of $S^\infty W$[17] and let $\mathcal{H}$ denote the event that each player's first-order beliefs are independent across opponents' behavior and states of nature.[18] Then analogously to Theorem 2.3.1 above, we can establish that $S^\infty W_I(G)$ coincides with $\text{Proj}_A \, CB^p([\phi] \cap \mathcal{R} \cap \mathcal{P} \cap \mathcal{H})$ for p close enough to 1. We can also obtain an analog of Theorem 2.3.5 under a suitably modified definition of the ε-perturbed perfect p-rationalizable set. Finally, it is easy to see that in Example 2.3.6 we also have $\text{Proj}_A \, CB^p([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \cap \mathcal{H}) = A \supsetneq S^\infty W_I(G)$ for all $p \in (0,1)$ and $\varepsilon > 0$.

---

[17]Formally, $S^\infty W_I(G)$ is obtained by imposing on each instance of $\mu_i \colon \Theta \to \Delta(\hat{A}_{-i})$ in Definition 2.2.1 the requirement that $\mu_i(\theta) = \mu_i(\theta')$ for all θ, θ'. Note that we do allow for correlation across opponents' actions, because $\mu_i(\theta)$ is not assumed to be a product measure; thus our definition is less stringent than "interim independence" in the sense of Dekel *et al.* (2007).

[18]$\mathcal{H}_i := \{\omega : f_i(t_i(\omega))(\theta, a_{-i}) = \text{marg}_\Theta \, f_i(t_i(\omega))(\theta) \cdot \text{marg}_{A_{-i}} \, f_i(t_i(\omega))(a_{-i}) \, \forall (\theta, a_{-i})\}$ and $\mathcal{H} := \bigcap_{i \in I} \mathcal{H}_i$.

Chapter 3

Monotone Threshold Representations[1]

3.1 Introduction

The classical model of rational choice is based on two fundamental postulates: When choosing from a menu A, an agent considers acceptable precisely those alternatives in A that are optimal according to some underlying preference ranking; this preference ranking is assumed (1) to be independent of the particular menu at hand and (2) to satisfy the axioms of a weak order.

These two postulates are jointly called into question by a growing literature (spanning psychology, marketing, and behavioral economics) on the phenomenon of "choice overload".[2] This literature has sought to corroborate the intuition that individuals have limited cognitive resources, which are put under greater strain

[2]For a detailed survey, including additional references, see Broniarczyk (2008).

by larger menus of alternatives. In line with this intuition, but contrary to (1), experimental studies suggest that agents' choice procedures vary with the menu, with consumers faced with larger menus resorting to greater use of simplifying choice heuristics (e.g., Payne (1976); Payne *et al.* (1993)) and achieving lower levels of choice accuracy relative to their "ideal" benchmark (e.g., Jacoby *et al.* (1974); Malhotra (1982)).[3] Contrary to (2), such choice heuristics typically do not involve the maximization of an underlying weak order.[4] [5]

This paper proposes a parsimonious extension of the classical model that accommodates these findings. As is well known, the classical model is equivalent to the Weak Axiom, and when the domain X of alternatives is finite, to the existence of a utility function $v: X \to \mathbb{R}$ such that the set $c(A)$ of acceptable alternatives in menu $A \subseteq X$ can be represented as

$$c(A) = \{x \in A : \max_{y \in A} v(y) - v(x) = 0\}. \tag{3.1}$$

We generalize the classical utility maximizing representation in equation (3.1) to a *monotone threshold representation* of the form

$$c(A) = \{x \in A : \max_{y \in A} v(y) - v(x) \leq \delta(A)\}, \tag{3.2}$$

[3] In these studies, subjects are asked to choose from sets of hypothetical alternatives (e.g., houses in Malhotra (1982)), each of which is described in terms of a range of attributes. The ideal benchmark is obtained by first eliciting consumers' most preferred levels for each attribute on which the study provides information.

[4]E.g., in Payne (1976), subjects facing menus of 6 or more alternatives frequently reported using choice procedures reminiscent of Simon's (1997) "satisficing" model and/or Tversky's (1972) "elimination-by-aspects" model.

[5]Another well documented manifestation of "choice overload" is that consumers are more likely to "walk away" from larger menus without making any choice (e.g., Iyengar and Lepper (2000)). This is not the focus of the present paper. However, if "walking away" is modeled as an outside option x^* that is available in every menu A and has value $v(x^*)$, then the monotone threshold model in equation (3.2) could accommodate agents who never choose x^* from some binary menu $\{x^*, y_1\}$ where $v(y_1) > v(x^*)$, but sometimes choose x^* from the larger menu $\{x^*, y_1, y_2, \ldots, y_n\}$.

where δ is a threshold function mapping menus of alternatives to nonnegative real numbers and we assume that δ is weakly increasing with respect to set inclusion.

The fully rational agent of the classical model ranks alternatives in any menu according to the weak order represented by v. The monotone threshold model captures a boundedly rational agent who departs from the maximization of v in a menu-dependent way. To model this departure in a parsimonious manner, the agent is assumed to maximize a menu-dependent *semiorder*, according to which y is preferred to x in A if and only if $v(y)$ exceeds $v(x)$ by more than $\delta(A)$. The semiorder is consistent with the underlying rational benchmark v in the sense that $v(y)$ is preferred to $v(x)$ only if $v(y) > v(x)$; but it is less discriminating, with the threshold $\delta(A)$ quantifying the menu-dependent extent of the departure from v.

Going back to Luce (1956), menu-*independent* semiorders have been used to model cognitively constrained agents who either deliberately resort to simplifying heuristics that ignore small differences in some decision-relevant criteria (e.g., price differences of a few cents) and/or are simply unable to discriminate between some alternatives (e.g., similar varieties of toothpaste). Moving beyond this, the threshold δ of the agent's menu-*dependent* semiorder in our model is monotonic with respect to set inclusion, making the departure from the rational benchmark more severe the larger the menu. Thus, Lucean "limited discrimination" is paired with an "overload effect", enabling us to capture the finding that larger menus exacerbate agents' cognitive limitations, decreasing choice accuracy and increasing their use of simplifying heuristics.

Relying on observable choice data alone, how can an external observer test whether an agent's behavior is consistent with the monotone threshold model? Remarkably, Section 3.2 shows that all observable implications of the monotone threshold model are fully encapsulated by the acyclicity of a simple relation S^c

67

derived from the agent's choice data c (Theorem 3.2.4); S^c encodes two intuitive ways in which c reveals one alternative to be superior to another according to *any* rational benchmark from which the agent could conceivably be departing. This characterization answers a question left open in Aleskerov *et al.* (2007)[6] and can equivalently be stated in terms of a relaxation of the Weak Axiom which we call Occasional Optimality. Our proof provides a fully constructive procedure for obtaining a monotone threshold representation $\langle v, \delta \rangle$ for given choice data c.

In Section 3.3, we relate the monotone threshold model to other threshold models in the literature and build on the characterization result of Section 3.2 to provide new foundations for these models. We show that adding the well-known Contraction axiom (Sen's α) to Occasional Optimality yields Luce's (1956) model of choice generated by a menu-independent semiorder, which differs from the monotone threshold model in that it is consistent with postulate (1) of the classical model and does not accommodate the "overload effect". On the other hand, adding the Strong Expansion axiom (Sen's β) to Occasional Optimality yields Tyson's (2008) "expansive satisficing" model; this is consistent with postulate (2) of the classical model, in that the agent's preference over the options in any given menu is assumed to be a *weak order* rather than a semiorder, and hence cannot accommodate agents who in the face of large menus resort to heuristics that ignore small differences between alternatives. Finally, the intersection of Luce's and Tyson's models is precisely the classical model, and all aforementioned models are special cases of Simon's theory of "satisficing" as axiomatized by Aleskerov *et al.* (2007).

---

[6]In Chapter 5, Aleskerov *et al.* (2007) introduce the concept of a monotone threshold representation, which they refer to as "utility maximization with an isotone threshold". However, while Aleskerov et al. provide necessary and sufficient conditions for a choice function to admit a *general* threshold representation, where the threshold map δ is not required to be nondecreasing with respect to set inclusion (we discuss this model in Section 3.3), their study of monotone threshold representations (pp. 190-193) only establishes a (straightforward) necessary condition for this more restrictive type of representation to exist.

Other Related Literature: In addition to the aforementioned papers, our model relates to Ortoleva's (2013) representation of an agent who dislikes large menus because of the greater "cost of thinking" involved in choosing from them. However, at a methodological level, Ortoleva's primitive is a preference over lotteries of menus, whereas our primitive is the agent's choice *from* menus, which is arguably more readily observable. At the conceptual level, Ortoleva's agent anticipates always choosing an *optimal* element from any menu, but might sometimes dislike larger menus because identifying their optimal elements entails a greater cost of thinking; our agent, on the other hand, does not always choose an optimal (according to the underlying rational benchmark) element from every menu, precisely because his cognitive limitations prevent him from doing so.

More broadly, our approach in this paper fits into an emerging literature in decision theory that seeks to characterize bounded rationality in terms of axioms on observable choice behavior. Some deviations from the fully rational paradigm that have been studied include status quo bias (Masatlioglu and Ok, 2005), framing effects (Salant and Rubinstein, 2008), sequential elimination of options (Mandler *et al.*, 2012; Manzini and Mariotti, 2007, 2012), limited attention (Masatlioglu *et al.*, 2012; Ellis, 2014), and sequential consideration of options (Caplin and Dean, 2011; Masatlioglu and Nakajima, 2013).

Among these, our paper relates most closely to Masatlioglu *et al.* (2012) and Manzini and Mariotti (2012). However, Masatlioglu *et al.* (2012) study an agent who maximizes a stable, menu-independent weak order in any given menu, but departs from the classical paradigm in that this maximization is carried out only over a limited subset of alternatives from the menu (his "consideration set"). This departure can be viewed as a reaction to choice overload which is in some sense the opposite of our model: Our agent always considers *all* items in any given menu,

but as a result of the taxing nature of this process perceives coarser preferences in larger menus. The agent in Manzini and Mariotti (2012) also employs the heuristic of ignoring small differences in some decision-relevant criteria, but their model prescribes the successive application of *multiple* semiorders (each representing a different decision-relevant dimension) until a *single* alternative from a given menu is left over; moreover, in contrast with our "overload effect", the semiorders that are applied are again assumed to be menu-independent.

3.2   Monotone Threshold Representations

Throughout this paper, let $X \neq \varnothing$ denote a finite set of *alternatives* and let $\mathcal{A} := \{A \subseteq X : A \neq \varnothing\}$ denote the set of *menus* (which are assumed to be non-empty). A and B always denote menus, and x, y, z alternatives. A *choice correspondence* on X is a map $c \colon \mathcal{A} \to \mathcal{A}$ such that $c(A) \subseteq A$ for all A; by definition, we only consider non-empty choice correspondences. We study the class of choice correspondences which admit a monotone threshold representation:

Definition 3.2.1. A choice correspondence c on X admits a *threshold representation* if there exist functions $v \colon X \to \mathbb{R}$ and $\delta \colon \mathcal{A} \to \mathbb{R}_+$ such that for every A,

$$c(A) = \{x \in A : \max_{y \in A} v(y) - v(x) \leq \delta(A)\}.$$

We call v the *fully rational benchmark* and δ the *departure threshold* of the representation. The threshold representation $\langle v, \delta \rangle$ is called a *monotone threshold representation* (MTR) if δ is nondecreasing with respect to set inclusion, i.e. $\delta(A) \leq \delta(B)$ whenever $A \subseteq B$.

As motivated in the introduction, the monotone threshold model captures two anomalies: The agent has "limited discrimination" between the alternatives in any menu A, represented by the fact that his choices from A do not maximize

v, but rather the *semiorder*[7] according to which y is preferred to x if and only if $v(y) - v(x) > \delta(A)$. Moreover, there is an "overload effect", captured by the assumption that the extent $\delta(A)$ of the departure in A from the rational benchmark v is more severe the larger A. We interpret $c(A)$ as the set of alternatives that the agent considers acceptable after being presented with and contemplating the entire menu A; these are the alternatives which we might observe him select on different occasions. To keep the departure from the classical model as parsimonious as possible, we are not concerned with predicting the relative *frequency* with which any particular alternative is chosen.[8] For the same reason, we do not seek to model other menu-dependent departures from fully rational choice, such as the "attraction effect",[9] which we view as orthogonal to the "choice overload" phenomenon.

3.2.1 Characterization

Our main result, Theorem 3.2.4, identifies testable conditions on an agent's choice behavior c that are equivalent to c admitting a monotone threshold representation. We first define the following revealed preference relations:

Definition 3.2.2. Given a choice correspondence c on X, the induced relations P^c,

---

[7] A relation K on X is a *semiorder* if it is *irreflexive* ($\forall x \in X \, \neg xKx$), *semitransitive* (for all w, x, y, $z \in X$ such that wKx and xKy, we have wKz or zKy), and satisfies the *interval order condition* (for all w, x, y, $z \in X$ such that wKx and yKz, we have wKz or yKx).

[8] By contrast, taking as primitive an agent's stochastic choice rule, Fudenberg *et al.* (2014) study so-called "Additive Perturbed Utility" representations. As a special case, they consider an agent who has "limited discrimination", in the sense that his choices are "more uniform" the larger the menu he faces. In contrast with the monotone threshold model, their representation assumes that for any menu A, *all* alternatives from A are chosen with strictly positive probability.

[9] Cf. Huber *et al.* (1982) for the original formulation of this effect and Ok *et al.* (2015) for an axiomatic model.

Q^c, R^c and S^c on X are defined as follows.[10] For all $x, y \in X$:

(i). $xP^c y$ if and only if $x \neq y$ and $c(\{x, y\}) = \{x\}$;

(ii). $xR^c y$ if and only if there exists A such that $x \in c(A)$ and $y \in A \smallsetminus c(A)$;

(iii). $xQ^c y$ if and only if there exists A such that $y \in A$ and $c(A) \nsubseteq c(A \cup \{x\})$;

(iv). $xS^c y$ if and only if $xR^c y$ or $xQ^c y$.

If c encodes a *rational* agent's choice behavior, then c is generated by a weak order W. In this case, it is easy to see that $W = P^c = R^c = Q^c$, so that W can be uniquely identified by observing the agent's choices from binary menus. Suppose, on the other hand, that c represents the choice behavior of a boundedly rational agent, in the sense captured by the monotone threshold model. Then there is an underlying rational benchmark, represented by v, but the agent's choice behavior departs from v-maximization in a menu-dependent manner captured by a threshold map δ. Consequently, observing c will not, in general, fully reveal the rational benchmark. However, as a result of the agent's "limited discrimination" and the "overload effect", R^c and Q^c each reveal useful information about v, and in general this exceeds the information that is revealed by P^c alone:

Due to the agent's "limited discrimination", his choices from any given menu A are generated by the semiorder L_A given by $xL_A y$ if and only if $v(x) - v(y) > \delta(A)$. This has the implication that even if $x \neq y$ and the agent does not perceive a preference for x over y in the direct *pairwise* comparison (i.e., $\neg xP^c y$), the agent's preferences over *other* alternatives might nevertheless reveal that v must rank x higher than y. Two instances of this are worth noting:

---

[10]Relations P^c and R^c are familiar to the literature, cf. Tyson (2008) and Aleskerov *et al.* (2007). The definitions of relations Q^c and S^c appear to be new.

First, there might be a menu $A \supseteq \{x, y\}$ and some $z \in A$ such that in A, z is perceived preferred to y but not to x (zL_Ay but $\neg zL_Ax$). This will be the case if and only if $x \in c(A)$ and $y \in A \smallsetminus c(A)$, i.e. xR^cy.

Second, there might be a menu $B \supseteq \{x, y\}$ and some $w \in B$ such that in B, x is perceived preferred to w but y is not (xL_Bw but $\neg yL_Bw$). Note that in this case y might still be chosen from B and from any other menu containing x and y, so we need not have xR^cy. However, because of the "overload effect" one observable manifestation of this situation is the following: Suppose that $w \in c(B \smallsetminus \{x\})$, but $w \notin c(B)$, so that xQ^cy. Then in $B \smallsetminus \{x\}$, *no* alternative is perceived preferred to w (in particular, $\neg yL_{B \smallsetminus \{x\}}w$). But the "overload effect" makes the agent *less* discriminating in B than in $B \smallsetminus \{x\}$. Thus, in menu B the agent also does not prefer any alternative from $B \smallsetminus \{x\}$ to w. Since $w \notin c(B)$, the only possible conclusion is that xL_Bw but $\neg yL_Bw$.

In both of the above cases we must have $v(x) > v(y)$. So if c admits a monotone threshold representation, then *any* rational benchmark v from which the agent might conceivably be departing must extend R^c and Q^c. Hence, even if the choice data does not allow us to fully identify the rational benchmark v, we can at the very least conclude that the relation S^c, which subsumes both R^c and Q^c, must be *acyclic*.[11] Remarkably, we will see in Theorem 3.2.4 that acyclicity of S^c is not only necessary for c to admit a monotone threshold representation, but in fact completely encapsulates *all* the observable implications of the model. To fully bring out the connection with the classical model, we note that acyclicity of S^c can also be stated in terms of a relaxation of the Weak Axiom. Consider the following two equivalent formulations of the Weak Axiom:

---

[11]A relation K on X is called *acyclic* if for all $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n \in X$ such that x_iKx_{i+1} for all $i = 1, \ldots, n - 1$, we have $\neg x_nKx_1$.

(Equivalent Formulations of the Weak Axiom).

- **Formulation 1:** For all A, for all $x \in c(A)$, and for any B including x: If $c(B) \cap A \neq \varnothing$, then $x \in c(B)$.

- **Formulation 2:** For all A, for all $x \in c(A)$, and for any B including x: $c(B) \subseteq c(B \cup \{y\})$ for all $y \in A$.

In both formulations, a particular kind of optimality requirement is imposed on *all* elements of $c(A)$. In the language of Definition 3.2.2, Formulation 1 requires all elements in $c(A)$ to be R^c-maximal. It is easy to see that this is equivalent to Formulation 2, which imposes Q^c-maximality on all elements in $c(A)$.

On the other hand, a relation K on a finite set X is acyclic if and only if every non-empty subset $A \subseteq X$ has a K-maximal element.[12] Therefore, acyclicity of $S^c = R^c \cup Q^c$ is equivalent to every menu A containing *at least one* distinguished element x_A which is *both R^c-maximal and Q^c-maximal*. This is the content of the following relaxation of the Weak Axiom, where the changes are highlighted in boldface:

Condition 3.2.3 (Occasional Optimality). For all A, **there exists** $x_A \in A$ such that for any B including x_A:

(i) If $c(B) \cap A \neq \varnothing$, then $x_A \in c(B)$. **And**

(ii) $c(B) \subseteq c(B \cup \{y\})$ for all $y \in A$.

If Occasional Optimality holds, then applying (i) with $B = A$ it is clear that $x_A \in c(A)$. Hence, while the Weak Axiom requires that *any* alternative we might observe the agent choose from A will be S^c-optimal, Occasional Optimality requires

---

[12]Cf., for example, Kreps (1988), Propositions 2.7 and 2.8 (p. 12-13).

only that *at least some* of the agent's potential choices from *A* are optimal, thus justifying the name of the axiom.[13]

We are now ready to state the representation theorem:

Theorem 3.2.4. *Suppose c is a choice correspondence on X. The following are equivalent:*

(i). *c admits a monotone threshold representation;*

(ii). *the relation S^c is acyclic;*

(iii). *c satisfies Occasional Optimality.*

The proof is in Appendix C.1. The argument that (ii) implies (i) is fully constructive: If S^c is acyclic, then we can construct a weak order W^c on *X* that extends S^c. The utility *v* is then constructed in such a way as to represent W^c while also satisfying the following "increasing differences" property:

$$\text{If } v(x) > v(y) \text{ and } v(x) > v(w), \text{ then } v(x) - v(w) > v(y) - v(z) \text{ for all } z. \quad (3.3)$$

This property allows us to inductively define the threshold δ as follows:

- set $\delta(\{x\}) := 0$ for all $x \in X$;

- if $|A| \geq 2$ and $c(A) = \text{argmax}_A v$, set $\delta(A) := \max_{B \subsetneq A} \delta(B)$;

- if $|A| \geq 2$ and $c(A) \supsetneq \text{argmax}_A v$, set $\delta(A) := \max_A v - \min_{c(A)} v$.

By an inductive argument involving several cases, we verify that $\langle v, \delta \rangle$ thus constructed is indeed an MTR of *c*.

---

[13]The move from the universal imposition of a certain property in the Weak Axiom to the requirement that it hold only on a distinguished subset of elements in the Occasional Optimality axiom is in the same spirit as other relaxations of the Weak Axiom in the recent literature on boundedly rational choice, for example the Reducibility axiom in Manzini and Mariotti (2012) or the WARP with Limited Attention axiom in Masatlioglu *et al.* (2012). Note that despite the existential formulation of these axioms, they are (at least in principle) testable by observation, because the global domain *X* of alternatives is assumed to be finite.

3.3 Related Threshold Models

In this section, we contrast the monotone threshold model with related models in the literature.

Taking their cue from Simon's theory of "satisficing", according to which agents choose "alternative[s] that meet or exceed specified criteria," but that are "not guaranteed to be either unique or in any sense the best,"[14] Aleskerov *et al.* (2007) study choice correspondences c which admit a *satisficing representation* (SR): There exist functions $u\colon X \to \mathbb{R}$ and $\theta\colon \mathcal{A} \to \mathbb{R}$ such that for every A, we have that $c(A) = \{x \in A : u(x) \geq \theta(A)\}$. Setting $u = v$ and $\delta(A) = \max_{y \in A} v(y) - \theta(A)$, it is clear that any satisficing representation can be converted into a *general* threshold representation and vice versa; but Example C.2.1 in Appendix C.2 exhibits a choice correspondence with a satisficing representation that does *not* admit a *monotone* threshold representation. Thus, the satisficing model incorporates the intuition of an agent who departs from maximization of a rational benchmark v in a menu-dependent manner, but unlike the monotone threshold model, it is too general to capture the fact that this departure is due to "choice overload" caused by larger menus. In the previous section, our argument that Q^c reveals the agent's preference relied on the "overload effect", but the argument for R^c did not. Hence, it comes as no surprise that the satisficing model is fully characterized by acyclicity of R^c, as is shown in Aleskerov et al.[15]

On the other hand, the monotone threshold model is strictly more general than the following two special cases of the satisficing model: Luce's (1956) model of choice

---

[14]Simon (1997) p. 125.

[15]Cf. Aleskerov *et al.* (2007), Corollary 5.4 (p. 167). Aleskerov et al. refer to acyclicity of R^c as the Strong Axiom of Revealed Strict Preference and to satisficing representations as "over-value" choice rules.

generated by a menu-*independent* semiorder[16] is equivalent to a *constant threshold representation* (CTR) with menu-independent threshold $\delta \in \mathbb{R}_+$. Tyson (2008) studies choice correspondences c which admit an *expansive satisficing representation* (ESR): There exists a satisficing representation $\langle u, \theta \rangle$ of c with the property that whenever $A \subseteq B$ and $\max_{y \in A} u(y) \geq \theta(B)$, then $\theta(A) \geq \theta(B)$.

Like the monotone threshold model, the constant threshold model satisfies "limited discrimination", violating postulate (2) of the classical model (as formulated in the introduction). But since the departure threshold is independent of menu size, it is consistent with postulate (1) and does not capture an additional "overload effect". By contrast, the expansive satisficing model can accommodate an "overload effect", whereby larger menus make agents less discriminating, but unlike the monotone threshold model, it is consistent with postulate (2) of the classical model, because Tyson's agent can be seen as "locally rational": Tyson[17] shows that the ESR model is equivalent to the agent maximizing a menu-dependent preference relation P_A in each A, where these preference relations are coarser than an underlying weak order W[18], and more so the larger the menu,[19] but additionally, *each menu-dependent preference P_A is itself a weak order*.

To axiomatically elucidate the gap between the monotone threshold and CTR and ESR models, the following lemma builds on Theorem 3.2.4 to obtain novel foundations for the latter two representations. Recall the following two well-known conditions, which are jointly equivalent to the Weak Axiom (cf. Sen and Bordes):[20]

---

[16]The model was later axiomatized by Jamison and Lau (1973) and Fishburn (1975).

[17]Cf. Tyson (2008) Theorem 3 (p. 58) and Theorem 5B (p. 59).

[18]In the sense that $\bigcup_{A \in \mathcal{A}} P_A \subseteq W$.

[19] In the sense that if $x, y \in A \subseteq B$ and $x P_B y$, then $x P_A y$ (i.e. the relations are "nested").

[20]Cited in Tyson (2008), p. 56.

Condition 3.3.1 (Contraction/Sen's α). For all A, B such that $A \subseteq B$, we have $c(B) \cap A \subseteq c(A)$.

Condition 3.3.2 (Strong Expansion/Sen's β). For all A, B such that $A \subseteq B$ and $c(B) \cap A \neq \varnothing$, we have $c(A) \subseteq c(B)$.

Lemma 3.3.3. *Suppose c is a choice correspondence on X. Then:*

(i). *c admits a CTR if and only if c satisfies Occasional Optimality and Contraction.*

(ii). *c admits an ESR if and only if c satisfies Occasional Optimality and Strong Expansion.*

Proof. See Appendix C.1. ∎

Since Contraction and Strong Expansion are jointly equivalent to the Weak Axiom, the class of choice correspondences admitting *both* a CTR and an ESR is precisely the class of choice correspondences with a classical utility maximizing representation. On the other hand, Example C.2.2 (respectively C.2.3) in Appendix C.2 exhibits a choice correspondence which admits a CTR but not an ESR (respectively, an ESR but not a CTR), showing that Strong Expansion ("local rationality") and Contraction ("menu-independence") are independent in the presence of Occasional Optimality. Finally, Example C.2.4 exhibits a choice correspondence with an MTR that admits *neither* a CTR nor an ESR, showing that the monotone threshold model can simultaneously accommodate failures of postulates (1) and (2) of the classical model. Figure 3.1 summarizes the relationships between the various threshold models.

We conclude with a brief discussion of the example showing that the monotone threshold model is strictly more general than the ESR model. The violation of Strong Expansion in this example takes the following form: For some menu A and alternatives $x, y \in A$ and $z \notin A$ with $v(z) > v(y) > v(x)$, we have that $x, y \in c(A)$,
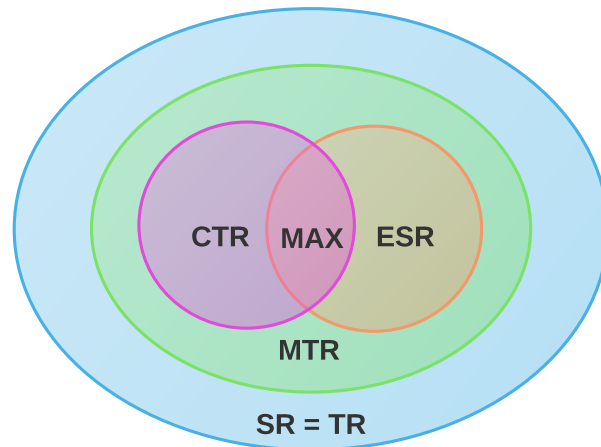
Figure 3.1: *Relationship between various threshold models (SR, TR, MTR, ESR, CTR and MAX denote the class of choice functions admitting a satisficing, threshold, monotone threshold, expansive satisficing, constant threshold and utility maximizing representation, respectively).*

$y, z \in c(A \cup \{z\})$, and $x \notin c(A \cup \{z\})$. Concretely, such a situation might arise if a consumer employs a "rule of thumb" of ignoring price differences of less than 10 cents when faced with menus of A's size or larger, and if x, y, z are (otherwise equally attractive) candy bars priced at \$.99, \$.95, and \$.87, respectively. The example relies crucially on the consumer's "limited discrimination": Adding z to A *indirectly* helps him choose between x and y, even though his heuristic does not *directly* discriminate between the two—a situation that cannot arise if a consumer is "locally rational" in the sense of Tyson's model.[21]

This example appears consistent with findings from consumer psychology. As discussed in the introduction, a consumer's choice accuracy from menus that exceed

---

[21]More precisely, in Tyson's model, if $\neg z P_{A \cup z} y$ and $z P_{A \cup z} x$, then because $P_{A \cup \{z\}}$ is a weak order, we must in fact have that $y P_{A \cup \{z\}} x$ (i.e., the consumer *directly* perceives a preference for y over x in $A \cup \{z\}$). But by nestedness (see footnote 19), this implies $y P_A x$, contradicting $x \in c(A)$.

a certain size is in general far from perfect,[22] but various studies also suggest that the *addition* of alternatives to a menu can have an ambiguous effect on choice accuracy— summarizing these findings, Broniarczyk (2008) writes that "the addition of product alternatives to a choice set initially increases a consumer's choice accuracy, but the continued addition of product options results in a decrease in a consumer's choice accuracy."[23] While the measures of choice accuracy employed by this literature vary, one indicator is a consumer's *worst-possible* choice from a menu. Contrary to these findings, the ESR model implies that either $c(A \cup \{z\}) = \{z\}$ (so that choice accuracy is perfect) or $c(A) \subseteq c(A \cup \{z\})$ (so that the worst-possible choice from $c(A \cup \{z\})$ is at least as bad as from $c(A)$). By contrast, the monotone threshold model allows that $c(A) \nsubseteq c(A \cup \{z\}) \neq \{z\}$, and hence can accommodate improved, but still imperfect, choice accuracy as a result of adding z to A.

---

[22]E.g., in Malhotra (1982), consumers' average probability of choosing the alternative from a menu A that is closest (in terms of Euclidean distance) to their ideal benchmark is 0.34 if $|A| = 15$ and 0.23 if $|A| = 25$. Recall footnote 3 for an explanation of the ideal benchmark.

[23]Broniarczyk (2008) p. 12.

References

ALESKEROV, F., BOUYSSOU, D. and MONJARDET, B. (2007). *Utility Maximization, Choice and Preference*. Berlin: Springer, 2nd edn.

BANDIERA, O. and RASUL, I. (2006). Social networks and technology adoption in Northern Mozambique. *Economic Journal*, **116** (514), 869–902.

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

BAPTISTA, R. (1999). The diffusion of process innovations: A selective review. *International Journal of the Economics of Business*, **6** (1), 107–129.

BASS, F. M. (1969). A new product growth model for consumer durables. *Management Science*, **15** (5), 215–227.

— (1980). The relationship between diffusion rates, experience curves, and demand elasticities for consumer durable technological innovations. *Journal of Business*, **53** (3), S51–S67.

BATTIGALLI, P. and SINISCALCHI, M. (2003). Rationalization and incomplete information. *Advances in Theoretical Economics*, **3** (1), 57.

BERGEMANN, D. and HEGE, U. (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance*, **22** (6), 703–735.

— and — (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics*, pp. 719–752.

— and VÄLIMÄKI, J. (1997). Market diffusion with two-sided learning. *RAND Journal of Economics*, pp. 773–795.

BESLEY, T. and CASE, A. (1993). Modeling technology adoption in developing countries. *American Economic Review*, **83** (2), 396–402.

— and — (1994). Diffusion as a learning process: Evidence from HYV cotton, Woodrow Wilson School: Discussion Paper #174.

BIKHCHANDANI, S., HIRSHLEIFER, D. and WELCH, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, **100** (5), 992–1026.

BOARD, S. and MEYER-TER VEHN, M. (2013). Reputation for quality. *Econometrica*, **81** (6), 2381–2462.

BOLTON, P. and HARRIS, C. (1999). Strategic experimentation. *Econometrica*, **67** (2), 349–374.

BONATTI, A. and HÖRNER, J. (2011). Collaborating. *American Economic Review*, **101** (2), 632–663.

BÖRGERS, T. (1994). Weak dominance and approximate common knowledge. *Journal of Economic Theory*, **64**, 265–265.

BRANDENBURGER, A. and DEKEL, E. (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, **59** (1), 189–198.

—, FRIEDENBERG, A. and KEISLER, H. J. (2008). Admissibility in games. *Econometrica*, **76** (2), 307–352.

BRONIARCZYK, S. (2008). Product assortment. In C. P. Haugtvedt, P. Herr and F. Kardes (eds.), *Handbook of Consumer Psychology*, New York: Taylor & and Francis, pp. 755–79.

BUCHWALD, H. and OIEN, D. M. (2009). Metabolic/bariatric surgery worldwide 2008. *Obesity Surgery*, **19** (12), 1605–1611.

— and — (2013). Metabolic/bariatric surgery worldwide 2011. *Obesity Surgery*, **23** (4), 427–436.

CABRAL, L. (2012). Lock in and switch: Asymmetric information and new product diffusion. *Quantitative Marketing and Economics*, **10** (3), 375–392.

CAPLIN, A. and DEAN, M. (2011). Search, choice, and revealed preference. *Theoretical Economics*, **8**, 19–48.

CHAMLEY, C. and GALE, D. (1994). Information revelation and strategic delay in a model of investment. *Econometrica*, **62** (5), 1065–1085.

CHE, Y.-K. and HÖRNER, J. (2014). Optimal design for social learning, mimeo.

CHEVALIER, J. A. and MAYZLIN, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, **43** (3), 345–354.

CONLEY, T. G. and UDRY, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review*, **100** (1), 35–69.

CONSUMERS' UNION (1973). Microwave ovens: Not recommended. *Consumer Reports*, **38** (4), 221–228.

DAVID, P. A. (1969). *A Contribution to the Theory of Diffusion*. Stanford University: Research Center in Economic Growth.

DAVIES, S. (1979). *The Diffusion of Process Innovations*. Cambridge, UK: CUP Archive.

DEKEL, E. and FUDENBERG, D. (1990). Rational behavior with payoff uncertainty. *Journal of Economic Theory*, **52** (2), 243–267.

—, — and MORRIS, S. (2007). Interim correlated rationalizability. *Theoretical Economics*, **2** (1), 15–40.

— and SINISCALCHI, M. (2014). Epistemic game theory. In P. Young and S. Zamir (eds.), *Handbook of Game Theory, Volume 4*, Amsterdam: North Holland.

ELLIS, A. (2014). Foundations for optimal inattention, mimeo.

ELLISON, G. and FUDENBERG, D. (1993). Rules of thumb for social learning. *Journal of*

Political Economy, **101** (4), 612–643.

Farrell, J. and Saloner, G. (1985). Standardization, compatibility, and innovation. *RAND Journal of Economics*, **16** (1), 70–83.

— and — (1986). Installed base and compatibility: Innovation, product preannouncements, and predation. *American Economic Review*, **76** (5), 940–955.

Fishburn, P. (1975). Semiorders and choice functions. *Econometrica*, **43**, 975–977.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, **103** (6), 1176–1209.

Frick, M. and Ishii, Y. (2014). Innovation adoption by forward-looking social learners, February 2014 version (available at http://scholar.harvard.edu/mfrick).

Fudenberg, D., Iijima, R. and Strzalecki, T. (2014). Stochastic choice and revealed perturbed utility, mimeo.

—, Kreps, D. and Levine, D. (1988). On the robustness of equilibrium refinements. *Journal of Economic Theory*, **44** (2), 354–380.

Geroski, P. A. (2000). Models of technology diffusion. *Research Policy*, **29** (4), 603–625.

Gort, M. and Klepper, S. (1982). Time paths in the diffusion of product innovations. *Economic Journal*, **92** (367), 630–653.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica*, **25** (4), 501–522.

Guenthner, J., Lin, B.-H. and Levi, A. E. (1991). The influence of microwave ovens on the demand for fresh and frozen potatoes. *Journal of Food Distribution Research*, **22** (3), 45–52.

Halac, M., Liu, Q. and Kartik, N. (2013). Optimal contracts for experimentation, mimeo.

—, — and — (2014). Contests for experimentation, mimeo.

— and PRAT, A. (2014). Managerial attention and worker engagement, mimeo.

HÖRNER, J. and SAMUELSON, L. (2013). Incentives for experimenting agents. *RAND Journal of Economics*, **44** (4), 632–663.

HOYER, W. D., MacINNIS, D. J. and PIETERS, R. (2012). *Consumer Behavior*. Mason, OH: South-Western, 6th edn.

HU, T.-W. (2007). On p-rationalizability and approximate common certainty of rationality. *Journal of Economic Theory*, **136** (1), 379–391.

HUBER, J., PAYNE, J. and PUTO, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, **9**, 90–98.

IYENGAR, S. and LEPPER, M. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, **79**, 995–1006.

JACOBY, J., SPELLER, D. and KOHN, C. (1974). Brand choice behavior as a function of information load. *Journal of Marketing Research*, **11** (1), 63–69.

JAMISON, D. T. and LAU, L. J. (1973). Semiorders and the theory of choice. *Econometrica*, **41**, 901–912.

JENSEN, R. (1982). Adoption and diffusion of an innovation of uncertain profitability. *Journal of Economic Theory*, **27** (1), 182–193.

JOVANOVIC, B. and LACH, S. (1989). Entry, exit, and diffusion with learning by doing. *American Economic Review*, **79** (4), 690–699.

KAJII, A. and MORRIS, S. (1997). The robustness of equilibria to incomplete information. *Econometrica*, **65** (6), 1283–1309.

KAPUR, S. (1995). Technological diffusion with social learning. *Journal of Industrial Economics*, **43** (2), 173–195.

KARSHENAS, M. and STONEMAN, P. L. (1993). Rank, stock, order, and epidemic effects

in the diffusion of new process technologies: An empirical model. *RAND Journal of Economics*, **24** (4), 503–528.

KEILLOR, B. D. (2007). *Marketing in the 21st Century*, vol. 4. Westport, CT: Greenwood Publishing Group.

KELLER, G. and RADY, S. (2010). Strategic experimentation with Poisson bandits. *Theoretical Economics*, **5** (2), 275–311.

— and — (2014). Breakdowns. *Theoretical Economics*, forthcoming.

—, — and CRIPPS, M. (2005). Strategic experimentation with exponential bandits. *Econometrica*, **73** (1), 39–68.

KLEIN, N. and RADY, S. (2011). Negatively correlated bandits. *Review of Economic Studies*, **78**, 693–732.

KOHLBERG, E. and MERTENS, J.-F. (1986). On the strategic stability of equilibria. *Econometrica*, **55** (5), 1003–1037.

KREPS, D. (1988). *Notes on the Theory of Choice*. Boulder: Westview Press.

LUCE, R. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, **24**, 178–191.

LUCE, R. D. and RAIFFA, H. (1957). *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.

MACLEOD, W. B. (2007). Reputations, relationships, and contract enforcement. *Journal of Economic Literature*, **45** (3), 595–628.

MAHAJAN, V., MULLER, E. and BASS, F. M. (1990). New product diffusion models in marketing: A review and directions for research. *Journal of Marketing*, **54**, 1–26.

— and PETERSON, R. A. (1985). *Models for Innovation Diffusion*, vol. 48. Newbury Park, CA: Sage Publications.

MALHOTRA, N. (1982). Information load and consumer decision making. *Journal of Consumer Research*, **8** (4), 419–430.

MANDLER, M., MANZINI, P. and MARIOTTI, M. (2012). A million answers to twenty questions: Choosing by checklist. *Journal of Economic Theory*, **147**, 71–92.

MANSFIELD, E. (1961). Technical change and the rate of imitation. *Econometrica*, **29** (4), 741–766.

— (1968). *Industrial Research and Technological Innovation: An Econometric Analysis*. New York: Norton.

MANZINI, P. and MARIOTTI, M. (2007). Sequentially rationalizable choice. *American Economic Review*, **97**, 1824–1839.

— and — (2012). Choice by lexicographic semiorders. *Theoretical Economics*, **7**, 1–23.

MASATLIOGLU, Y. and NAKAJIMA, D. (2013). Choice by iterative search. *Theoretical Economics*, **8**, 701–728.

—, — and OZBAY, E. (2012). Revealed attention. *American Economic Review*, **102**, 2183–2205.

— and OK, E. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, **121**, 1–29.

MUNSHI, K. (2004). Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution. *Journal of Development Economics*, **73** (1), 185–213.

MURTO, P. and VÄLIMÄKI, J. (2011). Learning and information aggregation in an exit game. *Review of Economic Studies*, **78** (4), 1426–1461.

OK, E., ORTOLEVA, P. and RIELLA, G. (2015). Revealed (p)reference theory. *American Economic Review*, **105**, 299–321.

ORTOLEVA, P. (2013). The price of flexibility: Towards a theory of thinking aversion. *Journal of Economic Theory*, **148**, 903–934.

PAYNE, J. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human*

Performance, **16**, 366–387.

—, Bettman, J. and Johnson, E. (1993). *The Adaptive Decision Maker*. New York: Cambridge University Press.

Pearce, D. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, **52** (4), 1029–1050.

Rosenberg, D., Solan, E. and Vieille, N. (2007). Social learning in one-arm bandit problems. *Econometrica*, **75** (6), 1591–1611.

Salant, Y. and Rubinstein, A. (2008). (a, f): Choice with frames. *Review of Economic Studies*, **75**, 1287–1296.

Schuhmacher, F. (1999). Proper rationalizability and backward induction. *International Journal of Game Theory*, **28** (4), 599–615.

Simon, H. (1997). *Models of Bounded Rationality*, vol. 3. Cambridge, MA: MIT Press.

Smith, L. and Sørensen, P. (2000). Pathological outcomes of observational learning. *Econometrica*, **68** (2), 371–398.

Strulovici, B. (2010). Learning while voting: Determinants of collective experimentation. *Econometrica*, **78** (3), 933–971.

Szpilrajn, E. (1930). Sur l'extension de l'ordre partiel. *Fundamenta Mathematica*, **16**, 386–389.

Tan, T. C.-C. and da Costa Werlang, S. R. (1988). The Bayesian foundations of solution concepts of games. *Journal of Economic Theory*, **45** (2), 370–391.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, **79** (4), 281–299.

Tyson, C. (2008). Cognitive constraints, contraction consistency, and the satisficing criterion. *Journal of Economic Theory*, **138**, 51–70.

Weinstein, J. and Yildiz, M. (2007). A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica*, **75** (2), 365–400.

Wiersema, F. D. and Buzzell, R. D. (1979). Note on the microwave oven industry, Harvard Business School Case Study.

Williams, A. (2014). Surveys of microwave ovens in US homes, Lawrence Berkeley National Laboratory Paper #5947E.

Young, H. P. (2009). Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American Economic Review*, **99** (5), 1899–1924.

Zhao, Y. and Encinosa, W. (2007). Bariatric surgery utilization and outcomes in 1998 and 2004, Agency for Healthcare Research and Quality: Statistical Brief #23.

Appendix A

Appendix to Chapter 1

A.1 Proof of Theorem 1.3.1

This appendix establishes the quasi-single crossing property for equilibrium incentives (Theorem 1.3.1). All remaining proofs are in Appendix A.2. We will make use of the following five lemmas which are proved in Appendix A.2.1. For an equilibrium adoption flow $N_{t \geq 0}$, denote the associated value to waiting by $W^N_{t \geq 0}$ and the no-news posterior by $p^N_{t \geq 0}$.

Lemma A.1.1. *If $N_{t \geq 0}$ is an an equilibrium, then W^N_t is continuous in t.*

Lemma A.1.2. *Suppose that $N_{t \geq 0}$ is an equilibrium and that $W^N_t < 2p^N_t - 1$ for some $t > 0$. Then there exists some $v > 0$ such that W^N_t is continuously differentiable in t on the interval $(t - v, t + v)$ and*

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

Lemma A.1.3. *Suppose that $N_{t \geq 0}$ is an equilibrium and that $W^N_t > 2p^N_t - 1$ for some $t > 0$. Then there exists some $v > 0$ such that W^N_t is continuously differentiable in t on the*

interval $(t - v, t + v)$ and

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

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

Lemma A.1.4. *Let $N_{t \geq 0}$ be an equilibrium under PBN. Suppose that $\varepsilon > 0$ or $p_0 > \frac{1}{2}$. Then $\lim_{t \to \infty} p_t^N = \mu(\varepsilon, \Lambda_0, p_0)$ and $\lim_{t \to \infty} W_t^N = \frac{\rho}{\rho + r}(2\mu(\varepsilon, \Lambda_0, p_0) - 1)$, where*

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

Lemma A.1.5. *Suppose that learning is via PBN. Suppose that $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$. Then the unique equilibrium satisfies $N_t = 0$ for all t.*

Henceforth we drop the superscript N from W and p.

Proof of Theorem 1.3.1 under Perfect Good News:

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

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

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

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

By continuity, $s^* < s'$, $W_{s^*} = 2p_{s^*} - 1$, and $W_s < 2p_s - 1$ for all $s \in (s^*, s')$. Then by Lemma A.1.2, the right hand derivative of $W_s - (2p_s - 1)$ at s^* exists and satisfies:

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

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

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

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

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

By continuity, $s^* > t$, $W_{s^*} = 2p_{s^*} - 1$, and $W_s > 2p_s - 1$ for all $s \in (t, s^*)$. Note that $p_{s^*} \geq \frac{1}{2}$, because W_{s^*} is bounded below by 0. Moreover, by Lemma A.1.3 the left-hand derivative of $W_s - (2p_s - 1)$ at s^* exists and is given by:

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

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

Proof of Theorem 1.3.1 under Perfect Bad News:

Let $\varepsilon = \varepsilon_B \geq 0 = \varepsilon_G$ and $\lambda = \lambda_B > 0 = \lambda_G$. If $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$, then by Lemma A.1.5 $N_t = 0$ for all t, so the proof of Theorem 1.3.1 is obvious. We now prove the theorem under the assumption that either $\varepsilon > 0$ or $p_0 > \frac{1}{2}$.

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

Suppose that $W_t = 2p_t - 1$ and suppose for a contradiction that $W_{s'} > 2p_{s'} - 1$ for some $s' > t$. Let $\bar{s} := \inf\{s > s' : W_t \leq 2p_s - 1\} < \infty$, since by Lemma A.1.4 $\lim_{t \to \infty} 2p_t - 1 > \lim_{t \to \infty} W_t$. Let $\underline{s} := \sup\{s < s' : W_s \leq 2p_s - 1\}$. Then $\underline{s} < \bar{s}$, $W_{\underline{s}} = 2p_{\underline{s}} - 1$, $W_{\bar{s}} = 2p_{\bar{s}} - 1$, and $W_s > 2p_s - 1$ for all $s \in (\underline{s}, \bar{s})$. Lemma A.1.3

together with the fact that $N_s = 0$ for all $s \in (\underline{s}, \bar{s})$ implies the following two limits:

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

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

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

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

$$\text{and}$$

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

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

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

Suppose that $W_t < 2p_t - 1$, let $\underline{s} := \inf\{s' > t : W_{s'} \geq 2p_{s'} - 1\}$, and suppose for a contradiction that $\underline{s} < \infty$. By continuity, $W_\tau < 2p_\tau - 1$ for all $\tau \in [t, \underline{s})$ and $W_{\underline{s}} = 2p_{\underline{s}} - 1$. Furthermore, by Lemma A.1.4, there exists some $\bar{s} \geq \underline{s}$ such that $2p_{\bar{s}} - 1 = W_{\bar{s}}$ and $2p_s - 1 > W_s$ for all $s > \bar{s}$. Lemma A.1.2 implies the following two limits:

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

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

As usual, because $W_s < 2p_s - 1$ for all $s \in (t, \underline{s})$ and for all $s > \bar{s}$, we must have

93

$H_{\underline{s}} \geq 0$ and $H_{\bar{s}} \leq 0$. But since $p_{\bar{s}} \geq p_{\underline{s}}$, this is only possible if $\underline{s} = \bar{s} =: s^*$ and $H_{s^*} = H_{\underline{s}} = H_{\bar{s}} = 0$.

Thus,

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

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

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

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

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

Rearranging we obtain:

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

By Lemma A.1.2, the right-hand side is precisely the derivative $\frac{d}{ds}(2p_s - 1) - \dot{W}_s$. But then for all $s \in [t, s^*)$, $2p_s - 1 > W_s$ and $2p_s - 1 - W_s$ is strictly increasing, contradicting continuity and the fact that $2p_{s^*} - 1 = W_{s^*}$. ∎

A.2 Remaining Proofs

A.2.1 Proofs of Lemmas A.1.1–A.1.5

Proof of Lemma A.1.1: Note the following recursive formulations for W_t^N. If learning is via perfect bad news, then

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

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

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

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

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

Proof of Lemma A.1.2: Suppose that $W_t^N < 2p_t^N - 1$ for some $t > 0$. By Lemma A.1.1 W_t^N is continuous in t, and so is $2p_t^N - 1$. Hence there exists $\nu > 0$ such that $W_\tau^N < 2p_\tau^N - 1$ for all $\tau \in (t - \nu, t + \nu)$. Because N is an equilibrium this implies that $N_\tau = \rho \bar{N}_\tau$ for all $\tau \in (t - \nu, t + \nu)$. Thus, N_τ is continuous at all $\tau \in (t - \nu, t + \nu)$. From this it is immediate that W_τ^N is continuously differentiable in τ for all for all $\tau \in (t - \nu, t + \nu)$, because we have that

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

$$+ e^{-(r+\rho)(t+\nu-\tau)} \left(p_\tau^N e^{-\int_\tau^{t+\nu} (\varepsilon_G + \lambda_G N_x)\, dx} + (1 - p_\tau^N) e^{-\int_\tau^{t+\nu} (\varepsilon_B + \lambda_B N_x)\, dx} \right) W_{t+\nu}^N$$

$$+ \int_\tau^{t+\nu} \rho e^{-(\rho+r)(s-\tau)} p_\tau^N \left(1 - e^{-\int_\tau^s (\varepsilon_G + \lambda_G N_x)\, dx} \right) ds$$

$$+ e^{-(r+\rho)(t+\nu-\tau)} p_\tau^N \left(1 - e^{-\int_\tau^{t+\nu} (\varepsilon_G + \lambda_G N_x)\, dx} \right) \frac{\rho}{\rho + r}.$$

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

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

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

$$+ e^{-(r+\rho)\Delta} \left(p_\tau^N e^{-\int_\tau^{\tau+\Delta} (\varepsilon_G + \lambda_G N_x) dx} + (1 - p_\tau^N) e^{-\int_\tau^{\tau+\Delta} (\varepsilon_B + \lambda_B N_x) dx} \right) W_{\tau+\Delta}^N$$

$$+ \int_\tau^{\tau+\Delta} \rho e^{-(r+\rho)(s-\tau)} p_\tau^N \left(1 - e^{-\int_\tau^s (\varepsilon_G + \lambda_G N_x) dx} \right) ds$$

$$+ e^{-(r+\rho)\Delta} p_\tau^N \left(1 - e^{-\int_\tau^{\tau+\Delta} (\varepsilon_G + \lambda_G N_x) dx} \right) \frac{\rho}{\rho+r}.$$

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 $\dot{W}_\tau^N = \lim_{\Delta \to 0} \frac{d}{d\tau} W_{\tau+\Delta}^N$ by continuous differentiability, we then obtain:

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

$$- \rho(2p_\tau - 1) - p_\tau(\varepsilon_G + \lambda_G N_\tau) \frac{\rho}{\rho+r}.$$

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

Proof of Lemma A.1.3: The proof of continuous differentiability of W_t^N follows along the same lines as in the proof of Lemma A.1.2. Lemma A.1.1 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

$$W_\tau^N = e^{-r(t+\nu-\tau)} \left(p_\tau^N e^{-\varepsilon_G(t+\nu-\tau)} + (1 - p_\tau^N) e^{-\varepsilon_B(t+\nu-\tau)} \right) W_{t+\nu}^N$$

$$+ p_\tau^N \int_\tau^{t+\nu} \varepsilon_G e^{-(\varepsilon_G+r)s} \frac{\rho}{\rho+r} ds.$$

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

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

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

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

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. ∎

Proof of Lemma A.1.4: Consider first the case in which $\varepsilon > 0$. Then trivially $p_t^N \to 1$ as $t \to \infty$. But for any t, $\frac{\rho}{\rho+r}\left(2p_t^N - 1\right) \leq W_t^N \leq \frac{\rho}{\rho+r}$. This implies that $\lim_{t\to\infty} W_t^N = \frac{\rho}{\rho+r}$ as claimed.

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

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

as:

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

$$\leq \frac{\rho}{r + \rho} \left(p_t^N - (1 - p_t^N)(1 - v) \right)$$

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

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

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

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

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

Proof of Lemma A.1.5: Suppose that $N_{t \geq 0}$ is an equilibrium and suppose for a contradiction that $t_1^* := \inf\{t : N_t > 0\} < \infty$. Pick $t \geq t_1^*$ such that $N_t > 0$. By right-continuity of N, we have $N_\tau > 0$ for all $\tau > t$ sufficiently close to t. This implies that

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

(A.1)

where the second inequality holds because $p_{t_1^*}^N = p_0 \leq \frac{1}{2}$. But the integral on the left-hand side is the expected payoff at time t_1^* to adopting at the first opportunity in the future, conditional on no breakdown having occurred prior to this opportunity. By optimality of the value to waiting, this is weakly less than $W_{t_1^*}^N$. Hence, (A.1)

implies that $W_{t_1^*}^N > 2p_{t_1^*} - 1$. By continuity of W^N and p^N, it follows that for all $s \geq t_1^*$ sufficiently close to t_1^*, $W_s^N > 2p_s^N - 1$ and hence $N_s = 0$, contradicting the definition of t_1^*.

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

$\blacksquare$

A.2.2 Equilibrium under Perfect Bad News (Theorem 1.3.2)

In this section we prove Theorem 1.3.2. For this we do *not* impose Conditions 1.3.3 or 1.3.4. Recall the following definitions which we motivated in Section 1.3.2: Define

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

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

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

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

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

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

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

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

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

The proof of Theorem 1.3.2 proceeds in three steps. Suppose that $N_{t\geq 0}$ is an equilibrium with associated cutoff times t_1^* and t_2^* as defined by Equation (1.2). We first show in Lemma A.2.1 that if $t_1^* < t_2^*$, then at all $t \in (t_1^*, t_2^*)$, N_t is pinned down by a simple ODE. Second, Lemma A.2.3 provides a characterization of t_1^* and t_2^* in terms of the evolution of (p_t, Λ_t). Given these two steps, it is easy to see that if an equilibrium exists, it is unique and must take the Markovian form in Equation (1.4) of Theorem 1.3.2. Finally, to verify equilibrium existence, Lemma A.2.4 shows that the adoption flow implied by Equation (1.4) is feasible.

Characterization of Adoption between t_1^* and t_2^*

Lemma A.2.1. *Suppose $N_{t\geq 0}$ is an equilibrium with associated no-news posterior $p_{t\geq 0}$ and cutoff times t_1^* and t_2^* as defined by Equation (1.2). Suppose that $t_1^* < t_2^*$. Then at all times $t \in (t_1^*, t_2^*)$,*

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

Proof. By definition of t_1^* and t_2^* and Theorem 1.3.1, we have $2p_t - 1 = W_t^N$ at all $t \in (t_1^*, t_2^*)$. Because p_t is weakly increasing, this implies that p_t and W_t^N are differentiable at almost all $t \in (t_1^*, t_2^*)$ (with respect to Lebesgue measure).

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

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

where the second equality follows from Equation (1.1). Consider any $t \in (t_1^*, t_2^*)$ at which W_t^N and p_t are differentiable. Combining the fact that $\dot{p}_t = p_t(1 - p_t)(\varepsilon + \lambda N_t)$

with (A.2), we obtain:

$$\dot{W}_t^N = (r + (\varepsilon + \lambda N_t)(1 - p_t)) \, W_t^N. \tag{A.3}$$

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

$$\dot{W}_t^N = 2\dot{p}_t = 2p_t(1 - p_t)(\varepsilon + \lambda N_t). \tag{A.4}$$

Combining (A.3), (A.4) and the fact that $W_t^N = 2p_t - 1$ then yields

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

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

As an immediate corollary of Lemma A.2.1 we obtain:

Corollary A.2.2. *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:

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

Characterization of Cutoff Times

Lemma A.2.3. *Let $N_{t \geq 0}$ be an equilibrium with corresponding no-news posterior $p_{t \geq 0}$ and cutoff times t_1^* and t_2^* as defined by Equation (1.2), and let $\Lambda_{t \geq 0} := \lambda \bar{N}_{t \geq 0}$ describe the evolution of the economy's potential for social learning. Then*

(i). $t_2^* = \inf\{t \geq 0 : \Lambda_t < \Lambda^*(p_t)\}$; *and*

(ii). $t_1^* = \min\{t_2^*, \sup\{t \geq 0 : p_t < p^*\}\}$.[1]

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.1.4 implies that $\lim_{t\to\infty} 2p_t - 1 > \lim_{t\to\infty} W_t$, whence $t_2^* < +\infty$. Moreover, p_t is strictly increasing for all $t > 0$.

For the first bullet point, note that by definition of $t_2^* := \sup\{t \geq 0 : N_t < \rho \bar{N}_t\}$, we have that $2p_t - 1 \geq W_t = G(p_t, \Lambda_t)$ for all $t \geq t_2^*$. This implies that $\Lambda_{t_2^*} \leq \Lambda^*(p_{t_2^*})$. Moreover, for all $t > t_2^*$, $\Lambda_t < \Lambda_{t_2^*}$ and $p_t > p_{t_2^*}$, so since Λ^* is non-decreasing we have $\Lambda_t < \Lambda^*(p_t)$. Suppose that $0 < t_2^*$. Then by continuity we must have $2p_{t_2^*} - 1 = W_{t_2^*} = G(p_{t_2^*}, \Lambda_{t_2^*})$ and so $\Lambda_{t_2^*} = \Lambda^*(p_{t_2^*})$. But since for all $s < t_2^*$ we have $\Lambda_s \geq \Lambda_{t_2^*}$ and $p_s < p_{t_2^*}$, this implies $\Lambda_s \geq \Lambda^*(p_s)$. This establishes (i).

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

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

(b) If $t_1^* > 0$, then $p_{t_1^*} \leq \bar{p}$.

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

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

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

For claim (b), note that if $t_1^* > 0$, then for all $t < t_1^*$, we have $N_t = 0$. Then for all $t < t_1^*$, $W_t \geq 2p_t - 1$ and by the proof of Lemma A.1.3, $\dot{W}_t = (r + (1 - p_t)\varepsilon)W_t$.

---

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

Since $W_{t_1^*} = 2p_{t_1^*} - 1$, we must then have

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

$$= r(2p_{t_1^*} - 1) - \varepsilon(1 - p_{t_1^*}),$$

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 A.2.1 implies that for all $\tau \in (t_1^*, t_2^*)$,

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

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

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

and hence by continuity $p_{t_1^*} \geq \overline{p}$ as claimed. This proves the lemma when either $\varepsilon > 0$ or $p_0 > \frac{1}{2}$. Finally, if $\varepsilon = 0$ and $p_0 \leq \frac{1}{2}$, then by Lemma A.1.5 $N_t = 0$ for all t. Thus, by definition, $t_1^* = t_2^* = +\infty$. Moreover, $p_t = p_0 \leq \frac{1}{2}$ and $\Lambda_t = \Lambda_0 > 0$ for all t, so $\inf\{t : \Lambda_t < \Lambda^*(p_t) = 0\} = \sup\{t : p_t < p^* = \frac{1}{2}\} = +\infty$, as required. ∎

Given Lemmas A.2.1 and A.2.3, it is immediate that if an equilibrium exists, then it must take the form of the adoption flow given by Equation (1.4) in Theorem 1.3.2. Moreover, it is easy to see that given initial parameters, Equation (1.4) uniquely pins down the times t_1^* and t_2^* as well as the joint evolution of p_t and N_t at all times (we elaborated on this in the main text), and that whenever $t_1^* < t_2^* < +\infty$, then $2p_t - 1 = W_t$ for all $t \in [t_1^*, t_2^*]$. Provided feasibility is satisfied, it is then easy to check that this adoption flow constitutes an equilibrium.

Feasibility

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

Lemma A.2.4. *Suppose $N_{t \geq 0}$ is an adoption flow satisfying Equation (1.4) in Theorem 1.3.2 such that $t_1^* < t_2^*$. Then for all $t \in (t_1^*, t_2^*)$, $N_t \leq \rho \bar{N}_t$.*

Proof. It suffices to show that

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

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

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

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

Equilibrium under Perfect Bad News without Condition 1.3.3

In this section, we discuss the case where $\rho \leq \varepsilon$, so that Condition 1.3.3 is violated. The previous sections established the equilibrium characterization of Theorem 1.3.2 without assuming Condition 1.3.3 . If $\rho \geq \varepsilon$, then $p^* = p^\sharp$, so because $\Lambda^*(p) = +\infty$ for all $p > p^\sharp$, we have:

$$N_t = \begin{cases} 0 & \text{if } \Lambda_t > \Lambda^*(p_t), \\ \rho \bar{N}_t & \text{if } \Lambda_t \leq \Lambda^*(p_t). \end{cases}$$
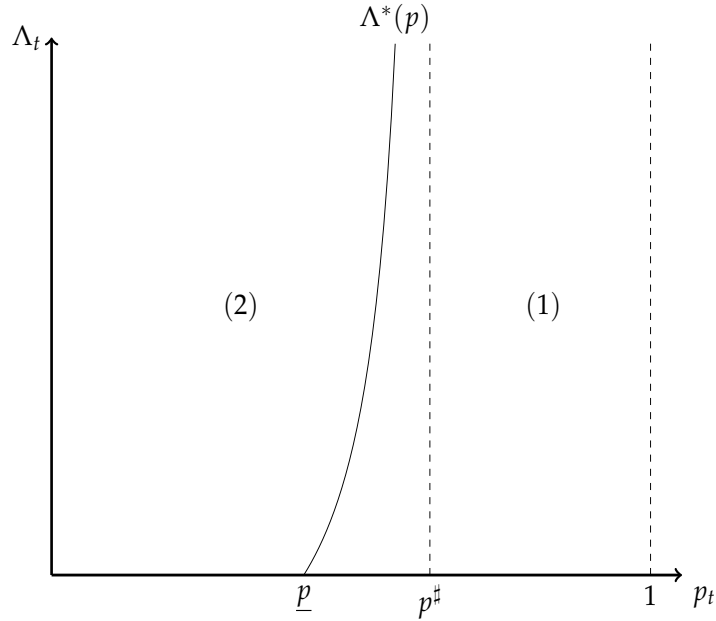
Figure A.1: *Partition of (p_t, Λ_t) when $\varepsilon \geq \rho$*

Thus, there is no region of partial adoption. As a result, it is easy to see that the saturation effect discussed in Section 1.4.2 is no longer present and welfare always strictly increases in response to an increase in the potential for social learning:

Proposition A.2.5. *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.*

A.2.3 Equilibrium under Perfect Good News (Theorem 1.3.6)

Theorem 1.3.6 follows immediately from Lemma A.2.6 and Lemma A.2.7:

Lemma A.2.6. *Let $N_{t \geq 0}$ be an equilibrium with associated cutoff times t_1^* and t_2^* given by Equation (1.3). Then $t_1^* = t_2^* =: t^*$.*

Proof. Suppose for a contradiction that $t_1^* < t_2^*$. From the definition of these cutoffs and Theorem 1.3.1 , we have that $2p_t - 1 = W_t$ for all $t \in (t_1^*, t_2^*)$. Then for all

$t \in (t_1^*, t_2^*)$ and $\Delta \in (0, t_2^* - t)$ we have:

$$W_t = p_t \int\limits_t^{t+\Delta} (\varepsilon + \lambda N_\tau) \, e^{-\int_t^\tau (\varepsilon + \lambda N_s) ds} e^{-r(\tau - t)} \frac{\rho}{\rho + r} d\tau +$$

$$\left((1 - p_t) + p_t e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) e^{-r\Delta} \left(2 p_{t+\Delta} - 1 \right),$$

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

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

$$p_t \int\limits_t^{t+\Delta} (\varepsilon + \lambda N_\tau) \, e^{-\int_t^\tau (\varepsilon + \lambda N_s) ds} e^{-r(\tau - t)} \frac{\rho}{\rho + r} d\tau$$

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

This implies that

$$W_t < p_t \left(1 - e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) \frac{\rho}{\rho + r}$$

$$+ \left((1 - p_t) + p_t e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) \left(2 p_{t+\Delta} - 1 \right)$$

$$\leq p_t \left(1 - e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) + \left((1 - p_t) + p_t e^{-\int_t^{t+\Delta} (\varepsilon + \lambda N_s) ds} \right) \left(2 p_{t+\Delta} - 1 \right)$$

$$= 2 p_t - 1,$$

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

Lemma A.2.7. *Let $N_{t\geq 0}$ be an equilibrium with corresponding cutoff time $t^* := t_1^* = t_2^*$ and no-news posterior $p_{t\geq 0}$. Then*

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

where

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

Proof. Define

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

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

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

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

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

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

A.2.4 Comparative Statics under PBN (Proposition 1.4.2)

As in the text, we impose Conditions 1.3.3 and 1.3.4 throughout this section. Define $\overline{\Lambda}_0 := \max\{\Lambda^*(p_0), \Lambda^*(\overline{p})\}$. We first prove Lemma 1.3.5 from Section 1.3.2:

Proof of Lemma 1.3.5: We show that $t_1^*(\Lambda_0) < t_2^*(\Lambda_0)$ if and only if $\Lambda_0 > \overline{\Lambda}_0$.

Suppose first that $\Lambda_0 > \overline{\Lambda}_0$. Then by the proof of the first part of Lemma A.2.3, we must have $t_2^* > 0$ and $\Lambda_{t_2^*} = \Lambda^*(p_{t_2^*})$. If $t_1^* = t_2^* =: t^*$, then by claims (a) and (b) in the proof of Lemma A.2.3, we must have $p_{t^*} \leq \overline{p}$. But combining these statements, we get

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

which is a contradiction.

Suppose conversely that $t_1^* < t_2^*$. Then by the proof of Lemma A.2.3, we have that $\Lambda^*(p_{t_1^*}) < \Lambda_{t_1^*} = \Lambda_0$. That proof also implies that if $0 < t_1^* < t_2^*$, then $p_{t_1^*} = \overline{p} \geq p_0$; and if $0 = t_1^* < t_2^*$, then $p_{t_1^*} = p_0 \geq \overline{p}$. Thus, either way $\Lambda_0 > \overline{\Lambda}_0$, as claimed. ∎

The following three subsections prove Proposition 1.4.2, by considering the effect of an increase in Λ_0 on welfare, learning, and adoption behavior, respectively.

Comparative Statics of Welfare

We prove a slightly more general result than in Proposition 1.4.2: We allow for any $p_0 \in (0,1)$ and show that

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

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

If $p_0 \in (\overline{p}, p^\sharp)$ as in Proposition 1.4.2, then $\overline{\Lambda}_0 = \Lambda^*(p_0)$, so we get the result in Proposition 1.4.2.

To prove the first bullet point, consider $\hat{\Lambda}_0 := \Lambda_0^2 > \Lambda_0 := \Lambda_0^1 > \overline{\Lambda}_0$ with corresponding cutoff times t_1^i and t_2^i, value to waiting W_t^i, and no-news posteriors p_t^i for $i = 1, 2$. By Lemma 1.3.5, we have $t_1^i < t_2^i$ for $i = 1, 2$. Moreover, by the proof of Lemma A.2.3, we have $\max\{p_0, \overline{p}\} = p_{t_1^1}^1 = p_{t_1^2}^2$. Because $N_t^i = 0$ for all $t < t_1^i$ for both $i = 1, 2$, this implies that $t_1^1 = t_1^2 = t_1$. Then

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

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

For the second bullet point, suppose $\Lambda_0^1 < \Lambda_0^2 \leq \overline{\Lambda}_0$. By Lemma 1.3.5, we must have $t_1^i = t_2^i =: t^i$. Let $\hat{t} := \min\{t^1, t^2\}$. Then note that for all $t \leq \hat{t}$, $p_t^1 = p_t^2$ and $\Lambda_{\hat{t}}^i = \Lambda_0^i$. By Lemma A.2.3 this implies that either $0 = t^1 = t^2$ or $t^1 < t^2$. If $0 = t^1 = t^2$, then for all $t > 0$, we have $2p_t^i - 1 > W_t^i$ and

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

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

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

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

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

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

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

Comparative Statics of Learning

In this section and Section A.2.4, we assume as in Proposition 1.4.2 that $p_0 \in (\bar{p}, p^\sharp)$. This implies that $t_1^* = 0$ and $\overline{\Lambda}_0 = \Lambda^*(p_0) < +\infty$.

Note first that $p_t^{\Lambda_0}$ is strictly increasing in Λ_0 for all $\Lambda_0 \in (0, \Lambda^*(p_0))$ since in this case $t_2^*(\Lambda_0) = 0$ so that

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

Suppose next that $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. To prove the non-monotonicity result in item (ii) of Proposition 1.4.2, we first prove the following lemma:

Lemma A.2.8. *Suppose that* $\hat{\Lambda}_0 = \hat{\lambda}\hat{N}_0 > \Lambda_0 = \lambda\bar{N}_0 > \Lambda^*(p_0)$, *with corresponding equilibrium flows of adoption* $\hat{N}_{t\geq 0}$ *and* $N_{t\geq 0}$. *Then*

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

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

Proof. Suppose that $\hat{\Lambda}_0 > \Lambda_0 > \overline{\Lambda}_0 = \Lambda^*(p_0)$. Then by Lemma 1.3.5, we have $t_2^*(\hat{\Lambda}_0), t_2^*(\Lambda_0) > 0$. Let $t_2^* = \min\{t_2^*(\hat{\Lambda}_0), t_2^*(\Lambda_0)\}$. Then because $p_0 = p_0^{\Lambda_0} = p_0^{\hat{\Lambda}_0}$, the ODE in Corollary A.2.2 implies that at all times $t < t_2^*$, we have $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0} = p_t$. By Lemma A.2.1, this implies that for all $t < t_2^*$,

$$\lambda N_t = \hat{\lambda}\hat{N}_t. \tag{A.5}$$

Note that Equation A.5 implies that

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

Because $p_{t_2^*}^{\Lambda_0} = p_{t_2^*}^{\hat{\Lambda}_0}$, Lemma A.2.3 implies that $t_2^* = t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$.

From this and Equation A.5, it is then immediate that $\lambda N_t = \hat{\lambda}\hat{N}_t$ for all $t < t_2^* = t_2^*(\Lambda_0)$. ∎

Proof of item (ii) of Proposition 1.4.2: Suppose that $\hat{\Lambda}_0 > \Lambda_0 \geq \Lambda^*(p_0)$. By Lemma A.2.8, $t^* := t_2^*(\Lambda_0) < t_2^*(\hat{\Lambda}_0)$, $\lambda N_t = \hat{\lambda}\hat{N}_t$, and $p_t^{\Lambda_0} = p_t^{\hat{\Lambda}_0}$ for all $t \leq t^*$, which proves the first bullet point.

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

$$\lim_{t \uparrow t^*} \lambda N_t < \lim_{t \downarrow t^*} \lambda N_t. \tag{A.6}$$

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 1.1,

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

which by Lemma A.2.3 implies

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

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

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

111

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

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

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

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

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

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

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

which as above implies that

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

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

$t > \bar{t}$, as claimed by the third bullet point. ∎

Comparative Statics of Adoption Behavior

Adoption of Good Products: *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. First note that because $p_0 \geq \bar{p}$, $t_1^*(\Lambda_0) = t_1^*(\hat{\Lambda}_0) = 0$.

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 1.3.4). Let $t^* = t_2^*(\Lambda_0)$. By Lemma A.2.8, $\lambda N_t = \hat{\lambda} \hat{N}_t$ for all $t < t^*$. Then for all $t < t^*$

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

. 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:

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

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

Adoption of Bad Products: *For all $t > 0$, $A_t(\Lambda_0, B)$ is strictly decreasing in Λ_0.*

Proof. Recall that $A_t(\lambda, \bar{N}_0, B)$ denotes the *expected* proportion of adopters at time t

conditional on $\theta = B$, that is, letting $N_{t \geq 0}$ denote the associated equilibrium

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

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

where the final equality follows from integration by parts. Moreover, from the Markovian description of equilibrium in Equation (1.4), it is easy to see that this expression depends on λ and $\bar{N}_0$ only through $\Lambda_0 = \lambda \bar{N}_0$, so we can denote it by $A_t(\Lambda_0, B)$. Then it suffices to prove the claim when λ is increased to $\hat{\lambda} > \lambda$ holding fixed $\bar{N}_0$, because for any $\hat{\Lambda}_0 > \Lambda_0$ there exists $\bar{N}_0$ and $\hat{\lambda} > \lambda$ such that $\hat{\Lambda}_0 = \hat{\lambda} \bar{N}_0$ and $\Lambda_0 = \lambda \bar{N}_0$.

Let $N_{t \geq 0}$ and $\hat{N}_{t \geq 0}$ be the equilibrium under λ and $\hat{\lambda}$, respectively. Note that when $\bar{p} \leq p_0$, $N_t > 0$ for all $t > 0$. Given an arbitrary strictly positive adoption flow $M_{s \geq 0}$ and $t > 0$, note that the map

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

is strictly decreasing in λ. This implies that for all $t > 0$,

$$\int_0^t N_\tau e^{-\int_0^\tau (\varepsilon + \lambda N_s) ds} d\tau > \int_0^t N_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau. \tag{A.7}$$

We now show that

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

which together with (A.7) 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

$$\int_0^t N_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau < \int_0^t \hat{N}_\tau e^{-\int_0^\tau (\varepsilon + \hat{\lambda} \hat{N}_s) ds} d\tau. \tag{A.8}$$

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

$$\int_0^t \varepsilon e^{-\int_0^\tau (\varepsilon + \hat{\lambda} N_s) ds} d\tau \leq \int_0^t \varepsilon e^{-\int_0^\tau (\varepsilon + \hat{\lambda} \hat{N}_s) ds} d\tau \tag{A.9}$$

for all $t \geq 0$. Inequalities (A.8) and (A.9) together imply:

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

But this is equivalent to

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

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

A.2.5 Comparative Statics under PGN (Proposition 1.4.3)

We prove Proposition 1.4.3.

Strict Welfare Gains

Proof. 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 $\hat{\Lambda}_0$:

$\pi^* = \frac{p_0 - p^*}{1 - p^*}$. Because learning occurs at the same exogenous rate ε once p^* is reached, the continuation value W^* conditional on p^* being reached is also the same: $W^* = p^* \int_0^\infty \varepsilon e^{-(\varepsilon + r)t} \frac{\rho}{r + \rho} dt = 2p^* - 1$. So the only difference is that conditional on no breakthroughs, the time t^* at which p^* is reached occurs earlier under $\hat{\Lambda}_0$. To see that this is strictly beneficial, note that W_0 is composed of the following two terms:

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

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

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

∎

Learning Speeds Up

Proof. If $p_0 > p^*$, then conditional on no breakthroughs, all consumers adopt immediately upon an opportunity until the time t^* at which the cutoff posterior p^* is reached. By Theorem 1.3.6, we have that for all $t < \min\{t^*(\hat{\Lambda}_0), t^*(\Lambda_0)\}$, $\lambda N_t = \rho e^{-\rho t} \Lambda_0 < \rho e^{-\rho t} \hat{\Lambda}_0 = \hat{\lambda} \hat{N}_t$. Since $p^* = \frac{(\varepsilon + r)(\rho + r)}{2(\varepsilon + \rho)(\varepsilon + r) - \varepsilon \rho}$ is independent of the

potential for social learning, this implies that $t^*(\hat{\Lambda}_0) < t^*(\Lambda_0)$ and that $p_t^{\hat{\Lambda}_0} < p_t^{\Lambda_0}$ for all $t > 0$. Moreover, once the cutoff posterior is reached, information is generated at the constant exogenous rate ε, which means that conditional on $t > t^*$, beliefs depend only on $t - t^*$, as summarized in the third bullet point. ∎

No Initial Slow-Down in Adoption

Proof. From Section A.2.5, $t^*(\hat{\Lambda}_0) < t^*(\Lambda_0)$. Thus, at all times $t \leq t^*(\hat{\Lambda}_0)$, all consumers adopt upon first opportunity in both equilibria. ∎

A.2.6 Proof of Theorem 1.4.4

We first establish the following basic mathematical fact:

Lemma A.2.9. *Suppose $\bar{t} > t^* \geq 0$ and consider $f, g : [0, \infty) \to \mathbb{R}$ such that $f(\tau) = g(\tau)$ for all $\tau \leq t^*$, $f(\tau) < g(\tau)$ for $\tau \in (t^*, \bar{t})$, and $f(\tau) > g(\tau)$ for all $\tau > \bar{t}$. Suppose that $\int_0^\infty e^{-r\tau} f(\tau) d\tau = \int_0^\infty e^{-r\tau} g(\tau) d\tau$ for some $r > 0$. Then for all $\hat{r} > r$,*

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

Proof. We have

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

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

To prove Theorem 1.4.4, fix $0 < r_p < r_i$, $\rho > 0$, $\bar{N}_0^p > 0$ and $p_0 \in (\frac{1}{2}, \frac{\rho+r_p}{\rho+2r_p})$. Consider $\hat{\lambda} > \lambda > 0$ such that $\hat{\lambda}\bar{N}_0^p > \lambda\bar{N}_0^p > \Lambda_{r_p}^*(p_0)$. As in the text, we assume that there is no exogenous news source. The following lemma derives the equilibrium of the game with a sufficiently small mass of impatient types:

Lemma A.2.10. *There exists $\eta > 0$ such that whenever $\bar{N}_0^i < \eta$, the unique equilibrium for $\gamma \in \{\lambda, \hat{\lambda}\}$ takes the following form: There exists some $t^*(\gamma)$ such that the equilibrium flows N^i and N^p of impatient and patient adopters satisfy:*

$$
N_t^i = \rho\bar{N}_t^i \text{ for all } t,
$$

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

Proof. Fix $\gamma \in \{\lambda, \hat{\lambda}\}$. Pick $\eta > 0$ such that $p_0 > \frac{\eta+r}{\eta+2r}$. Consider first the game consisting only of mass $\bar{N}_0^p$ consumers of type r_p (and no consumers of type r_i). If there were an exogenous news source in this game which generated signals at rate

118

$\varepsilon \leq \eta$, then by Theorem 1.3.2 type r_p would always weakly prefer to adopt absent breakdowns. Then it is easy to see that in the game with no exogenous news source but with mass $\bar{N}_0^i < \eta$ of types r_i, type r_p will also always weakly prefer to adopt. This implies that type r_i must always *strictly* prefer to adopt.

Thus, $N_t^i = \rho \bar{N}_t^i$ for all t. Given this, the game reduces to one in which patient types view the information generated by the impatient types as a non-stationary exogenous news source which generates signals at rate $\varepsilon_t = \gamma \rho \bar{N}_t^i$. Modifying the arguments in the proof of Theorem 1.3.1, there must exist some $t^*(\gamma) > 0$ such that r_p is indifferent between adoption and delay for $t \leq t^*(\gamma)$, and r_p strictly prefers to adopt at all times $t > t^*(\gamma)$. Then the unique equilibrium can be derived in the same manner as in the proof of Theorem 1.3.2. ∎

Given Lemma A.2.10, we can follow the arguments in the proof of Proposition 1.4.2 to show that $t^*(\lambda) < t^*(\hat{\lambda})$ and that there exists some $\bar{t} > t^*(\lambda)$ such that

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

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

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

Since r_p is initially indifferent between adoption and delay under both λ and $\hat{\lambda}$, we have $W_0^p(\lambda) = W_0^p(\hat{\lambda}) = 2p_0 - 1$. But then applying Lemma A.2.9 yields $W_0^i(\lambda) > W_0^i(\hat{\lambda})$. This completes the proof of Theorem 1.4.4.

Appendix B

Appendix to Chapter 2

B.1 Extension of Dekel and Fudenberg (1990)

In this section, we briefly discuss Dekel and Fudenberg's (1990) characterization of $S^\infty W$: They show that in a complete information game G, $S^\infty W(G)$ represents the set of strategies which survive iterated admissibility in a sequence of *elaborations* $\tilde{G}_n$ converging to G, which differ from G in allowing for vanishingly small amounts of payoff uncertainty. We extend this characterization to the incomplete information game G defined in Section 2.2, but show that this result is once again very sensitive to the way in which the small amounts of additional payoff uncertainty are taken to interact with the uncertainty already present in G.

In our incomplete information setting, a sequence of elaborations is a sequence of I-player games $\tilde{G}_n$ with the same underlying strategy sets A_i and the same state space Θ as G. There are finite type spaces T_i for each player, and in each $\tilde{G}_n$, payoff functions are given by $\tilde{u}_i^n \colon A_i \times A_{-i} \times \Theta \times T_i \to [0,1]$ and beliefs by $\kappa_i^n \colon T_i \to \Delta^\circ(\Theta \times T_{-i})$. To capture the idea that payoff uncertainty is vanishingly small, we consider sequences of elaborations $\tilde{G}_n$ which *converge* to G, in the sense that

in each elaboration there is a "sane" type for each player which has exactly the same payoffs as in G and assigns increasingly large probability to his opponents being sane. Analogously to the two notions of ACC of admissibility we defined in Section 2.2.2, the question arises whether or not to allow sane types' uncertainty about opponents' payoffs to coincide with small amounts of uncertainty about opponents' beliefs about Θ. Consider first the following strong notion of convergence, which requires the sane types in each elaboration to have *exactly* the same marginal beliefs on Θ as in G:

Definition B.1.1. A sequence of elaborations $\tilde{G}_n$ **converges strongly** to G, denoted $\tilde{G}_n \xrightarrow{S} G$, if for every i there exists a "sane" type $\bar{t}_i \in T_i$ such that

(i). $\text{marg}_\Theta \kappa_i^n(\bar{t}_i) = \phi_i$ for all n;

(ii). $\tilde{u}_i^n(\cdot, \theta, \bar{t}_i) = u_i(\cdot, \theta)$ for all $\theta \in \Theta$ and for all n;

(iii). for all θ, $\kappa_i^n(\bar{t}_i)(\{\theta\} \times \{\bar{t}_{-i}\}) \to \phi_i(\theta)$ as $n \to \infty$.

Then interpreting iterated dominance in both G and in each $\tilde{G}_n$ in the interim correlated sense,[1] we have the following extension of Dekel and Fudenberg's main result:

Proposition B.1.2 (Extension of Proposition 3.1 in Dekel and Fudenberg (1990))**.** *Let $\bar{a} \in A$. Then $\bar{a} \in S^\infty W(G)$ if and only if there exists a sequence of elaborations $\tilde{G}_n \xrightarrow{S} G$ such that $\bar{a}_i \in W^\infty(\tilde{G}_n)(\bar{t}_i)$ for the sane type $\bar{t}_i$ of each player i and for all n.*

However, if we allow for vanishingly small perturbations to sane types' beliefs about Θ (i.e. in condition (i) of Definition B.1.1 we impose equality only in the limit as $n \to \infty$), then $S^\infty W(G)$ is once again only a (generally strict) subset of the behavior

---

[1] Interim correlated dominance in $\tilde{G}_n$ allows each type to hold correlated beliefs over states of nature, opponent behavior, and opponent types; cf. Dekel *et al.* (2007) for the formal definition.

predicted by W^∞ in a sequence of elaborations. This is easy to see by considering the game in Example 2.3.6, along with the sequence of elaborations with a single, sane type $\bar{t}_i$ for each player who has beliefs $\phi^n \in \Delta(\Theta)$ as in Example 2.3.6—here $D \in W^\infty(\tilde{G}_n)(\bar{t}_1)$ for all n even though $D \notin S^\infty W(G)$.[2]

B.2 Proofs

Proof of Lemma 2.2.2. Fix i and consider the 2-player game complete information game $\tilde{G}$ with strategy sets $\tilde{A}_1 = \hat{A}_i$ (with typical elements a_i) and $\tilde{S}_2 = \hat{A}^\Theta_{-i}$ (with typical elements $\lambda_i \colon \Theta \to \hat{A}_{-i}$), where player 1's payoffs are given by $g_1(a_i, \lambda_i) = \sum_{\theta, a_{-i}} \phi_i(\theta) \lambda_i(\theta)[a_{-i}] u_i(a_i, a_{-i}, \theta)$ and player 2's payoffs are arbitrary.

Then a_i is not strongly (respectively weakly) dominated in $\hat{A}_i$ if and only if a_i is not strongly (respectively weakly) dominated for player 1 in $\tilde{G}$, which by Pearce (1984, Lemma 3) (respectively Pearce (1984, Lemma 4)) is equivalent to the existence of some $\lambda_i \in \Delta(\hat{A}^\Theta_{-i})$ (respectively $\lambda_i \in \Delta^\circ(\hat{A}^\Theta_{-i})$) to which a_i is a best response in $\tilde{G}$. But this is equivalent to there being a belief $\lambda_i \colon \Theta \to \Delta(\hat{A}_{-i})$ (respectively $\lambda_i \colon \Theta \to \Delta^\circ(\hat{A}_{-i})$) to which a_i is a best response in $\hat{A}_i$. ∎

Proof of Lemma 2.2.3. Immediate from the definitions. ∎

Proof of Theorem 2.3.1. Given Corollary 2.3.3 and Proposition 2.3.4, which are proved below, the proof follows from the discussion in Section 2.3.1. ∎

Proof of Lemma 2.3.2. We adapt Börgers's (1994) proof to our setting: By finiteness of I and of the strategy sets for each player, it suffices to find $\pi \in (0,1)$ with the required properties for some fixed i, $\hat{A}_i \subseteq A_i$ and $\hat{A}_{-i} \subseteq A_{-i}$. Note that by finiteness

---

[2]See the earlier working paper version of this paper for a general characterization of the robust extension of W^∞ under this notion of "convergence with perturbed priors".

of the strategy sets and because the correspondence mapping every $p \in (0,1)$ to the set $D_i^p(\tilde{A}_i, \hat{A}_{-i})$ is non-empty valued and decreasing in p (with respect to set inclusion), there exists $\pi \in (0,1)$ such that for all $p \geq \pi$, $D_i^p(\hat{A}_i, \hat{A}_{-i}) = D_i^\pi(\hat{A}_i, \hat{A}_{-i})$. We claim that π is as required.

For the "only if" part, suppose that $a_i \in D_i^\pi(\hat{A}_i, \hat{A}_{-i})$. By definition of $D_i^\pi(\hat{A}_i, \hat{A}_{-i})$, a_i is a best response in $\hat{A}_i$ to some $\mu_i \colon \Theta \to \Delta^\circ(A_{-i})$. Hence, by Lemma 2.2.2, there is no $\alpha_i \in \Delta(\hat{A}_i)$ which weakly dominates a_i on A_{-i}, so condition (i) holds. To establish condition (ii), consider a sequence (p_n) such that $p_n \in (\pi, 1)$ for all n and $\lim_{n \to \infty} p_n = 1$. Then $a_i \in D_i^{p_n}(\hat{A}_i, \hat{A}_{-i}) = D_i^\pi(\hat{A}_i, \hat{A}_{-i})$ for all n. So for all n, a_i is best response in $\hat{A}_i$ to some belief $\mu_i^n \colon \Theta \to \Delta(A_{-i})$ such that $\mu_i^n(\theta)[\hat{A}_{-i}] \geq p_n$ for all θ. Passing to a subsequence if necessary, let μ_i denote the limit of the μ_i^n. Then $\mu_i \colon \Theta \to \Delta(\hat{A}_{-i})$ and a_i is a best response in $\hat{A}_i$ to μ_i. Hence by Lemma 2.2.2, there is no $\alpha_i \in \Delta(\hat{A}_i)$ which strongly dominates a_i on $\hat{A}_{-i}$, so condition (ii) is satisfied.

For the "if" part, suppose that $a_i \in \hat{A}_i$ satisfies conditions (i) and (ii). By Lemma 2.2.2, condition (i) implies that a_i is a best response in $\hat{A}_i$ to some $\lambda_{i,1} \colon \Theta \to \Delta^\circ(A_{-i})$. Condition (ii) implies that a_i is a best response in $\hat{A}_i$ to some $\lambda_{i,2} \colon \Theta \to \Delta(\hat{A}_{-i})$. Given any $p \in [\pi, 1)$, a_i is still a best response in $\hat{A}_i$ to the convex combination $\lambda_i^p := (1 - p)\lambda_{i,1} + p\lambda_{i,2}$. Moreover, $\lambda_i^p \colon \Theta \to \Delta^\circ(A_{-i})$ and $\lambda_i^p(\theta)[\hat{A}_{-i}] \geq p\lambda_{i,2}(\theta)[\hat{A}_{-i}] = p$ for all θ. Hence, $a_i \in D_i^\pi(\hat{A}_i, \hat{A}_{-i})$. ∎

Proof of Corollary 2.3.3. We again adapt the argument in Börgers (1994): It is easy to see that for any $p \in (0,1)$ and $n \geq 0$, we have $\Lambda_i^{p,n+1}(A) := D_i^p(A_i, \Lambda_{-i}^{p,n}(A)) = D_i^p(\Lambda_i^{p,n}(A), \Lambda_{-i}^{p,n}(A))$. Hence, $\tilde{R}_i^p(G)$ can be determined as follows: First delete all strategies for every player j that are not a best response to any belief $\lambda_j \colon \Theta \to \Delta^\circ(A_{-j})$. Then, *among all the remaining strategies*, delete all strategies for every player j that are not a best response to any belief $\lambda_j \colon \Theta \to \Delta^\circ(A_{-j})$ that in every state

assigns probability at least p to the remaining strategies of the opponents. Iterate this procedure until no further strategies can be deleted, which by the finiteness of the strategy sets will happen in some finite number m of steps. The resulting sets of strategies for each player i will be $\tilde{R}_i^p(G) = \Lambda_i^{p,m}(A)$.

Now let π be as found in Lemma 2.3.2. Then Lemma 2.3.2 implies that for $p \geq \pi$, the above procedure is equivalent to the procedure in which at each step all strategies are eliminated that are either weakly dominated in the original game or strongly dominated in the remaining reduced game. This in turn is equivalent to first deleting all weakly dominated strategies, and then at each later step deleting all strategies that are strongly dominated in the remaining game. This proves that for all $p \geq \pi$, we have $\tilde{R}_i^p(G) = S^\infty W(E)_i$, as claimed. ∎

Proof of Proposition 2.3.4 and of Theorem 2.3.5. We give a combined proof of Proposition 2.3.4 and of Theorem 2.3.5 by showing that for all $p \in (0,1)$ and for all $\varepsilon \geq 0$,

$$\text{Proj}_A CB^p(([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) = \tilde{R}^{\varepsilon, p}(G).$$

Adapting the argument in Hu (2007) to our setting, we note first that for all n

$$\text{Proj}_A B^{p,n}([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) = \tilde{\Lambda}^{\varepsilon, p, n}(A).$$

To see this, we proceed by induction on n and repeatedly invoke Lemma B.2.1, which is stated and proved below. For $n = 0$, applying Lemma B.2.1 with $B = A_{-i} \times T_{-i}$ yields $\text{Proj}_{A_i}([\phi_i, \varepsilon] \cap \mathcal{R}_i \cap \mathcal{P}_i) = \tilde{\Lambda}_i^{\varepsilon, p, 0}(A)$. Since $\text{Proj}_A([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P}) = \prod_i \text{Proj}_{A_i}([\phi_i, \varepsilon] \cap \mathcal{R}_i \cap \mathcal{P}_i)$, it follows that $\text{Proj}_A B^{p,0}([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) := \text{Proj}_A([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P}) = \tilde{\Lambda}^{\varepsilon, p, 0}(A)$ as required. Assuming the claim holds for n,

apply Lemma B.2.1 with $B = \text{Proj}_{A_{-i} \times T_{-i}} B^{p,n} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta))$. This yields

$$\text{Proj}_{A_i} \left(B_i^p \left(\Theta \times A_i \times T_i \times \text{Proj}_{A_{-i} \times T_{-i}} B^{p,n} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) \mid E(\Theta) \right) \right.$$
$$\left. \cap ([\phi_i, \varepsilon] \cap \mathcal{R}_i \cap \mathcal{P}_i) \right) = \Lambda_i^{\varepsilon, p} (A_i \times \text{Proj}_{A_{-i}} B^{p,n} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)))$$
$$= \Lambda_i^{\varepsilon, p, n+1}(A),$$

where the last equality uses the induction hypothesis. Taking the product across all i yields

$$\text{Proj}_A \left(\bigcap_i B_i^p \left(\Theta \times A_i \times T_i \times \text{Proj}_{A_{-i} \times T_{-i}} B^{p,n} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) \mid E(\Theta) \right) \right.$$
$$\left. \cap ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P}) \right) = \Lambda^{\varepsilon, p, n+1}(A).$$

Finally, since $\Lambda^{\varepsilon, p, n+1}(A) \supseteq \Lambda^{\varepsilon, p, n}(A) = \text{Proj}_A B^{p,n} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta))$, we can intersect the LHS of the previous equality with $\text{Proj}_A B^{p,n} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta))$ without affecting the equality. Performing this intersection and simplifying yields $\text{Proj}_A B^{p,n+1} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) = \Lambda^{\varepsilon, p, n+1}(A)$, as required.

Given this it is immediate that $\text{Proj}_A CB^p(([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) \subseteq \tilde{R}^{\varepsilon, p}(G)$. For the converse, suppose $a_i \in \tilde{R}_i^{\varepsilon, p}(G)$. Then by the previous paragraph, there is a sequence $\{t_i^n\}_{n=1}^\infty \subseteq T_i$ such that $(a_i, t_i^n) \in \text{Proj}_{A_i \times T_i} B_i^p \left(B^{p,n-1} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) \mid E(\Theta) \right)$ for all n. Since T_i is compact metric, there is a subsequence $\{t_i^{n_k}\}_{k=1}^\infty$ of $\{t_i^n\}_{n=1}^\infty$ which converges to some $t_i \in T_i$. By continuity of β_i, this implies $(a_i, t_i) \in \text{Proj}_{A_i \times T_i} B_i^p \left(B^{p,n_k-1} ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta)) \mid E(\Theta) \right)$ for all k. It follows that $a_i \in \text{Proj}_{A_i} CB^p ([\phi, \varepsilon] \cap \mathcal{R} \cap \mathcal{P} \mid E(\Theta))$, as required. ∎

Lemma B.2.1. *For all $B \subseteq A_{-i} \times T_{-i}$, $p \in (0, 1)$ and $\varepsilon \geq 0$,*

$$\text{Proj}_{A_i} \left(B_i^p (\Theta \times A_i \times T_i \times B \mid E(\Theta)) \cap ([\phi_i, \varepsilon] \cap \mathcal{R}_i \cap \mathcal{P}_i) \right) = \tilde{\Lambda}_i^{\varepsilon, p} \left(A_i \times \text{Proj}_{A_{-i}} B \right).$$

Proof of Lemma B.2.1. Suppose that

$$(a_i, t_i) \in \text{Proj}_{A_i \times T_i} \left(B_i^p \left(\Theta \times A_i \times T_i \times B \mid E(\Theta) \right) \cap \left([\phi_i, \varepsilon] \cap \mathcal{R}_i \cap \mathcal{P}_i \right) \right).$$

Then $f_i(t_i)$ has full support on $\Theta \times A_{-i}$; for all $\theta \in \Theta$, we have $\frac{f_i(t_i)(\{\theta\} \times \text{Proj}_{A_{-i}} B)}{f_i(t_i)(\{\theta\} \times A_{-i})} \geq p$; a_i maximizes expected payoffs given $f_i(t_i)$; and there exists $\phi_i' \in \Delta(\Theta)$ such that $||\phi_i' - \phi||_\infty \leq \varepsilon$ and such that $\phi_i'(\theta) = \text{marg}_\Theta \beta_i(t_i)(\theta) = f_i(t_i)(\{\theta\} \times A_{-i})$ for all θ. Defining $\mu: \Theta \to \Delta^\circ(A_{-i})$ by $\mu(\theta)(a_{-i}) = \frac{f_i(t_i)(\theta, a_{-i})}{f_i(t_i)(\{\theta\} \times A_{-i})}$, we get that $\mu(\theta)(\text{Proj}_{A_{-i}} B) \geq p$ for all θ and $a_i \in BR^{\phi_i'}(\mu)$. This implies that $a_i \in \tilde{\Lambda}_i^{\varepsilon, p} \left(A_i \times \text{Proj}_{A_{-i}} B \right)$, as required.

For the converse, suppose that $a_i \in \tilde{\Lambda}_i^{\varepsilon, p} \left(A_i \times \text{Proj}_{A_{-i}} B \right)$. Then there exists $\mu: \Theta \to \Delta^\circ(A_{-i})$ such that $\mu(\theta)(\text{Proj}_{A_{-i}} B) \geq p$ for all θ and such that $a_i \in BR_i^{\phi_i'}(\mu)$ for some $\phi_i' \in \Delta(\Theta)$ with $||\phi_i' - \phi_i||_\infty \leq \varepsilon$. Define a map $\Phi_i: A_{-i} \to A_{-i} \times T_{-i}$ by $\Phi_i(a_{-i}) = (a_{-i}, t_{-i}[a_{-i}])$, where if $a_{-i} \in \text{Proj}_{A_{-i}} B$ then $t_{-i}[a_{-i}]$ is chosen such that $(a_{-i}, t_{-i}[a_{-i}]) \in B$ and otherwise $t_{-i}[a_{-i}] \in T_{-i}$ is arbitrary. Consider the type $t_i \in T_i$ given by $\beta_i(t_i)(\theta, a_{-i}, t_{-i}) = \phi_i'(\theta) \cdot \mu(\theta) \left(\Phi_i^{-1}(a_{-i}, t_{-i}) \right)$ for all θ, a_{-i}, and t_{-i}. Then note that $\text{marg}_\Theta \beta_i(t_i) = \phi_i'$; $f_i(t_i)(\theta, a_{-i}) = \phi_i'(\theta) \cdot \mu(\theta)(a_{-i}) > 0$ for all θ and a_{-i}; and $\frac{\beta_i(t_i)(\{\theta\} \times B)}{\beta_i(t_i)(\{\theta\} \times A_{-i} \times T_{-i})} = \mu(\theta)(\text{Proj}_{A_{-i}} B) \geq p$ for all θ. Thus, $(a_i, t_i) \in \text{Proj}_{A_i} \left(B_i^p \left(\Theta \times A_i \times T_i \times B \mid E(\Theta) \right) \cap \left([\phi_i, \varepsilon] \cap \mathcal{R}_i \cap \mathcal{P}_i \right) \right)$, as required. $\blacksquare$

Proof of Proposition B.1.2. We adapt the proof technique of Proposition 3.1 in Dekel and Fudenberg (1990) to our setting: For the "only if" direction, suppose $\bar{a} \in S^\infty W(G)$. We construct a sequence of elaborations $\tilde{G}_n$ converging strongly to G such that for the sane type $\bar{t}_i$ of every player, we have $\bar{a}_i \in W^\infty(\tilde{G}_n)(\bar{t}_i)$. Let $T_i = \{\bar{t}_i, \hat{t}_i\}$, with $\tilde{u}_i^n(\cdot, \theta, \bar{t}_i) = u_i(\cdot, \theta)$ and $\tilde{u}_i^n(\cdot, \theta, \hat{t}_i) := 0$. Define $\kappa_i^n: T_i \to \Delta(\Theta \times T_{-i})$ by $\kappa_i^n(t_i)[\theta, \bar{t}_{-i}] = (1 - \frac{1}{n})\phi_i(\theta)$ and $\kappa_i^n(t_i)[\theta, \hat{t}_{-i}] = \frac{1}{n}\phi_i(\theta)$ for each t_i. Clearly, $\tilde{G}_n \overset{S}{\to} G$.

Step 1: $\bar{a}_i \in W(\tilde{G}_n)(\bar{t}_i)$.

Since $\bar{a}_i \in W(G)_i$, Lemma 2.2.2 yields $\bar{\mu}_i: \Theta \to \Delta^\circ(A_{-i})$ to which $\bar{a}_i$ is a best

response in G. We can regard $\bar{\mu}_i$ as a map from $\Theta \times T_{-i}$ to $\Delta^\circ(A_{-i})$ which does not depend on the opponents' types. Since $\text{marg}_\Theta \kappa_i^n(\bar{t}_i) = \phi_i$ and $\tilde{u}_i^n(\cdot, \theta, \bar{t}_i) = u_i(\cdot, \theta)$, it follows that for type $\bar{t}_i$, $\bar{a}_i$ is a best response to $\bar{\mu}_i$ in $\tilde{G}_n$. Hence, by the analog of Lemma 2.2.2 for elaborations, $\bar{a}_i \in W(\tilde{G}_n)(\bar{t}_i)$.

Step 2: $\bar{a}_i \in W^2(\tilde{G}_n)(\bar{t}_i)$.

It suffices to find a map $\lambda_i \colon \Theta \times T_{-i} \to \Delta^\circ(A_{-i})$ such that $\text{supp}(\lambda_i)(\theta, t_{-i}) = \prod_{j \neq i} W(\tilde{G}_n)(t_j)$ for all (θ, t_{-i}) and such that $\bar{a}_i$ is a best response to λ_i in $\tilde{G}_n$. Because $\kappa_i^n(\bar{t}_i)$ assigns positive probability only to opponent type profiles $\bar{t}_{-i}$ and $\hat{t}_{-i}$, it suffices to define $\bar{\lambda}_i = \lambda_i(\cdot, \bar{t}_{-i}) \colon \Theta \to \Delta^\circ(A_{-i})$ and $\hat{\lambda}_i = \lambda_i(\cdot, \hat{t}_{-i}) \colon \Theta \to \Delta^\circ(A_{-i})$, where we need that $\text{supp}(\bar{\lambda}_i)(\theta) = \prod_{j \neq i} W(\tilde{G}_n)(\bar{t}_j)$ and $\text{supp}(\hat{\lambda}_i)(\theta) = \prod_{j \neq i} W(\tilde{G}_n)(\hat{t}_j) = A_{-i}$ for all θ.

Since $\bar{a}_i \in SW(G)_i$, $\bar{a}_i$ is a best response to some $\mu_i \colon \Theta \to \Delta(A_{-i})$ with $\text{supp}(\mu_i)(\theta) \subseteq W(G)_{-i} \subseteq \prod_{j \neq i} W(\tilde{G}_n)(\bar{t}_j)$, where the last inclusion follows from Step 1. Let $\mu_i' \colon \Theta \to \Delta(A_{-i})$ be any map such that $\text{supp}(\mu_i')(\theta) = \prod_{j \neq i} W(\tilde{G}_n)(\bar{t}_j)$ for all θ. Let $\bar{\mu}_i \colon \Theta \to \Delta^\circ(A_{-i})$ be as in Step 1, and let $\delta := \min_{\theta, a_{-i}} \bar{\mu}_i(\theta)[a_{-i}] \in (0, 1)$. Let $\beta \in (0, 1)$ be sufficiently small (more precisely, we require $\beta < \frac{\delta}{(n-1)}$). Set $\bar{\lambda}_i = (1 - \beta)\mu_i + \beta\mu_i'$, so that for all θ, $\text{supp}(\bar{\lambda}_i(\theta)) = \text{supp}(\mu_i'(\theta)) = \prod_{j \neq i} W(\tilde{G}_n)(\bar{t}_j)$, as required. Set $\hat{\lambda}_i = \bar{\mu}_i + \beta(n - 1)(\bar{\mu}_i - \mu_i')$; then for all θ, $\hat{\lambda}_i(\theta) \in \Delta^\circ(A_{-i})$, as required, because β was chosen sufficiently small and because $\text{supp}(\bar{\mu}_i(\theta)) = A_{-i}$. Finally, by construction, we have for all θ that

$$\kappa_i^n(\bar{t}_i)[\theta, \bar{t}_{-i}]\bar{\lambda}_i(\theta) + \kappa_i^n(\bar{t}_i)[\theta, \hat{t}_{-i}]\hat{\lambda}_i(\theta)$$
$$= \phi_i(\theta)\left[\frac{(n-1)(1-\beta)}{n}\mu_i(\theta) + \frac{1 + \beta(n-1)}{n}\bar{\mu}_i(\theta)\right].$$

So, since $\tilde{u}_i^n(\cdot, \theta, \bar{t}_i) = u_i(\cdot, \theta)$ and since $\bar{a}_i$ is a best response to both $\bar{\mu}_i$ and μ_i in G, $\bar{a}_i$ is a best response for $\bar{t}_i$ to $\lambda_i = (\bar{\lambda}_i, \hat{\lambda}_i)$ in $\tilde{G}_n$.

Iterating Step 2, we obtain that $\bar{a}_i \in W^\infty(\tilde{G}_n)(\bar{t}_i)$, as claimed.

127

For the "if" direction, suppose $\tilde{G}_n \xrightarrow{S} G$ and consider $\bar{a} \in A$ such that $\bar{a}_i \in W^\infty(\tilde{G}_n)(\bar{t}_i)$ for the sane type $\bar{t}_i$ of each player i and for all n. We show that $\bar{a}_i \in S^\infty W(G)_i$ for all i.

Step 1: $\bar{a}_i \in W(G)_i$.

Since $\bar{a}_i \in W(\tilde{G}_n)(\bar{t}_i)$ for all n, there is a sequence of beliefs $\lambda_i^n \colon \Theta \times T_{-i} \to \Delta^\circ(A_{-i})$ to which $\bar{a}_i$ is a best response for $\bar{t}_i$ in $\tilde{G}_n$. That is,

$$\bar{a}_i \in \text{argmax}_{a_i' \in A_i} \sum_{\theta, t_{-i}, a_{-i}} \kappa_i^n(\bar{t}_i)[\theta, t_{-i}] \lambda_i^n(\theta, t_{-i})[a_{-i}] \tilde{u}_i^n(a_i', a_{-i}, \theta, \bar{t}_i).$$

Using conditions (i) and (ii) of the definition of strong convergence, we can rewrite this as

$$\bar{a}_i \in \text{argmax}_{a_i' \in A_i} \sum_\theta \phi_i(\theta) \sum_{a_{-i}} (\sum_{t_{-i}} \kappa_i^n(t_{-i}|\theta, \bar{t}_i) \lambda_i^n(\theta, t_{-i})[a_{-i}]) u_i(a_i', a_{-i}, \theta),$$

where $\kappa_i^n(t_{-i}|\theta, t_i) := \frac{\kappa_i^n(t_i)[\theta, t_{-i}]}{\text{marg}_\Theta \kappa_i^n(t_i)[\theta]}$ (so that by condition (i) of strong convergence, $\kappa_i^n(t_{-i}|\theta, \bar{t}_i) = \frac{\kappa_i^n(\bar{t}_i)[\theta, t_{-i}]}{\phi_i(\theta)}$). So, setting $\hat{\lambda}_i^n(\theta)[a_{-i}] = \sum_{t_{-i}} \kappa_i^n(t_{-i}|\theta, \bar{t}_i) \lambda_i^n(\theta, t_{-i})[a_{-i}]$ for all θ and a_{-i} yields beliefs $\hat{\lambda}_i^n \colon \Theta \to \Delta^\circ(A_{-i})$ to which $\bar{a}_i$ is a best response in G, where $\hat{\lambda}_i^n$ has range $\Delta^\circ(A_{-i})$ because λ_i^n does. By Lemma 2.2.2, $\bar{a}_i \in W(G)_i$, as claimed.

Step 2: $\bar{a}_i \in SW(G)_i$.

Since $\bar{a}_i \in W^2(\tilde{G}_n)(\bar{t}_i)$ for all n, there are beliefs $\mu_i^n \colon \Theta \times T_{-i} \to \Delta(A_{-i})$ such that $\text{supp}\,\mu_i^n(\theta, t_{-i}) = \prod_{j \neq i} W(\tilde{G}_n)(t_j)$ for all θ and t_{-i} and such that $\bar{a}_i$ is a best response for $\bar{t}_i$ to μ_i^n in $\tilde{G}_n$. As in Step 1, setting $\hat{\mu}_i^n(\theta)[a_{-i}] = \sum_{t_{-i}} \kappa_i^n(t_{-i}|\theta, \bar{t}_i) \mu_i^n(\theta, t_{-i})[a_{-i}]$ yields beliefs $\hat{\mu}_i^n \colon \Theta \to \Delta(A_{-i})$ to which $\bar{a}_i$ is a best response in G. Taking limits (passing to a subsequence if necessary) and using condition (iii) of the definition of strong convergence, we have that $\lim_n \hat{\mu}_i^n(\theta)[a_{-i}] = \lim_n \mu_i^n(\theta, \bar{t}_{-i})[a_{-i}]$. But $\text{supp}\,\mu_i^n(\theta, \bar{t}_{-i}) = \prod_{j \neq i} W(\tilde{G}_n)(\bar{t}_j) \subseteq W(G)_{-i}$, where the last inclusion follows from

128

Step 1. So, $\hat{\mu}_i := \lim_n \hat{\mu}_i^n$ is a map from Θ into $\Delta(W(G)_{-i})$. Moreover, by continuity $\bar{a}_i$ is a best response to $\hat{\mu}_i$ in G. Hence, by Lemma 2.2.2, $\bar{a}_i \in SW(G)_i$, as claimed.

Iterating the argument of Step 2, we conclude that $\bar{a}_i \in S^\infty W(G)_i$, as required. ■

Appendix C

Appendix to Chapter 3

C.1 Proofs

Proof of Theorem 3.2.4. We will prove the equivalence of (i) and (ii). The equivalence of (ii) and (iii) is immediate from the discussion in the main text.

<u>(i) $\Rightarrow$ (ii):</u> Suppose c admits an MTR $\langle v, \delta \rangle$. We prove that for any $x, y \in X$, xS^cy implies $v(x) > v(y)$—acyclicity of S^c then follows from acylicity of $>$ on $\mathbb{R}$. Suppose that xS^cy, so either xR^cy or xQ^cy. If xR^cy, then there is some A with $x \in c(A)$ and $y \in A \smallsetminus c(A)$. Then $\max_A v - v(x) \leq \delta(A) < \max_A v - v(y)$, so $v(x) > v(y)$. If xQ^cy, then there exists A and $z \in A$ such that $y \in A$, $z \in c(A)$ and $z \notin c(A \cup \{x\})$. Then $\max_A v - v(z) \leq \delta(A) \leq \delta(A \cup \{x\}) < \max_{A \cup \{x\}} v - v(z)$, so we must have $v(x) = \max_{A \cup \{x\}} v > \max_A v \geq v(y)$.

<u>(ii) $\Rightarrow$ (i):</u> Suppose S^c is acyclic. By Szpilrajn's (1930) Embedding Theorem, there is a weak order (in fact even a strict total order) on X extending S^c.[1] Thus, the set of

---

[1]Since X is finite, we can also directly construct such a strict total order T^c by means of a simple inductive argument on the cardinality of $|X|$: If $|X| = 1$, there is nothing to prove. If $|X| = n > 1$, pick any S^c-minimal element $a \in X$ (such an a exists, because S^c is acyclic and X is finite). Supposing that S^c has been extended to a strict total order T^c on $X \smallsetminus \{a\}$, setting xT^ca for all $x \in X \smallsetminus \{a\}$ gives a strict total order on X which extends S^c.

weak orders on X extending S^c is non-empty, finite (since X is finite), and partially ordered by set inclusion, so we can pick a minimal element W^c.[2] Let E^c denote the associated equivalence relation, i.e. xE^cy iff $\neg xW^cy$ and $\neg yW^cx$.

To construct the utility $v \colon X \to \mathbb{R}$, pick a single element from each equivalence class of E^c and enumerate these elements in W^c-increasing order by $x_0, x_1, \ldots, x_{n-1}$ (where n is the number of equivalence classes). Then set $v(y) := 2^i$ for any $y \in X$ with yE^cx_i. Clearly v represents W^c, which extends S^c, so for all $x, y \in X$ such that xS^cy, we have $v(x) > v(y)$. Note also that v satisfies the following "increasing differences" condition: For all $w, x, y, z \in X$,

$$\text{if } v(x) > v(y) \text{ and } v(x) > v(w), \text{ then } v(x) - v(w) > v(y) - v(z). \tag{C.1}$$

For every A, note that the set $\operatorname{argmax}_A v$ of v-maximal elements in A is contained in $c(A)$. Indeed, if not, then xR^cy for some $x \in c(A)$ and $y \in \operatorname{argmax}_A v$, implying $v(x) > v(y) = \max_A v$, a contradiction. So we can inductively define the threshold map $\delta \colon \mathcal{A} \to \mathbb{R}_+$ as follows:

- set $\delta(\{x\}) := 0$ for all $x \in X$;

- if $|A| \geq 2$ and $c(A) = \operatorname{argmax}_A v$, set $\delta(A) := \max_{B \subsetneq A} \delta(B)$;

- if $|A| \geq 2$ and $c(A) \supsetneq \operatorname{argmax}_A v$, set $\delta(A) := \max_{x,y \in c(A)}(v(x) - v(y)) = \max_A v - \min_{c(A)} v$.

To prove that v and δ constitute an MTR of c, we show by induction on the cardinality of A that

(a) $c(A) = \{x \in A : \max_A v - v(x) \leq \delta(A)\}$;

---

[2] The proof does not rely on W^c being minimal. The reason we choose a minimal weak order extension of S^c is so as not to attribute departures from rationality to the agent "unless absolutely necessary"—specifically, if c satisfies the Weak Axiom, then the construction in the proof ensures that $W^c = S^c$ and $\delta \equiv 0$.

(b) if $B \subseteq A$, then $\delta(B) \leq \delta(A)$;

(c) $\delta(A) \leq \max_{x,y \in A}(v(x) - v(y))$, and if $|A| \geq 2$ and $c(A) = \operatorname{argmax}_A v$, then
$$\delta(A) \leq \max_{x,y \in A \smallsetminus \operatorname{argmax}_A v}((v(x) - v(y))).$$

If $|A| = 1$, then (a), (b) and (c) obviously hold. If $|A| \geq 2$ and (a), (b) and (c) hold for sets of cardinality less than $|A|$, then we consider separately the cases where $c(A) = \operatorname{argmax}_A v$ or $c(A) \supsetneq \operatorname{argmax}_A v$.

Case 1: First, suppose that $c(A) = \operatorname{argmax}_A v$. Then $\delta(A) := \max_{B \subsetneq A} \delta(B) = \delta(B_0)$, say. Thus, property (b) is immediate. If $|B_0| = 1$, then $\delta(A) = \delta(B_0) = 0$ and properties (a) and (c) are equally obvious. So suppose that $|B_0| \geq 2$. Since $\delta(B_0) \geq 0$ by inductive hypothesis, we certainly have $c(A) = \operatorname{argmax}_A v \subseteq \{x \in A : \max_A v - v(x) \leq \delta(B_0)\}$. To prove (a), we must also show that if $w \in A \smallsetminus \operatorname{argmax}_A v$, then $\max_A v - v(w) > \delta(B_0)$. There are two cases to consider:

Suppose first that $\max_A v = \max_{B_0} v$. By (b) applied to sets of cardinality less than $|A|$, we can assume that $|B_0| = |A| - 1$, so $B_0 \cup \{z\} = A$ for some $z \in A$. Then we must have $c(B_0) = \operatorname{argmax}_{B_0} v$: Indeed, if not, we have $c(B_0) \not\subseteq c(B_0 \cup \{z\}) = c(A) = \operatorname{argmax}_A v$. But then $z Q^c y$ for all $y \in B_0$, so $v(z) > \max_{B_0} v$, contradicting $\max_A v = \max_{B_0} v$. Thus, by inductive hypothesis and since $|B_0| \geq 2$, the second part of (c) applies to B_0 and yields $\delta(B_0) \leq \max_{x,y \in B_0 \smallsetminus \operatorname{argmax}_{B_0} v}(v(x) - v(y)) = \max_{x,y \in B_0 \smallsetminus \operatorname{argmax}_A v}(v(x) - v(y))$.

Suppose now that $\max_A v > \max_{B_0} v$. Then by (c) applied to B_0, we again get $\delta(B_0) \leq \max_{x,y \in B_0}(v(x) - v(y)) = \max_{x,y \in B_0 \smallsetminus \operatorname{argmax}_A v}(v(x) - v(y))$.

So in either case, $\delta(B_0) \leq \max_{x,y \in B_0 \smallsetminus \operatorname{argmax}_A v}(v(x) - v(y))$. But note that the increasing differences property (C.1) implies that $\max_{x,y \in B_0 \smallsetminus \operatorname{argmax}_A v}(v(x) - v(y)) < \max_A v - v(w)$ for all $w \in A \smallsetminus \operatorname{argmax}_A v$. Thus, for all $w \in A \smallsetminus \operatorname{argmax}_A v$, $\max_A v - v(w) > \delta(B_0)$, establishing (a). Finally, property (c) holds for A, be-

132

cause by the above we have $\delta(A) = \delta(B_0) \leq \max_{x,y \in B_0 \smallsetminus \mathrm{argmax}_A v}(v(x) - v(y)) \leq \max_{x,y \in A \smallsetminus \mathrm{argmax}_A v}(v(x) - v(y))$.

<u>Case 2</u>: Now suppose that $c(A) \supsetneq \mathrm{argmax}_A v$. Then $\delta(A) := \max_A v - \min_{c(A)} v > 0$. Thus, $c(A) \subseteq \{x \in A : \max_A v - v(x) \leq \delta(A)\}$ is immediate. Conversely, if $z \in A \smallsetminus c(A)$, then $yR^c z$ for any $y \in c(A)$, so $v(z) < \min_{c(A)} v$, whence $\max_A v - v(z) > \delta(A)$. This proves (a). Property (c) is immediate by definition of $\delta(A)$. Finally, to prove (b), consider $B \subseteq A$. If $|B| = 1$, then $\delta(B) = 0 < \delta(A)$. So suppose $|B| \geq 2$. Again there are two cases:

First suppose that $\max_B v = \max_A v$. By inductive hypothesis, we can assume $|B| = |A| - 1$, say $B \cup \{z\} = A$ for some $z \in A$. Then $\max_B v = \max_A v \geq v(z)$, so $\neg z Q^c y$ for some $y \in B$, whence $c(B) \subseteq c(B \cup \{z\}) = c(A)$. Thus, $\min_{c(B)} v \geq \min_{c(A)} v$. So if $B \supsetneq \mathrm{argmax}_B v$, then $\delta(B) := \max_B v - \min_{c(B)} v \leq \max_A v - \min_{c(A)} v = \delta(A)$, as required. On the other hand, if $c(B) = \mathrm{argmax}_B v$, then the second part of (c) applied to B yields $\delta(B) \leq \max_{x,y \in B \smallsetminus \mathrm{argmax}_B v}(v(x) - v(y)) < \max_A v - \min_{c(A)} v$ by the increasing differences property (C.1). So again $\delta(B) \leq \delta(A)$.

On the other hand, suppose $\max_B v < \max_A v$. Then part (c) applied to B yields $\delta(B) \leq \max_{x,y \in B}(v(x) - v(y)) < \max_A v - \min_{c(A)} v$ by the increasing differences property (C.1). So again $\delta(B) \leq \delta(A)$, completing the proof. ∎

Proof of Lemma 3.3.3.

(i): Suppose first that c admits a CTR $\langle v, \delta \rangle$. Then Occasional Optimality holds by Theorem 3.2.4. To prove Contraction, suppose that $x \in c(B) \cap A$ with $A \subseteq B$. Then $\max_B v - v(x) \leq \delta$ and $\max_B v \geq \max_A v$, so $\max_A v - v(x) \leq \delta$, whence $x \in c(A)$.

For the converse, suppose that c satisfies Occasional Optimality and Contraction. By Theorem 3.2.4, c admits an MTR, $\langle v, \delta \rangle$, say. By Luce (1956), c admits a CTR

if and only if c is generated by a semiorder (a simple proof can be found in Aleskerov *et al.* (2007)).[3] Hence, it is sufficient to prove that c is generated by P^c and that P^c is a semiorder,[4] where P^c is as in Definition 3.2.2. To see that P^c generates c, fix B. If $x \in c(B)$ and $y \in B$, then applying Contraction with $A = \{x, y\}$ yields $x \in c(\{x, y\})$, so $\neg y P^c x$. Conversely, if $x \in B \smallsetminus c(B)$, then there exists $y \in B$ such that $v(y) - v(x) > \delta(B) \geq \delta(\{x, y\})$. So $c(\{x, y\}) = \{y\}$, whence $y P^c x$. Hence, $c(B) = \{x \in B : \forall y \in B \ \neg y P^c x\}$, i.e. P^c generates c. We now show that P^c is a semiorder: Irreflexivity is clear. If $x P^c y$ and $y P^c z$, then since P^c generates c and c is non-empty, we must have $c(\{x, y, z\}) = \{x\}$. Thus, $v(x) - v(z) > \delta(\{x, y, z\}) \geq \delta(\{x, z\}) \geq 0$, so that $x \neq z$ and $c(\{x, z\}) = \{x\}$, i.e. $x P^c z$. This shows that P^c is transitive. To prove that P^c is semitransitive, suppose to the contrary that $x P^c y$ and $y P^c z$, but $\neg x P^c w$ and $\neg w P^c z$. $x P^c y$ and $\neg x P^c w$ together with transitivity of P^c implies that $\neg y P^c w$, so $w \in c(\{x, y, w\})$ and $y \notin c(\{x, y, w\})$, whence $w S^c y$. But $y P^c z$ and $\neg w P^c z$ implies $z \in c(\{w, z\})$ and $z \notin c(\{w, z, y\})$, so $y Q^c w$, whence $y S^c w$. This contradicts acyclicity of S^c (which holds by Theorem 3.2.4). Finally, to prove that P^c satisfies the interval order condition, suppose to the contrary that $x P^c y$ and $w P^c z$, but $\neg x P^c z$ and $\neg w P^c y$. Then by transitivity of P^c, we also have $\neg y P^c z$ and $\neg z P^c y$. Hence, $z \in c(\{x, y, z\})$ and $y \notin c(\{x, y, z\})$, so $z S^c y$; also, $y \in c(\{w, y, z\})$ and $z \notin c(\{w, y, z\})$, so $y S^c z$. This again contradicts acylicity of S^c.

(ii): Tyson (2008) proves that c admits an ESR if and only if P^c is acyclic and c satisfies Strong Expansion.[5] By Theorem 3.2.4, Occasional Optimality is equivalent to acyclicity of S^c, which implies acyclicity of P^c since $P^c \subseteq S^c$. So it suffices to prove that if c satisfies Strong Expansion, then $S^c \subseteq P^c$. Suppose that $x S^c y$, so either $x R^c y$

---

[3]Cf. Aleskerov *et al.* (2007) Theorem 3.2 (p. 66).

[4]Recall the definition of a semiorder in footnote 7.

[5]Cf. Tyson (2008) Theorem 5B (p. 59). Note that Tyson refers to acyclicity of P^c as Base Acyclicity.

or xQ^cy. If xR^cy, then there is B such that $x \in c(B)$ and $y \in B \setminus c(B)$. Applying Strong Expansion with $A = \{x,y\}$ gives $c(\{x,y\}) = \{x\}$, so xP^cy. If xQ^cy, then there is A such that $y \in A$ and $c(A) \not\subseteq c(A \cup \{x\})$. Applying Strong Expansion with $B = A \cup \{x\}$ yields $c(A \cup \{x\}) = \{x\}$. Thus xR^cy, and we are back in the previous case. ∎

C.2 Separating Examples for Section 3.3

Example C.2.1 (SR $\not\subseteq$ MTR). Let c be the choice correspondence on $X = \{x,y,z\}$ with satisficing representation $\langle u, \theta \rangle$, where $u(x) = 1$, $u(y) = 2$, $u(z) = 3$, $\theta(\{a\}) = \theta(\{a,b\}) = 1$ for all a, b, and $\theta(\{x,y,z\}) = 3$. Then $c(\{x,z\}) = \{x,z\}$ and $c(\{x,y,z\}) = \{z\}$, so yQ^cz and zR^cy, producing the cycle yS^czS^cy. Hence, by Theorem 3.2.4, c does not admit an MTR.

Example C.2.2 (CTR $\not\subseteq$ ESR). Let c be the choice correspondence on $X = \{x,y,z\}$ induced by the CTR $\langle v, \delta \rangle$, where $v(z) = 3$, $v(y) = 2$, $v(x) = 1$, $\delta = 1$. Then $c(\{x,y\}) = \{x,y\}$ and $c(\{x,y,z\}) = \{y,z\}$. So setting $A = \{x,y\}$ and $B = \{x,y,z\}$ yields a violation of Strong Expansion. Thus, by Lemma 3.3.3 c does not admit an ESR.

Example C.2.3 (ESR $\not\subseteq$ CTR). Let c on $X = \{x,y,z\}$ be given by $c(\{x,y\}) = \{x,y\}$, $c(\{x,z\}) = \{z\} = c(\{y,z\})$, $c(X) = X$. Setting $A = \{x,z\}$, $B = X$ yields a violation of Contraction, so by Lemma 3.3.3 c does not admit a CTR. But $\langle u, \theta \rangle$ with $u(z) = 3$, $u(y) = 2 = u(x)$, $\theta(X) = 1$, $\theta(\{y,z\}) = \theta(\{x,z\}) = \theta(\{z\}) = 3$, $\theta(A) = 2$ for all other A is an ESR of c.

Example C.2.4 (MTR $\not\subseteq$ ESR $\cup$ CTR). Let c be the choice correspondence on $X = \{x,y,z\}$ induced by the MTR $\langle v, \delta \rangle$, where $v(z) = 4$, $v(y) = 2$, $v(x) = 1$, $\delta(\{a\})$,

$\delta(\{a,b\}) = 1$ for all $a, b \in X$ with $a \neq b$, and $\delta(X) = 2$. Then $c(\{y,z\}) = \{z\}$, $c(\{x,y\}) = \{x,y\}$, and $c(\{x,y,z\}) = \{y,z\}$. So setting $A = \{y,z\}$ and $B = \{x,y,z\}$ yields a violation of Contraction, whence c does not admit a CTR. And setting $A = \{x,y\}$ and $B = \{x,y,z\}$ yields a violation of Strong Expansion, whence c does not admit an ESR.