CERGE
Center for Economics Research and Graduate Education
Charles University

Essays on Information Economics

Ludmila Matysková

Dissertation

Prague, July 2018

Ludmila Matysková

Essays on Information Economics

Dissertation

Prague, July 2018

Dissertation Committee

Jakub Steiner (CERGE-EI; chair)

Filip Matějka (CERGE-EI)

Avner Shaked (CERGE-EI)

Jan Zápal (CERGE-EI)

Referees

Mark Dean (Columbia University)

Emir Kamenica (University of Chicago)

Contents

Abstract

In the first chapter, we study the effect on the economy of platforms for online consumer reviews. Consumer reviews may have perverse effects, including delays in the adoption of new products of unknown quality when consumers are boundedly rational. When consumers fail to take into account that past reviewers self-select into purchases, a monopolist may manipulate the posterior beliefs of consumers who observe the reviews, because the product price determines the self-selection bias. The monopolist will charge a relatively high price because the positive selection of the early adopters increases the quality reported in the reviews.

In the second chapter, we study a game between a sender and a receiver in a framework of Bayesian persuasion. A sender choosing a signal to be disclosed to a receiver can often influence the receiver's actions. Is persuasion more difficult when the receiver has additional information sources? Does the receiver benefit from having the additional sources? We extend a Bayesian persuasion model to a receiver's acquisition of costly information. The game can be solved as a standard Bayesian persuasion model under an additional constraint: the receiver never gathers her own costly information. The 'threat' of learning hurts the sender. However, the resulting outcome can also be worse for the receiver. We further propose a new solution method which does not rely on concavification, and which is also applicable to standard Bayesian persuasion.

The last chapter focuses on habit formation. Using a laboratory experiment, we study whether habits form mechanically as a consequence of past decisions, or whether they form only if they help the person to improve her decision-making in a complex, stochastic environment. We define habits as action inertia and study them in multiple-period binary tasks with serially correlated incentives. For a fixed sequence of decision problems, habits form when (i) the subjects do not observe past payoff states, and they do not form when (ii) the past states are observed. Since the past actions contain useful information in treatment (i) but not in (ii), the data suggests that habits are a functional adaptation.

Abstract

V první kapitole zkoumáme, jaký efekt mají platformy umožňující spotřebitelské recenze na ekonomiku. Když jsou spotřebitelé omezeně racionální, spotřebitelské recenze mohou mít zvrácené účinky zahrnující zpoždění v ujímání se nových produktů neznámé kvality. Pokud spotřebitelé opomíjejí, že u recenzentů došlo k samoselekci ohledně zakoupení či nezakoupení produktu, monopolista může manipulovat přesvědčení spotřebitelů ohledně kvality produktu aktualizované po přečtení recenzí, protože cena produktu určuje zkreslení dané samoselekcí (tzv. self-selection bias). Monopolista nastaví relativně vysokou cenu, protože pozitivní samoselekce u počátečních kupců zvýší reportovanou kvalitu v jejich následných recenzích.

Ve druhé kapitole studujeme hru mezi odesílatelem a příjemcem informací v rámci modelu Bayesovského přesvědčování. Odesílatel, který rozhoduje o tom, jaké informace se dostanou k příjemci, může často ovlivnit příjemcovy následné akce. Je přesvědčování těžší, pokud příjemce má i své vlastní zdroje informací? Přinese možnost získávání dalších informací i z vlastních zdrojů prospěch příjemci? V této studii rozšiřujeme model Bayesovského přesvědčování o možnost, že příjemce si může za určitou cenu navíc vyhledat dodatečné informace z vlastních zdrojů. Řešení této hry můžeme získat jako řešení standardního modelu Bayesovského přesvědčování za dodatečné podmínky, která musí platit: příjemce se nikdy nerozhodne získat další (nákladné) informace ze svých zdrojů. Ona 'hrozba' potenciálního dodatečného učení však snižuje užitek odesílatele. Nicméně, celkový výsledek může být horší i pro příjemce, který tak nutně nemusí mít prospěch z možnosti vyhledat své vlastní nákladné informace. Dále nabízíme alternativní postup pro řešení daného modelu, který se neodvíjí od standardní metody tzv. konkavifikace, a který je aplikovatelný i na standardní model Bayesovského přesvědčování.

Poslední kapitola se zaměřuje na formování návyků. Pomocí laboratorního experimentu zkoumáme, zda se návyky vytvoří mechanicky jakožto důsledek předchozích rozhodnutí, nebo zda se vytvoří pouze v případě, když mohou člověku pomoci zlepšit jeho rozhodování v komplexním, stochastickém prostředí. Návyky definujeme jako setrvačnost v

akcích a studujeme je v binárních úlohách o více periodách s podněty, které jsou sériově korelované. Pro danou sekvenci rozhodovacích problémů se návyk (i) vytvoří tehdy, když jsou minulé stavy světa pozorovány, zatímco (ii) se nevytvoří tehdy, když minulé stavy světa pozorovány nejsou. Vzhledem k tomu, že minulé akce obsahují užitečnou informaci pouze v případě (i), ale ne v případě (ii), tato data naznačují, že vytváření návyků je funkční adaptací.

Acknowledgments

I am extremely grateful to my supervisor, Jakub Steiner, for his guidance throughout my entire PhD studies, for his patience and constructive criticism, for making time for me whenever I needed it, and for his encouragement when I needed it most. I am grateful to other members of my dissertation committee, Filip Matějka, Avner Shaked, and Jan Zápal, with whom I could discuss any problems that have emerged along the way.

I am indebted to many other inspiring people, especially Mark Dean and Margaret Meyer who both took me under their wings when I left my home institution to explore their departments, and to fellow students Vladimír Novák and Andrei Maatvenko for simply being there for me during various ups and downs. Deborah Nováková and Gray Krueger carefully edited all three chapters, considerably enhancing their readability. I want to thank our SAO who helped me tremendously with many administrative issues. In the last stage, I benefited from insightful comments of Mark Dean and Emir Kamenica, who refereed the thesis. I also want to thank three institutions: CERGE-EI for enabling me to obtain an education in Economics in my home town, and Columbia University and Nuffield College of the University of Oxford for their generous hospitality during my stays.

I am thankful to my family, especially to my father, whom I successfully bothered about my research for so long that he did not have another choice but to engage in lengthy discussions.

Financial support from the H2020-MSCA-RISE project GEMCLIME-2020 GA No. 681228, Charles University project GA UK No. 178715, the ERC-2015-STG (H2020) project no. 678081, and Nadace "Nadání Josefa, Marie a Zdeňky Hlávkových" is gratefully acknowledged.

All errors remaining in this text are my responsibility.

Introduction

The unifying topic of all three chapters in this dissertation is acquiring and processing information, especially in strategic situations. In the first two chapters, I theoretically study strategic manipulation of truthful information in two different settings. In the last chapter, I experimentally examine information processing in a non-strategic setting. The first chapter concerns a monopolist who indirectly manipulates information about his product observed in online reviews via pricing decisions: his initial price determines the composition of actual buyers, hence the composition of reviewers. The second chapter studies a game between a sender and a receiver. The sender controls what information is provided to the uninformed receiver while taking into account that the receiver can later potentially gather her own information (at a cost). The last chapter is a laboratory experiment, in which we study habit formation in a single-agent dynamic information acquisition task.

In the first chapter, "Manipulation of Cursed Beliefs in Online Reviews " (joint with Jan Šípek), we study the effects of platforms for online reviews of new products on firm's pricing decisions. We demonstrate that such platforms can have negative implications for the society, such as slowing down the diffusion of new products, because the firm has an incentive to set a high initial price. When potential buyers have heterogeneous preferences, the initial price of the product affects the composition of actual buyers. Since only the actual buyers are also the reviewers, the selection bias translates to a bias in reviews. When consumers are boundedly rational and ignore the self-selection aspect (which has been empirically documented and which we model via Eyster and Rabin's (2005) concept of the 'cursed equilibrium'), the initial price is an indirect instrument to bias posterior beliefs. This new strategic aspect of the price, created by the existence of the platforms, then drives up the initial price of new products.

In the second chapter, "Bayesian Persuasion with Costly Information Acquisition", I combine two current streams of literature in Information Economics, Bayesian persuasion (Kamenica and Gentzkow 2011) and Rational inattention (Sims 2003). Bayesian persuasion studies a game in which a sender discloses information to a receiver in order to persuade her to change her original action. Rational inattention is a theory of endogenous costly information acquisition. I extend Bayesian persuasion to a receiver's endogenous information acquisition under an entropy-based cost, commonly used in rational inattention. I show that the possibility of the receiver to obtain additional information has a disciplinary effect on the sender (weakly decreasing his expected equilibrium payoff), for whom the persuasion becomes more difficult. However, the overall outcome can be worse also for the receiver, as the sender can prefer to strategically disclose significantly less information when facing the threat of additional learning. I exploit the (technical) similarity in both theories that allows for tractability. I propose a new method showing that the potentially very complicated problem can be simplified to a search through a finite set of information strategies, which is characterized by linear conditions. The new approach is also applicable to standard Bayesian persuasion and can simplify, sometimes dramatically, the search for a sender's optimal disclosure strategy (as opposed to the standard concavification technique usually used to solve these models).

In the final chapter, "Habit Formation: An Experimental Study" (joint with Keh Sun), we study why people form habits, tendencies to excessively repeat previous actions. One explanation argues that habits are an intrinsic part of preferences—people have (mental) switching costs (e.g. Abel (1990); Carroll, Overland, and Weil (2000); Constantinides (1990)). Alternatively, habits can arise as an adaptation in a complex, stochastic environment, rather than being a hard-wired feature. An agent who learns about a changing environment with positively serially correlated payoff states can save on learning costs by repeating their earlier actions. Such an agent then, *in some circumstances*, resembles behavior of an agent for whom habits are driven by preferences (Steiner, Stewart, and Matějka 2017). Habit formation is then information driven and varies with specific features of the environment. However, it is invariant to these variations under fixed switching costs. Recognizing whether habits are preference- or information-based is important for the evaluation of policy intervention, such as the impact of monetary policies during economic crises. We experimentally test these two different theories by observing how subjects take actions in binary perceptual decision problems over two periods. We consider a stochastic environment with imperfect stochastic information. We define habit as action inertia that is not explained by variations in optimal actions. For a fixed sequence of decision problems, habits form when (i) the subjects do not observe past payoff states, and they do not form when (ii) the past states are observed. Since the past actions contain useful information in treatment (i) but not in (ii), the data suggests that habits are a functional adaptation.

Chapter 1

Manipulation of Cursed Beliefs in Online Reviews

Co-authored with Jan Šípek (CERGE-EI).

1.1 Introduction

This paper challenges the general desirability of online reviews for experience products for the society. When consumers have heterogeneous preferences, each posted review reflects both the product's quality and the reviewer's idiosyncratic taste. With self-selection to buying decisions, reviews become a biased signal on unknown quality. If consumers do not correct for the self-selection, a monopolist may manipulate the posterior beliefs about the quality, because the product price determines self-selection bias. The monopolist faces a trade-off between higher demand today and higher posterior beliefs tomorrow: a higher price today lowers current demand, but intensifies self-selection to buying (only people with high enough idiosyncratic expectation/taste buy), resulting in higher posterior beliefs tomorrow. We show that the monopolist charges a *strictly higher* initial price compared to myopic initial price in a setting without reviews, in order to increase bias in the posterior beliefs.

We present a 2-period model, where a monopolist dynamically prices a new experience product of a quality initially unknown to the monopolist and consumers. Updating of beliefs occurs by observing reviews that are truthfully and automatically posted by all past buyers. While the monopolist is rational, consumers have heterogeneous preferences and are boundedly rational—they do not take into account that reviewers self-select to purchasing decisions—as in the cursed equilibrium of Eyster and Rabin (2005). The bounded rationality is motivated by Li and Hitt's (2008) empirical analysis of online book reviews: this suggests that consumers update beliefs *as if* the reviews reflected their own preferences, despite preference heterogeneity. In cursed equilibrium, agents correctly predict the distribution of actions, but neglect how these actions are correlated with other players' private information. Laboratory experiments (Parlour, Prasnikar, and Rajan 2007) as well as real world examples (Brown, Camerer, and Lovallo 2012) show behavior consistent with cursed equilibrium. We offer another application of such boundedly-rational behavior.

This paper is related to literature on the strategic manipulation of the social learning (SL) process. Liu and Schiraldi (2012), Bhalla (2013), Bose et al. (2006), and Bose et al. (2008) analyze a setting where consumers have homogeneous preferences and private information about the unknown value of the product. Pricing decisions may then screen private information. In contrast, our consumers hold identical ex-ante information about quality, but have heterogeneous preferences. The learning does not come from observing purchasing decisions (revealing ex-ante private information) but from observing reviews (revealing ex-post satisfaction). The pricing decision then affects bias in the learning.

In a specific setting of online reviews, Crapis et al. (2016), and Ifrach, Maglaras, and Scarsini (2014) characterize equilibrium dynamics with rational and boundedly rational consumers, respectively. Rather than asking what the dynamics of the learning process are, we ask what distortions in pricing decisions are induced by the existence of reviews. Ifrach, Maglaras, and Scarsini (2014) show that 'the optimal dynamic pricing strategy charges a *lower* price than the corresponding myopic policy, which ignores the effect of pricing on the SL process', since a low price speeds up learning. In contrast, we suggest that the review system also generates incentives for charging a *higher* initial price than the corresponding initial myopic price.

In a similar setting to ours, Papanastasiou, Bakshi, and Savva (2014) show that a monopolist, constrained to charging a fixed price, may deliberately under-supply the early demand in order to increase bias in the consumers' posterior beliefs. However, no optimality of such under-supplies remains under dynamic pricing. We analyze a similarly unconstrained model.

In the framework of online reviews, firms may post fake reviews to bias consumers' beliefs (Dellarocas 2006; Mayzlin, Dover, and Chevalier 2014). We show that even if we abstract from such practices, the monopolist is still able to manipulate consumers' beliefs using another tool: the price.

1.2 Model

A monopolist is selling an experience product of unknown quality, with zero marginal costs, over two periods. The monopolist maximizes total undiscounted expected profit by setting price $p_t \geq 0$ in period $t \in \{1, 2\}$.

Each period t, a continuum of 1-period-lived consumers of mass one enters the market (unsatisfied demand is not carried over to the next period). Each consumer i purchases the product whenever her expected utility $\mathbb{E}[u(a, \theta_i, p_t)|\theta_i, p_t]$ from purchase is nonnegative, where

$$u(q, \theta_i, p_t) = q + \theta_i - p_t, \tag{1.1}$$

q is an unknown *common quality*, and θ_i is a privately known *idiosyncratic taste*. Upon purchase, the consumer learns q and automatically posts her *experienced quality* $q + \theta_i$ in a public review.

Information: q is initially unknown to the monopolist and consumers. Common prior belief is $q \sim N(\mu, \sigma^2)$. Privately known taste parameters θ_i are iid random variables from uniform distribution on $[-\varepsilon, \varepsilon]$, $\varepsilon > 0$. We denote the random variables and their realizations by the same symbol. In period 2, all agents observe the average of past-period reviews $\overline{r} = \mathbb{E}_\theta[q + \theta|\mu + \theta - p_1 \geq 0]$, if there are any, as well as p_1 before taking action.

Consumers have the *cursed beliefs* of Eyster and Rabin (2005): each consumer correctly predicts the distribution of other consumers' actions, but do not take into account how these actions are correlated with idiosyncratic tastes. They incorrectly believe that others purchase the product randomly (with actual unconditional probability of purchase) irrespective of their taste, rather than in a way specified by their taste-contingent strategy[1]. The monopolist is rational and aware of the bounded rationality of consumers. Note that since all quality-relevant information is public, there is no signaling issue.

Timing: Nature chooses q. Consumers enter the market and privately learn their θ_i's. The monopolist sets p_1. Consumers make their purchasing decisions. Anyone buying learns q and posts a review $q + \theta_i$. Period 1 ends and all consumers leave the market. New consumers enter the market and privately learn their θ_i's. The average review $\overline{r}$ from the previous period and p_1 become public. The monopolist sets p_2. Consumers make purchasing decisions. The game ends.

1.2.1 Cursed posterior beliefs

Self-selection in period 1

If $p_1 \leq \mu+\varepsilon$, consumer i from period 1 buys iff $\theta_i \geq \hat{\theta}_1(p_1)$, where the threshold consumer $\widehat{\theta}_1(p_1)$ is

$$\widehat{\theta}_1(p_1) = \begin{cases} -\varepsilon & p_1 \leq \mu - \varepsilon, \\ p_1 - \mu & \mu - \varepsilon < p_1 \leq \mu + \varepsilon. \end{cases} \tag{1.2}$$

If $p_1 > \mu + \varepsilon$, nobody purchases the product.

Given the realization of quality q, the average review $\overline{r}(q, \widehat{\theta}_1(p_1))$, if there is any, is

$$\overline{r}(q, \widehat{\theta}_1(p_1)) = q + \mathbb{E}[\theta | \theta \geq \widehat{\theta}_1(p_1)]. \tag{1.3}$$

Rational (Bayesian) updating

Rational agents would account for self-selection. If past reviews exist, they 'de-bias' them appropriately, and recover the realized product quality by taking the inverse of (1.3). Ot-

[1] Eyster and Rabin (2005) parametrize the 'cursedness' by $\chi \in [0, 1]$. Agents believe others use a type-contingent strategy with probability $1 - \chi$ and act independently of their type otherwise. For simplicity, we set $\chi = 1$, but our results hold $\forall \chi > 0$.

herwise, the posterior beliefs coincide with priors (since there is no new information).

Cursed updating

Consumers mistakenly believe that reviews reflect the opinions of an unbiased random sample of the population, i.e., that $\overline{r}(q, \widehat{\theta}_1(p_1)) = \mathbb{E}[q+\theta] = q$, even though it is determined by (1.3). If past reviews exist, the cursed posterior belief is

$$q_c(q, \widehat{\theta}_1(p_1)) = \overline{r}(q, \widehat{\theta}_1(p_1)) \overset{(1.3)}{=} q + \underbrace{\mathbb{E}[\theta | \theta \geq \widehat{\theta}_1(p_1)]}_{bias\ (\geq 0)}. \tag{1.4}$$

Otherwise, the posterior beliefs coincide with priors (since there is no new information).

Since the monopolist is rational, her posterior beliefs are unbiased. The consumers' cursed posterior beliefs are unbiased only if the demand in period 1 was unity or zero. The monopolist endogenously manipulates the size of the bias through price p_1 (determining the intensity of the self-selection).

1.3 Inflation of price

Definition 1. An *equilibrium* is the monopolist's pair of prices $\{p_1^*, p_2^*(q, p_1^*)\}$ and consumers' pair of threshold tastes $\{\widehat{\theta}_1^*(p_1^*), \widehat{\theta}_2^*(q, p_1^*)\}$ such that:

1. In period 2, given p_1^*, q, and
   (a) given $p_2^*(q, p_1^*)$, a consumer i purchases the product iff $\theta_i \geq \widehat{\theta}_2^*(q, p_1^*)$;
   (b) given $\widehat{\theta}_2^*(q, p_1^*)$, $p_2^*(q, p_1^*)$ maximizes the second-period monopolist's profit.
   (c) Consumers' posterior belief is given by (1.4) for $p_1 = p_1^*$.
2. In period 1,
   (a) given p_1^*, a consumer i purchases the product iff $\theta_i \geq \widehat{\theta}_1^*(p_1^*)$;
   (b) given a second-period subgame equilibrium for any q (point 1.) and given $\widehat{\theta}_1^*(p_1)$, p_1^* maximizes the total undiscounted expected profit.

If $\mu \leq -\varepsilon$, nobody buys at any positive price. If $3\varepsilon \leq \mu$, heterogeneity is relatively small, compared to μ. The monopolist may find it optimal to satisfy the whole demand of period 1, yielding no bias in q_c. A1 ensures a nontrivial case, where the monopolist always sells to a strict subset of consumers in period 1 (heterogeneity matters relatively).

Assumption 1. *(**A1**): $-\varepsilon < \mu < 3\varepsilon$.*

Comparing p_1^* to its counterpart charged in a setting without review system, we establish the main result.

Proposition 1. *Let A1 hold. Then the optimal first-period price is strictly higher in the setting with the reviews than without them.*

Proof. Appendix 1.A. □

1.4 Conclusion

This paper illustrates that the presence of consumer reviews may generate undesirable incentives. A monopolist selling an experience good can manipulate anticipations of the product's quality, even with truthful reviews. Consumers with the heterogeneous preferences and cursed beliefs of Eyster and Rabin (2005) fail to take into account that past consumers self-select themselves into purchasing decisions, leading to excessively high anticipation of quality. Since the higher the price is, the higher is the self-selection bias, the monopolist can exacerbate this error by increasing the price. The monopolist charges a higher initial price than the corresponding myopic price which ignores the effect of pricing on learning. The presence of reviews may thus slow down the diffusion of new products. Addressing the motivation to post reviews, or the reliability of reviews based on their sample size (currently absent due to the continuum of consumers assumption) are some possible extensions of the model.

1.A Appendix

Proof of Proposition 1

Proof. From the ex-ante point of view, $q \sim N(\mu, \sigma^2)$. The monopolist's maximization problem is

$$\max_{p_1 \in [\mu-\varepsilon, \mu+\varepsilon)} p_1 \frac{\varepsilon - (p_1 - \mu)}{2\varepsilon} + \mathbb{E}[\Pi_2^*(q_c(q, \widehat{\theta}_1(p_1)))], \tag{1.5}$$

where

$$\Pi_2^*(q_c) = \begin{cases} 0 & q_c < -\varepsilon, \\ \left(\frac{1}{2\varepsilon}\right)\left(\frac{q_c+\varepsilon}{2}\right)^2 & -\varepsilon \leq q_c \leq 3\varepsilon, \\ q_c - \varepsilon & q_c > 3\varepsilon \end{cases} \tag{1.6}$$

is the equilibrium second-period profit (found by backward induction), and the expectation is taken over $q_c(q, \widehat{\theta}_1(p_1))$ given by (1.4). Given $p_1 \in [\mu-\varepsilon, \mu+\varepsilon)$ and the distribution assumptions[2], $q_c(q, \widehat{\theta}_1(p_1)) \sim N(\mu_u(p_1), \sigma^2)$ ex-ante, where

$$\mu_u(p_1) = \mu + \underbrace{\frac{1}{2}(p_1 - \mu + \varepsilon)}_{\text{bias}}. \tag{1.7}$$

The slope of (1.5) is

$$\frac{\mu - 2p_1 + \varepsilon}{2\varepsilon} + \frac{\mathrm{d}}{\mathrm{d}p_1}\mathbb{E}_{q_c}[\Pi_2^*(q_c(q, \widehat{\theta}_1(p_1)))]. \tag{1.8}$$

We show that the slope (1.8) is strictly positive for $p_1 \in [\mu - \varepsilon, \frac{\mu+\varepsilon}{2}]$, which implies that the solution to (1.5) is $p_1^* > \frac{\mu+\varepsilon}{2}$.

A1 ensures that $\frac{\mu+\varepsilon}{2} \in [\mu - \varepsilon, \mu + \varepsilon)$. Furthermore,

1. $\frac{\mu - 2p_1+\varepsilon}{2\varepsilon} > 0$ for $p_1 < \frac{\mu+\varepsilon}{2}$ and equals zero at $p_1 = \frac{\mu+\varepsilon}{2}$.

2. $\frac{\mathrm{d}}{\mathrm{d}p_1}\mathbb{E}_{q_c}[\Pi_2^*(q_c(q, \widehat{\theta}_1(p_1)))] > 0$ for any $p_1 \in (\mu - \varepsilon, \mu + \varepsilon)$:

[2] Optimal price $p_1^* \in [\mu - \varepsilon, \mu + \varepsilon)$, since charging $p_1 \geq \mu + \varepsilon$ or $p_1 < \mu - \varepsilon$ is strictly dominated by $p_1 = \mu - \varepsilon$.

(a) Since q_c is Gaussian with variance independent of p_1, and mean $\mu_u(p_1)$ increasing in p_1, the ex-ante distribution of posterior beliefs first-order stochastically increases in p_1 for $p_1 \in [\mu - \varepsilon, \mu + \varepsilon)$ (Levy 2015).

(b) (1.6) is a non-decreasing function of q_c, strictly increasing on some intervals. First-order stochastic dominance implies $\frac{\mathrm{d}}{\mathrm{d}p_1}\mathbb{E}_{q_c}[\Pi_2^*(q_c(q, \widehat{\theta}_1(p_1)))] \geq 0$. Gaussian distribution[3] and Hanoch and Levy (1969), Lemma 1, ensure that the inequality is strict, $\frac{\mathrm{d}}{\mathrm{d}p_1}\mathbb{E}_{q_c}[\Pi_2^*(q_c(q, \widehat{\theta}_1(p_1)))] > 0$.

In a setting without a review system, the monopolist's maximization problem in period 1 is $\max_{p_1 \in [\mu-\varepsilon, \mu+\varepsilon)} p_1 \frac{\varepsilon - (p_1 - \mu)}{2\varepsilon}$, resulting in optimal price $p_{1,a}^* = \frac{\mu+\varepsilon}{2} < p_1^*$. $\square$

[3]If Gaussian distribution A first-order stochastically dominates B, then $\mathrm{CDF}_A(x) < \mathrm{CDF}_B(x)\ \forall x \in \mathbb{R}$, where CDF denotes cummulative distribution function (Levy 2015).

Chapter 2

Bayesian Persuasion with Costly Information Acquisition

2.1 Introduction

A decision maker (a buyer, a politician) often relies on free information provided by an interested party (a seller, a lobbyist), but she may also be able to obtain her own information at a costly effort. Does she benefit from the ability to acquire her own information in addition to the given free information? Will she choose to acquire any?

We consider a Bayesian persuasion model (Kamenica and Gentzkow 2011; henceforth KG) extended to an *endogenous* acquisition of costly information. As in KG, a sender chooses a signal conveying information on the unknown state of the world to disclose to a receiver, a decision maker. However, unlike in KG, before taking an action, the receiver further chooses her own signal under an entropy-based cost, as in rational inattention (Sims 2003). We show that the possibility of additional learning lessens the sender's persuasive power (lowering his expected equilibrium utility). However, the outcome can be worse also for the receiver, as the sender can strategically prefer to disclose significantly less information when the receiver has her own learning option. For instance, the sender's strategic manipulation of a receiver's consideration set or his dislike for particular actions can lead to such a scenario, see Section 2.6.

We exploit a similarity in Bayesian persuasion and rational inattention allowing for tractability: any signal is feasible as long as it is consistent with prior beliefs. Signals can thus be explicitly modeled by posterior distributions over unknown states, under the martingale property (KG; Caplin and Dean 2013; henceforth CD). An optimal signal is then found by a *concavification* of an underlying value function of posterior beliefs related to the expected utility of the information designer. The concavification method remains valid in our model, but it requires solving the receiver's maximization problems for an entire space of beliefs first, which quickly becomes intractable. Further, the concavification method itself has limited applicability even for standard Bayesian persuasion problem as concavification of a general function is notoriously difficult.

We thus propose a new method not relying on concavification; it is sufficient to search through a relatively small finite set of the sender's signals, characterized by a series of linear conditions. The method is also applicable to a KG model, which is a limiting case of our model, and can then simplify, sometimes dramatically, the search for a solution. First, a Never-Learning Lemma states that our model can be solved as a standard Bayesian persuasion under an additional constraint: the receiver never costly learns. This results from both the sender's and the receiver's information technology being unconstrained (apart from the martingale property) and certain properties of the receiver's cost function. Second, Proposition 1 shows that, using a series of specific linear conditions, we can construct a finite set of the sender's signals satisfying the additional constraint and in which some optimal strategy must be contained. The new method complements the result of Lipnowski and Mathevet (2017) who show, in the standard Bayesian persuasion model, sufficiency to consider a properly chosen subset of the sender's signals. While they provide general abstract conditions on the subset, we provide readily applicable linear conditions, which follow from the entropy-based cost, but are also valid for the standard Bayesian persuasion at the limit.

We conclude the paper by examining the robustness to variations of the receiver's cost. The key simplification step (the Never-Learning Lemma) holds for a whole class of *posterior-separable* cost functions, for which the entropy-based cost is a prime example. Once the information technology is less flexible, this simplification need not hold;

the sender can take advantage of the restriction on the set of feasible receiver's signals. However, the possibility that the receiver can be hurt by having the option to learn is not unique to posterior-separable cost functions.

This paper is organized as follows. Section 2.2 provides an overview of the relevant literature. Section 2.3 sets up a motivating example. Section 2.4 provides a general model. Section 2.5 states the main simplification result and describes the new solution method. Section 2.6 describes comparative statics, giving examples in which the receiver does not benefit from having the learning option. Section 2.7 discusses the assumption of the receiver's information technology, and Section 2.8 concludes.

2.2 Related literature

Our model extends KG by enabling a receiver to *endogenously* acquire her own costly information. Extensions with an exogenously privately informed receiver have already been examined in Kolotilin (2015), Kolotilin (2018) and Kolotilin et al. (2017), and summarized in Bergemann and Morris (2016). Our result showing that the receiver's expected equilibrium utility can be non-monotone in her information cost parameter has a similar intuition as Kolotilin (2018), showing that the receiver's expected equilibrium utility can be non-monotone in the precision of her (costless) exogenous private information: in strategic environments, a sender can become discouraged and optimally respond to a receiver with better information technology by disclosing less information. Our model is more general than the above papers, because we neither restrict the set of actions to be binary, as they all do, nor do we assume linear environments as Kolotilin et al. (2017) do. The notion that an agent in a strategic setting can be hurt by having access to a better information technology is not unique for Bayesian persuasion, e.g., see Roesler and Szentes (2017) and Kessler (1998) for such a case in a contracting environment.

As the concavification approach has limited applicability for a number of problems, we also contribute to the Bayesian persuasion literature that focuses on alternative solution techniques not relying on concavification. For linear environments with binary actions, a more restricted model than ours, Kolotilin et al. (2017) reformulate the Bayesian persuasion problem as the maximization of a linear functional on a bounded set of a convex functions and use it to simplify the problem to a finite-variable optimization problem.

For a more general version, but also with binary actions, Kolotilin (2018) permits the verification of whether a candidate sender's strategy is optimal. However, this approach has limited applicability because it does not allow to directly characterize the optimal sender's strategy. On the other hand, we propose a solution method for a model with general utility functions and finite, but arbitrarily large action and state spaces, that allows one to directly find a solution. Our method is most similar to the Lipnowski and Mathevet (2017) method proposed for the standard Bayesian persuasion model. While they provide general abstract conditions, we give exact liner conditions that can be readily used not only for our version of the model, but also for standard Bayesian persuasion.

There is a rapidly growing Bayesian persuasion literature. Kolotilin (2015) studies Bayesian persuasion with exogenous public information. Rayo and Segal (2010), Perez-Richet and Prady (2011), Alonso and Câmara (2018), and Hedlund (2017) explore Bayesian persuasion with a privately informed sender. Gentzkow and Kamenica (2014) model situations in which the sender bears a cost associated with his information. They provide a class of cost functions (including entropy-based cost) that are compatible with the concavification approach. Similarly, we work with an entropy-based cost, but the receiver is the bearer of the cost. Other extensions include competition (Gentzkow and Kamenica 2016; Gentzkow and Kamenica 2017; Li and Norman 2017), heterogeneous priors (Alonso and Câmara 2016), or dynamic framework (Au 2015; Ely, Frankel, and Kamenica 2015; Ely 2017).

The assumptions about the receiver's cost function falls under rational inattention (Sims 2003). Single-agent rational inattention decision problems have been studied for investment decisions (Van Nieuwerburgh and Veldkamp 2009), rare events (Maćkowiak and Wiederholt 2015), static stochastic choice (Caplin and Dean 2013; Caplin and Dean 2015; Caplin, Dean, and Leahy 2016; Oliveira et al. 2017; Matějka and McKay 2015), or dynamic stochastic choice (Steiner, Stewart, and Matějka 2017). Yang (2011), Martin (2017), and Ravid (2017) examine rational inattention in strategic situations. While the overall framework of our model is strategic, the particular rational inattention problem is essentially a single-agent decision problem since the costly information acquisition occurs at the last stage of the game. We thus follow the papers on static stochastic choice when solving a receiver's rational inattention problem.

2.3 Motivation: Simple model

A seller (he) is persuading a buyer (she) to purchase his product (e.g., a music CD), which can be either a good match ($\omega = 1$) or a bad match ($\omega = 0$). The buyer can either buy ($a = 1$) or not buy ($a = 0$). $u(a, \omega)$ and $v(a)$ are the buyer's and seller's utilities,

$$u(a,\omega) = \begin{cases} 1 & a = 1 \ \wedge \ \omega = 1 \\ -1 & a = 1 \ \wedge \ \omega = 0 \\ 0 & \text{otherwise} \end{cases}, \qquad v(a) = \begin{cases} 1 & a = 1 \\ 0 & a = 0 \end{cases}.$$

We identify the beliefs with probability that $\omega = 1$. A common prior belief is $\mu_0 := \Pr[\omega = 1] < 0.5$ (under which the buyer's optimal action is not to buy).

The seller may persuade the buyer to take his preferred action (to buy) by providing further information (e.g., let her listen to a song). The buyer then updates her priors to an interim belief $\mu := \Pr[\omega = 1|\text{seller's information}]$. The seller's information strategy is a choice of a lottery $\tau \in \Delta([0, 1])$ over interim beliefs with mean μ_0.

After the buyer updates to a particular interim belief μ, she can gather additional information at a costly effort (e.g. search on the Internet), further updating her beliefs to a posterior belief $\gamma = \Pr[\omega = 1|\text{seller's and buyer's information}]$. Her information strategy is a choice of a lottery $\phi \in \Delta([0, 1])$ over posterior beliefs with mean μ. If the optimal lottery satisfies $\text{supp}(\phi) = \mu$, we say she *does not learn* at μ. Otherwise, we say she *learns* at μ.

While the seller's information is costless, the buyer bears a cost for her information (e.g. opportunity cost of time). Given μ, the cost of a lottery ϕ is $\lambda\left(H(\mu) - \mathbb{E}_\phi H(\gamma)\right)$, where $\lambda \geq 0$ is an information cost parameter and $H(\mu) - \mathbb{E}_\phi H(\gamma) \geq 0$ states how much uncertainty about the match, as measured by Shannon entropy[1] $H(\cdot)$, is expected to be reduced by ϕ.

To solve the game, we exploit one feature of the buyer's optimal behavior: once she obtains her chosen information (updates to a particular posterior γ), she never wishes to engage in another round of learning even if given a chance. This stems from a set of the buyer's information strategies being unconstrained (apart from the consistency requirement that

[1] The Shannon entropy at belief $p \in [0, 1]$ is $H(p) = -(p \ln p + (1 - p) \ln(1 - p))$ where $0 \ln 0 = 0$.

a mean is preserved) and from her cost function being *posterior-separable* (see Section 2.7). The latter guarantees that the cost is increasing in Blackwell informativeness and that it is invariant to intermediate stages[2]. As the seller's set of information strategies is also unconstrained, he can always skip the buyer's potential learning and directly 'send' her to the corresponding posteriors where she would have ended up by herself, without changing the outcome of the game. Since the seller's information is costless, it is thus sufficient to focus on a specific class of the seller's strategies under which the buyer never decides to further costly learn.

Buyer's optimal behavior

We follow an approach of CD to solve for the buyer's optimal behavior. Given μ, the buyer maximizes

$$\begin{array}{ll} \max_{\phi \in \Delta([0,1])} & \mathbb{E}_\phi[B(\gamma)] - \lambda\left(H(\mu) - \mathbb{E}_\phi[H(\gamma)]\right) \\ s.t. & \mathbb{E}_\phi[\gamma] = \mu, \end{array} \tag{2.1}$$

where the expectation is taken over posterior beliefs induced by ϕ and $B(\gamma)$ is the buyer's gross expected utility at posterior γ given that her subsequent action is optimal. Hence, $B(\gamma) = 0$ for $\gamma < 1/2$ (not buying) and $B(\gamma) = 2\gamma - 1$ otherwise (buying).

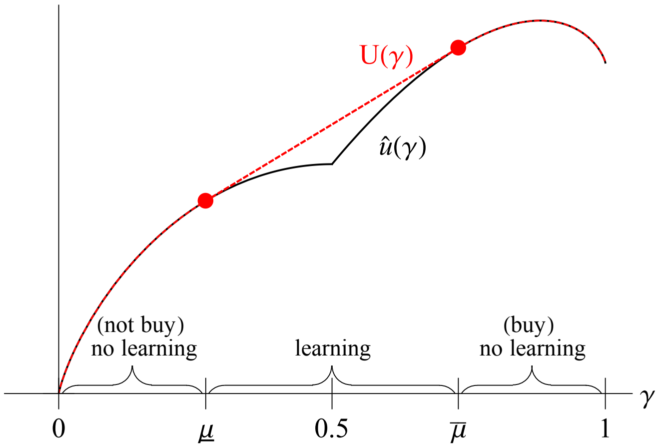
Note that the problem (2.1) can be rewritten as

$$\begin{array}{ll} \max_{\phi \in \Delta([0,1])} & \mathbb{E}_\phi[\hat{u}(\gamma)] - \underbrace{\lambda H(\mu)}_{=const.} \\ s.t. & \mathbb{E}_\phi[\gamma] = \mu, \end{array} \tag{2.2}$$

where $\hat{u}(\gamma) = B(\gamma) + \lambda H(\gamma)$ is the buyer's value function at posterior γ. The problem (2.2) has a geometric interpretation. Let $U(\gamma)$ be a *concavification* of $\hat{u}(\gamma)$ defined as the smallest concave function that is everywhere weakly greater than $\hat{u}(\gamma)$. CD showed that the support of an optimal lottery ϕ^* are those posterior beliefs that support the tangent hyperplane to the lower epigraph of the concavification above the interim belief μ, $U(\mu)$. Hence, whenever $\hat{u}(\mu) = U(\mu)$, the receiver does not learn at μ and whenever

[2] The cost of achieving a particular distribution of posterior beliefs would be the same regardless of whether the learning occurs in one or more stages.

Figure 2.1: Buyer's value function $\hat{u}(\gamma)$ and its concavification $U(\gamma)$



$\hat{u}(\mu) < U(\mu)$, she learns at μ, where the support of the optimal lottery is always the same: $\text{supp}(\phi^*) = \{\underline{\mu}, \overline{\mu}\}$, see Figure 2.1.

Hence, there are two threshold interim beliefs[3] $0 \leq \underline{\mu} \leq \overline{\mu} \leq 1$ that divide the space of interim beliefs into two *non-learning regions* and one *learning region*. A non-learning region of a particular action are all interim beliefs at which the buyer does not learn and optimally takes that action. A non-learning region of not buying is $[0, \underline{\mu}]$ and that of buying is $[\overline{\mu}, 1]$. For intermediate values of μ, when the buyer is very uncertain about what the right thing to do is, she learns and only sometimes buys (in the case of favorable information). Note that once she obtains her information (updates her beliefs either to a posterior $\underline{\mu}$ or $\overline{\mu}$), she does not wish to engage in another round of learning even if given the chance[4].

Bayesian persuasion s.t. never-learning constraint

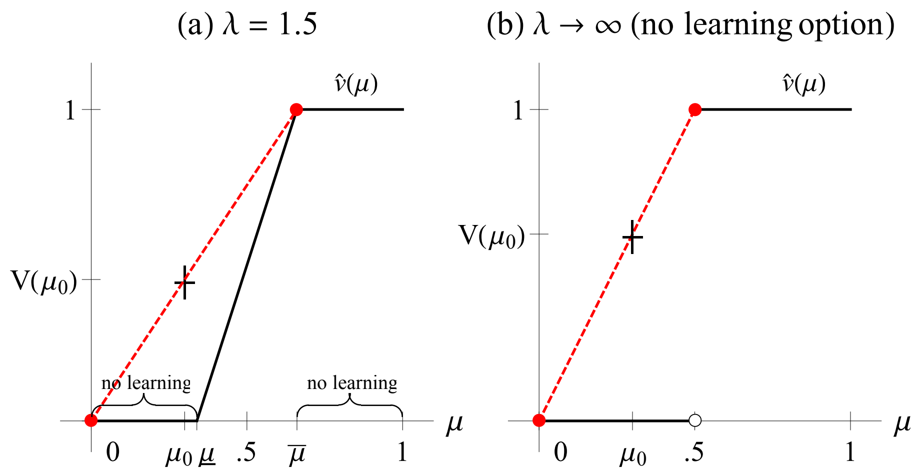
For each μ, let $\hat{v}(\mu)$ be the seller's expected utility which already accounts for the optimal buyer's behavior at μ.[5] Let $V(\mu)$ be a *concavification* of $\hat{v}(\mu)$ defined as the smallest concave function that is everywhere weakly greater than $\hat{v}$. Then the seller's expected

[3] Solving the buyer's maximization problem, we obtain $\underline{\mu} = \frac{1}{1+e^{\frac{1}{\lambda}}}$ and $\overline{\mu} = \frac{e^{\frac{1}{\lambda}}}{1+e^{\frac{1}{\lambda}}}$ (see Appendix 2.A).

[4] As the buyer's value function $\hat{u}(\gamma)$ and its concavification $U(\gamma)$ coincide at values $\underline{\mu}$ and $\overline{\mu}$, no learning is optimal either at $\underline{\mu}$ or at $\overline{\mu}$.

[5] Hence, $\hat{v}(\mu) = 0$ if $\mu \leq \underline{\mu}$ (the buyer does not learn and does not buy), $\hat{v}(\mu) = 1$ if $\mu \geq \overline{\mu}$ (the buyer does not learn and buys), and $\hat{v}(\mu) = \frac{1}{\overline{\mu}-\underline{\mu}}\mu - \frac{\underline{\mu}}{\overline{\mu}-\underline{\mu}}$ for $\mu \in (\underline{\mu}, \overline{\mu})$.

Figure 2.2: Function $\hat{v}(\mu)$ and sender's optimal information strategy



equilibrium utility is the concavification evaluated at the prior, $V(\mu_0)$, and the support of the optimal sender's lottery can be found from the graph in the same fashion as in the buyer's problem. See Figure 2.2 for an example of $\hat{v}(\mu)$ and the resulting optimal sender's strategy with $\lambda \to \infty$ (equivalent to KG's setting with a buyer who cannot gather her own information) and with $\lambda = 1.5$.

The support of the optimal lottery is $\text{supp}(\tau^*) = \{0, \overline{\mu}\}$ (when $\lambda \to \infty$, $\underline{\mu} = \overline{\mu} = 1/2$). These are beliefs that belong to non-learning regions. We show that it is generally sufficient to only consider the seller's information strategies under which the buyer never costly learns, i.e., the lotteries with the support over interim beliefs from non-learning regions only. This is because the buyer never wishes to have more than one round of costly learning, the set of the seller's information strategies is unconstrained, and his information is costless. The seller can then skip the receiver's learning part with his information without changing the outcome of the game.

Further, note that the support of the optimal lottery, $\text{supp}(\tau^*) = \{0, \overline{\mu}\}$, are *extreme points*[6] of the non-learning regions. We show that it is sufficient to consider the sender's strategies under which only extreme points of non-learning regions are chosen. Hence, in this example, one only needs to consider lotteries with support over interim beliefs from the set $\{0, \underline{\mu}, \overline{\mu}, 1\}$, where the thresholds $\underline{\mu}$, $\overline{\mu}$ are specified by particular linear equations

[6] An extreme point of a convex set S is a point in S which does not lie in any open line segment joining two points of S. Hence, extreme points of non-learning region $[0, \underline{\mu}]$ are $\{0, \underline{\mu}\}$ and extreme points of non-learning region $[\overline{\mu}, 1]$ are $\{\overline{\mu}, 1\}$.

resulting from the characterization of the optimal receiver's strategy.

2.4 General model

A receiver (she) chooses an action a from a finite set A. A payoff-relevant state ω is drawn from a finite set Ω according to an interior prior distribution $\mu_0 \in \Delta(\Omega)$. Before choosing her action, the receiver obtains free information about ω provided by a sender (he) and rationally updates her beliefs from the prior μ_0 to interim belief $\mu \in \Delta(\Omega)$. A sender's (information) strategy is a choice of a distribution $\tau \in \Delta(\Delta(\Omega))$ over the (updated) interim beliefs s.t. $\mathbb{E}_\tau[\mu] = \mu_0$ (martingale property).

After updating to a particular μ and before her move, the receiver can acquire additional costly information about ω, further rationally updating her beliefs from μ to a posterior belief $\gamma \in \Delta(\Omega)$. A receiver's strategy is comprised of an information strategy, which is a choice of a distribution $\phi \in \Delta(\Delta(\Omega))$ over the (further updated) posterior beliefs s.t. $\mathbb{E}_\phi[\gamma] = \mu$ (martingale property), and an action strategy $\sigma : \Delta(\Omega) \to A$, where $\sigma(\gamma)$ indicates the choice of action at a posterior belief γ. Let S be the set of all action strategies. We focus on sender-preferred subgame perfect equilibria: in case of indifference, the receiver uses a strategy that is (weakly) preferred by the sender[7].

The sender bears no information costs and derives utility $v(a, \omega)$. The *value* of the sender's strategy is the equilibrium expectation of $v(a, \omega)$ under that strategy profile. The *sender benefits from persuasion* if there exists τ whose value is strictly larger than the equilibrium expectation of $v(a, \omega)$ under no sender's information, defined as τ_0 with $\text{supp}(\tau_0) = \mu_0$.

The receiver derives gross utility $u(a, \omega)$, where the term 'gross' indicates that information costs are not included. As is standard in RI literature, we assume Shannon-entropy based cost. For a random variable T with finite support distributed according to $\mu \in \Delta(\text{supp}(T))$, the Shannon entropy is given by

$$H(T|\mu) = -\sum_{\theta \in \text{supp}(T)} \mu(\theta) \ln \mu(\theta), \tag{2.3}$$

[7] See 2.A for a sufficient assumption for a unique optimal receiver's strategy.

which is a measure of uncertainty about T (where $0 \log 0 = 0$ by convention). We assume the cost is proportional to the conditional mutual information[8]

$$I(\phi, \omega|\mu) = H(\omega|\mu) - \mathbb{E}_{\phi}[H(\omega|\gamma)] \tag{2.4}$$

between a receiver's information strategy ϕ and the state ω. Given μ, it captures how much uncertainty about ω is expected to be reduced by ϕ. The receiver solves the following problem.

Definition 2. Given interim belief μ, the *receiver's rational inattention problem* (henceforth receiver's RI problem) is

$$\begin{aligned} \max_{\phi \in \Delta(\Delta(\Omega)), \sigma \in S} \quad & \mathbb{E}_{\phi}\left[\sum_{\omega \in \Omega} \gamma(\omega) u\left(\sigma(\gamma), \omega\right)\right] - \lambda I(\phi, \omega|\mu) \\ s.t. \quad & \mathbb{E}_{\phi}[\gamma] = \mu, \end{aligned} \tag{2.5}$$

where $\lambda \geq 0$ is an information marginal cost parameter, and the expectation is over posterior beliefs γ distributed according to ϕ; $\gamma(\omega)$ denotes the probability of state ω at belief γ.

Matějka and McKay (2015) prove the existence of a solution to (2.5), which we denote by (ϕ^*_μ, σ^*)[9]. For a characterization of the receiver's optimal behavior, see 2.A. Say the receiver *does not learn* at μ if $\text{supp}(\phi^*_\mu) = \mu$. Otherwise, say the receiver *learns* at μ.

Applying backward induction, we can express the conditional sender's expected utility for each μ, denoted by $\hat{v}(\mu)$, where

$$\hat{v}(\mu) := \mathbb{E}_{\phi^*_\mu}\left[\sum_{\omega \in \Omega} \gamma(\omega) v(\sigma^*(\gamma), \omega)\right], \tag{2.6}$$

where the expectation is over posterior beliefs γ distributed according to ϕ^*_μ. $\hat{v}(\mu)$ is the sender's expected utility at an interim belief μ already accounting for the subsequent optimal receiver's behavior at μ. The sender solves the following problem.

[8] For more on the Shannon entropy and mutual information, see Cover and Thomas (2006).

[9] The optimal action strategy is independent of the interim belief as $\arg\max_a \mathbb{E}_\gamma u(a, \omega)$ is independent of the intermediate steps as to how one arrives at having posterior belief γ.

Definition 3. Given prior μ_0, the *sender's maximization problem* is

$$\begin{aligned} \max_{\tau \in \Delta(\Delta(\Omega))} \quad & \mathbb{E}_\tau[\hat{v}(\mu)] \\ s.t. \quad & \mathbb{E}_\tau[\mu] = \mu_0, \end{aligned} \tag{2.7}$$

where the expectation is over interim beliefs μ distributed according to τ.

Say the receiver *never learns* if the optimal sender's strategy τ^* satisfies: $\forall \mu \in \text{supp}(\tau^*)$, the receiver does not learn at μ.

2.5 Bayesian persuasion s.t. never-learning

While $\hat{v}(\mu)$ is straightforward when the receiver has no additional learning option[10], it becomes complicated once she can learn. It requires finding the receiver's optimal behavior for an entire space of interim beliefs, $\{(\phi^*_\mu, \sigma^*)\}_{\mu \in \Delta(\Omega)}$, which is intractable already for a small state space. We provide an approach that avoids such calculations.

First, we show that the game can be solved as a standard Bayesian persuasion model subject to an additional constraint: the receiver never costly learns. Determining the interim beliefs at which the receiver does not learn is sufficient. At such beliefs, the receiver's behavior is deterministic and $\hat{v}(\mu) = \sum_{\omega \in \Omega} \mu(\omega) v(\sigma^*(\mu), \omega)$. We then further specify a finite set of these beliefs on which some optimal sender's strategy must be supported.

2.5.1 Never-learning constraint

When the receiver is fairly uncertain about what the right thing to do is, she first learns to refine her beliefs before acting. However, when her interim belief is precise enough, she does not learn. Let us formalize the subsets of such interim beliefs.

[10] With no additional learning option, the receiver's optimal action is always deterministic at μ. In that case, $\hat{v}(\mu)$ is a piecewise-linear (upper semi-continuous) function: $\forall \mu \in \Delta(\Omega)$ we have $\hat{v}(\mu) = \sum_{\omega \in \Omega} \mu(\omega) v(\sigma^*(\mu), \omega)$.

Definition 4. A *non-learning region of action $a \in A$* is

$$NL^a \quad := \quad \{\mu \in \Delta(\Omega) : \text{supp}(\phi^*_\mu) = \mu \ \wedge \ a \in \arg\max_{a' \in A} \sum_{\omega \in \Omega} \mu(\omega)[u(a', \omega)]\}. \qquad (2.8)$$

The non-learning region of some action are all interim beliefs μ at which no learning and taking that action are optimal[11]. In the introductory example, a non-learning region of *not buy* is $\mu \in [0, \underline{\mu}]$, and that of *buy* is $\mu \in [\overline{\mu}, 1]$.

The following Never-Learning Lemma states that it is sufficient to focus on a subset of the sender's strategies under which the receiver never learns. The game can be solved as a standard Bayesian persuasion problem subject to a *never-learning* constraint[12].

Lemma 1 (Never-Learning)**.** *Let τ be a sender's information strategy of value v. Then there exists a sender's strategy τ' of value v where $\forall \mu \in$ supp(τ'): $\mu \in \cup_a NL^a$.*

In a proof of Never-Learning Lemma, we use a specific feature of the receiver's optimal behavior captured in Lemma 2.

Lemma 2. *The receiver wants to costly learn at most once, even if more rounds of costly learning were possible: $\forall \mu \in \Delta(\Omega)$, a receiver's optimal information strategy ϕ^*_μ satisfies: if $\gamma \in$ supp(ϕ^*_μ) then $\gamma \in \cup_{a \in A} NL^a$.*

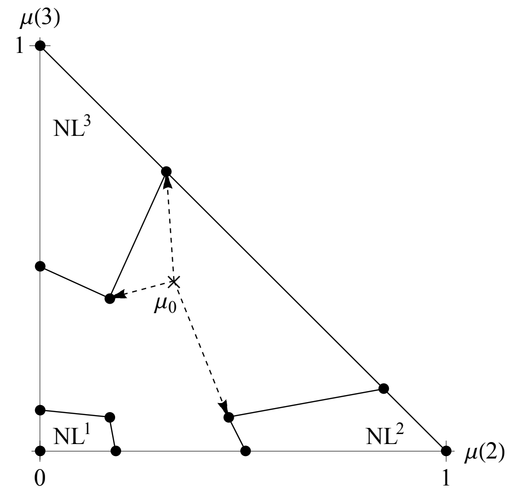
Lemma 2 follows from the set of the receiver's information strategies being unconstrained, the cost being increasing in Blackwell informativeness and invariant to intermediate stages[13]. Then, as the set of the sender's information strategies is also unconstrained, he can incorporate, at no cost, any receiver information strategy. Lemma 2 implies that doing so does not change the particular outcome of the game (a distribution of the receiver's actions conditional on the state) with respect to what outcome the original sender's information strategy induced. Thus, there is no need to solve the receiver's RI problems for an entire space of interim beliefs; finding non-learning regions is sufficient. The Never-Learning Lemma is an analogy to a revelation principle in mechanism design problems.

[11] When non-learning regions overlap, i.e. $\exists \mu \in NL^a$: $|\arg\max_{a' \in A} \sum_{\omega \in \Omega} \mu(\omega)[u(a', \omega)]| > 1$, the optimal action strategy, $\sigma^*(\mu)$, follows a sender-preferred equilibrium assumption.

[12] For any $\lambda \in \mathbb{R}$, $\hat{v}(\mu)$ differs from $\hat{v}(\mu)$ when the receiver has no additional learning option ($\lambda \to \infty$) only at μ not belonging to some non-learning region.

[13] The cost of achieving a particular distribution of posterior beliefs would be the same regardless of whether the learning occurs in one or more stages.

Figure 2.3: Extreme points of non-learning regions and a candidate sender's optimal strategy ($a, \omega \in \{1, 2, 3\}$, $u(a, \omega) = a$ if $a = \omega$ and 0 otherwise, $\lambda = 2$).



2.5.2 Extreme-points solution method

Recall that in the introductory example, there is a (unique) seller's strategy, under which only the extreme points of non-learning regions are induced. While the uniqueness property is not general (see Section 2.5.3), the optimality of some such strategy is generally guaranteed, which is captured in Proposition 2.

Let EP^a denote the *set of extreme points*[14] *of a non-learning region* NL^a.

Proposition 2. *The set* $\cup_{a \in A} EP^a$ *is non-empty and finite. Furthermore, whenever a sender's problem (2.7) has a solution, there exists an optimal sender's strategy* τ^**, for which* $|supp(\tau^*)| \leq |\Omega|$ *and* $\forall \mu \in supp(\tau^*)$*:* $\mu \in \cup_{a \in A} EP^a$.

Note that a sender-preferred equilibrium assumption implies an upper semi-continuity of $\hat{v}$, which guarantees the existence of the equilibrium. Proposition 2 then says that we can solve the game by comparing values of a finite number of the sender's strategies. These strategies induce (at most $|\Omega|$ of) extreme points of non-learning regions. See Figure 2.3 for an illustration of one such candidate sender's strategy. Once $v(a, \omega)$ is specified, we can determine the optimal one. The proof of Proposition 2 shows that for any sender strategy τ under which the receiver never learns, there exists a sender's strategy τ', inducing only

[14] An extreme point of a convex set B is a point in B which does not lie in any open line segment joining two points of B.

the extreme points, that has weakly higher value (where the sender-preferred assumption is used for a limit case of $\lambda \to \infty$). Carathéodory's theorem is then used in restricting the size of the support of an optimal sender's strategy.

The following lemmas characterize the set $\cup_a EP^a$ and the values of the candidate sender's strategies from Proposition 2. First, Lemma 3 characterizes the non-learning regions as being either a closed convex set determined by a finite series of linear inequalities (where the receiver's primitives—$u(a,\omega)$, λ—are the parameters) or an empty set. The linear conditions result from taking Shannon entropy as a measure of uncertainty in a *posterior-separable* cost function, see Section 2.7. The conditions follow from eq. (2.14) in Appendix 2.A, where the receiver's optimal behavior is characterized.

Lemma 3. *For any $\lambda \geq 0$ we have $\cup_{a\in A} NL^a \neq \emptyset$. Furthermore,*

$$NL^a = \left\{ \mu \in \Delta(\Omega) : \sum_{\omega\in\Omega} \mu(\omega) \left(\frac{e^{\frac{u(a',\omega)}{\lambda}}}{e^{\frac{u(a,\omega)}{\lambda}}} \right) \leq 1 \quad \forall a' \neq a \right\} \quad \forall a \in A. \tag{2.9}$$

Whenever $NL^a \neq \emptyset$, Lemma 3 implies that NL^a has (finitely many) extreme points (Krein-Milman Theorem). Lemma 4 states that an extreme point of a non-learning region is a belief in NL^a for which $|\Omega|$ of constraints from Lemma 3 are binding.

Lemma 4. *An extreme point of NL^a is $\mu \in \mathbb{R}^{|\Omega|}$ where $\sum_\omega \mu(\omega) = 1$ that satisfies (i) $\forall \omega \in \Omega$: $\mu(\omega) \geq 0$ and $\mu(\omega) \leq 1$; and (ii) $\sum_{\omega\in\Omega} \mu(\omega) \left(\frac{e^{\frac{u(a',\omega)}{\lambda}}}{e^{\frac{u(a,\omega)}{\lambda}}} \right) \leq 1 \quad \forall a' \neq a$, of which $|\Omega| - 1$ affine independent constraints are binding.*

Lemma 5 determines the value $\hat{v}(\mu)$ when μ is an extreme point of some non-learning region. When an extreme point belongs to more non-learning regions, the sender-preferred equilibrium assumption applies.

Lemma 5. *Let $\mu \in EP^a$ and $a = \sigma^*(\mu)$. Then $\hat{v}(\mu) = \sum_{\omega\in\Omega} \mu(\omega) v(a,\omega)$.*

Hence, to find an optimal sender's strategy, it suffices to:
(i) determine $\cup_{a\in A} EP^a$ (using Lemma 4);
(ii) evaluate $\hat{v}(\mu)$ at those beliefs (using Lemma 5);
(iii) compare the values of the sender's strategies that satisfy Bayes' law and induce at most $|\Omega|$ beliefs from the set $\cup_{a\in A} EP^a$.

Note that Lemmas 3, 4, 5 and Proposition 2 are applicable to KG, the case of a receiver with no additional learning option, by taking $\lambda \to \infty$.

2.5.3 Equilibrium with learning?

In a setting with binary action and state spaces, a setting used in a number of recent papers[15], the receiver never costly learns in an equilibrium as long as the sender benefits from persuasion, see Proposition 3. In a more general setting, however, this does not necessarily hold in case of the multiple equilibria (even when the sender benefits from persuasion), see Example 1. However, if we further assumed that the sender incurs a strictly positive cost whenever the receiver learns (e.g., waiting cost), equilibria with additional learning would disappear.

Let us first slightly restrict the sender's preferences, a necessary and sufficient condition for a unique equilibrium in a binary action and state spaces case. We rule out pathological cases that can lead to situations in which two actions a and a' are both induced (under no learning) by a sender's optimal strategy, but at the belief at which the receiver takes a, the sender is exactly indifferent between a and a'.

Assumption 2. *There exists no action $a \in A$ s.t. (i) $\forall \mu \in \Delta(\Omega) : \hat{v}(\mu) \leq \sum_{\omega\in\Omega} \mu(\omega)v(a,\omega)$, and (ii) $\exists \mu \in NL^{a'}$ where $a' \neq a$ and $\hat{v}(\mu) = \sum_{\omega\in\Omega} \mu(\omega)v(a,\omega)$.*

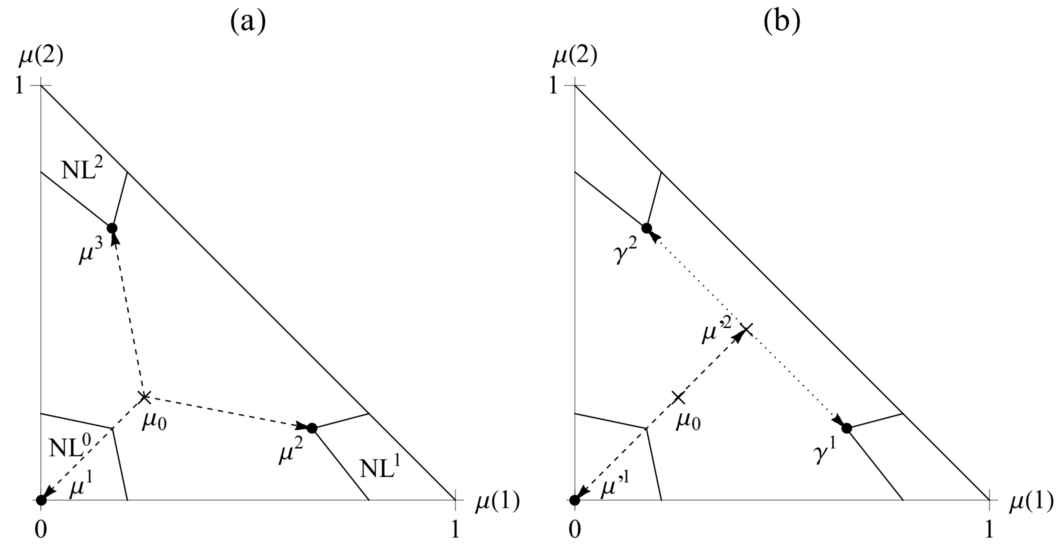
Proposition 3. *Suppose $|A| = |\Omega| = 2$ and the sender benefits from persuasion. Then, (i) the receiver never learns in any equilibrium; and (ii) A2 holds if and only if there exists unique equilibrium.*

Example 1. *Let $A = \Omega = \{0, 1, 2\}$; $u(a,\omega) = 1$ if $a = \omega$ and $u(a,\omega) = 0$ otherwise; $v(a,\omega) = 1$ if $a \neq 0, a \neq \omega$, and $v(a,\omega) = 0$ otherwise; $\lambda = 0.75$; prior belief: $\{\mu_0(0) = 0.5, \mu_0(1) = \mu_0(2) = 0.25\}$.*

Figure 2.4 depicts two optimal sender strategies for Example 1. In part a), an optimal sender's strategy τ^*: $\text{supp}(\tau^*) = \{\mu^1, \mu^2, \mu^3\}$ from Proposition 2 is shown. Under this strategy, the receiver never learns. In part b), a different sender's strategy τ'^*:

[15] Standard Bayesian persuasion was applied to bank regulation (Gick and Pausch 2012), electoral manipulation (Gehlbach and Simpser 2015), investment decisions (Bizzotto, Rüdiger, and Vigier 2015), and forecasting of disasters (Aoyagi 2014).

Figure 2.4: (a) Equilibrium with no learning; (b) Equilibrium with learning: the receiver learns at μ'^2 optimally inducing posteriors $\{\gamma^1, \gamma^2\}$.



$\text{supp}(\tau'^*) = \{\mu'^1, \mu'^2\}$ is considered. At μ'^2, the receiver learns and optimally induces posteriors $\{\gamma^1, \gamma^2\} = \{\mu^2, \mu^3\}$ with appropriate probabilities. Here, it is no longer true that the receiver never learns. As the outcome of the game (distribution of the receiver's strategies conditional on the state) is the same under both τ^* and τ'^*, τ'^* is also optimal. Note that the assumption A2 is satisfied in this example and hence A2 generally is not a sufficient assumption for uniqueness of equilibria.

2.5.4 Characterization of the sender's optimal strategies

Let us provide further characterization of candidate strategies from Proposition 2. Recall that in the introductory example, the optimal sender's strategy has $\text{supp}(\tau^*) = \{0, \overline{\mu}\}$; the buyer does not buy when $\mu = 0$ and buys when $\mu = \overline{\mu}$. Note two properties. First, whenever the buyer chooses the *least-preferred action* (*not to buy*), *she is certain of the state*, $\mu = 0$; she never rejects a good match. Second, whenever the buyer chooses an *action that is not the seller's least-preferred* (*to buy*), her beliefs are at a *border of a non-learning region*[16]. If her belief was inside a non-learning region when she buys ($\mu > \overline{\mu}$), the seller can increase the probability of buying by slightly decreasing μ.

[16] With a binary state, a border coincides with an extreme point.

The first of these properties holds in general. Say an action $\underline{a}$ is a *worst action* if $v(\underline{a}, \omega) < v(a, \omega)$ for all $a \neq \underline{a}$ and ω. Let $EP^{\Delta(\Omega)}$ denote the set of extreme points of the probability simplex $\Delta(\Omega)$.

Proposition 4. *If an optimal sender's strategy from Proposition 2 induces a belief $\mu \in NL^{\underline{a}}$, where $\underline{a}$ is a worst action, and $\underline{a} = \sigma^*(\mu)$, then $\mu \in EP^{\Delta(\Omega)}$.*

Proposition 4 states that whenever the receiver takes a worst action in an equilibrium (under the sender's strategy from Proposition 2), the state is fully revealed[17]. In the introductory example, the action *not to buy* is a worst action. When the receiver takes it, she is certain that the state is *bad*.

The second property holds under restriction on the sender's preference. When A2 is not satisfied, the sender can be exactly indifferent to a change in the probability mass between two different actions that are both induced (under no learning) by an optimal sender's strategy. This may break the second property.

Let $\mathrm{bd}(B)$ denote a boundary of the set B.

Proposition 5. *Let A2 hold and suppose the sender benefits from persuasion. If an optimal sender's strategy induces $\mu \in NL^a$, then $\mu \in bd(NL^a)$.*

Note that NL^a has a piecewise-linear boundary. Consider two different extreme points of NL^a from different 'linear segments' of its boundary. Proposition 5 implies that, under A2, both beliefs cannot be optimally induced (otherwise an optimal sender's strategy inducing an interim belief from the interior of NL^a also exists)[18].

2.6 Comparative statics

In this section, we examine the relationship between agents' expected equilibrium utilities and the receiver's information cost parameter λ. We show that the access to information

[17] If the receiver takes a worst action in an equilibrium under a different strategy than that of Proposition 2, there is no uncertainty left in the sense that the probability of states for which $\underline{a}$ is not optimal to choose is zero. Proposition 4 is an analogy of Proposition 4 in KG.

[18] Proposition 5 is a modification of Proposition 5 in KG, which states that any optimally induced interior belief leads the receiver to being indifferent between at least two different actions (given the analogy of A2 is satisfied).

has a disciplinary effect on the sender (decreases his expected equilibrium utility), but it is not necessarily beneficial for the receiver either.

2.6.1 Sender

Proposition 6. *The sender's expected equilibrium utility (weakly) increases in* λ.

As the receiver's access to her own information represents an additional *never-learning* constraint for the sender, it can only hurt him. The non-learning regions, and hence the set of the sender's strategies under which the receiver never learns, do not shrink as λ increases. The receiver's potential learning is less threatening as her information becomes more expensive.

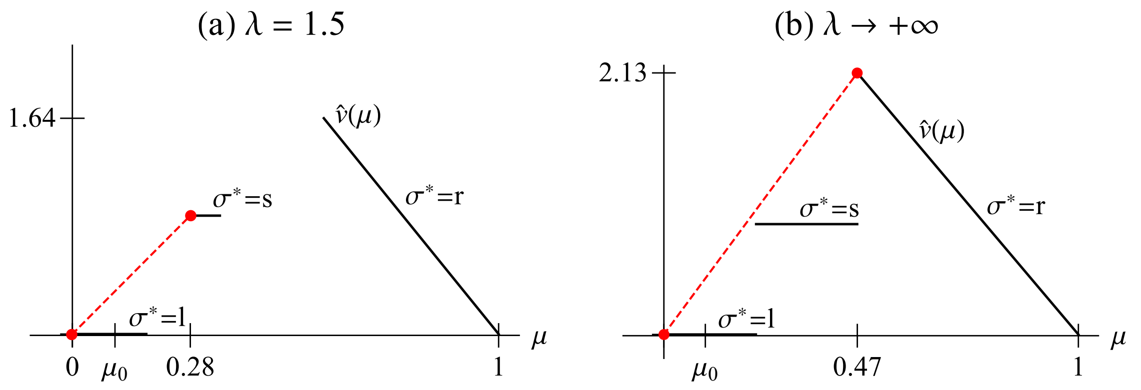
2.6.2 Receiver

Proposition 7. *Assume A2,* $|A| = |\Omega| = 2$, *and the sender benefits from persuasion* $\forall \lambda > 0$. *Then the receiver's expected equilibrium utility (weakly) decreases in* λ.

In a binary setting, the receiver (weakly) benefits from cheaper information, see Proposition 7. In general, however, the receiver does not necessarily gain from having the threat of learning; for intermediate cost, she can prefer commitment to not having the option to learn at all. Ceteris paribus, an agent would benefit from information being cheaper. However, in a strategic setting, the opponent responds to how expensive the information of the other agent is. With conflict of interest, the sender's choice under intermediate λ can be less informative (Blackwell sense[19]) than his choice under higher λ, making the receiver strictly prefer high to intermediate cost. We illustrate this in two examples. In Example 2, the sender targets a specific consideration set of the receiver (the set of actions chosen with strictly positive probability), and in Example 3, the sender dislikes a particular set of actions.

[19] An information strategy τ is more Blackwell-informative than τ' if and only if obtaining information via τ is preferred to information via τ' by all expected utility maximizers. Equivalently, τ is more Blackwell-informative than τ' if and only if supp(τ') lie inside the convex hull of supp(τ) (Blackwell and Girshick 1954), Thm 12.2.2.).

Figure 2.5: Example 2—Manipulation of the receiver's consideration set: $\hat{v}$ over non-learning regions and an optimal sender's strategy. The sender targets: (a) actions $\{l, s\}$ (less information); (b) actions $\{l, r\}$ (more information).

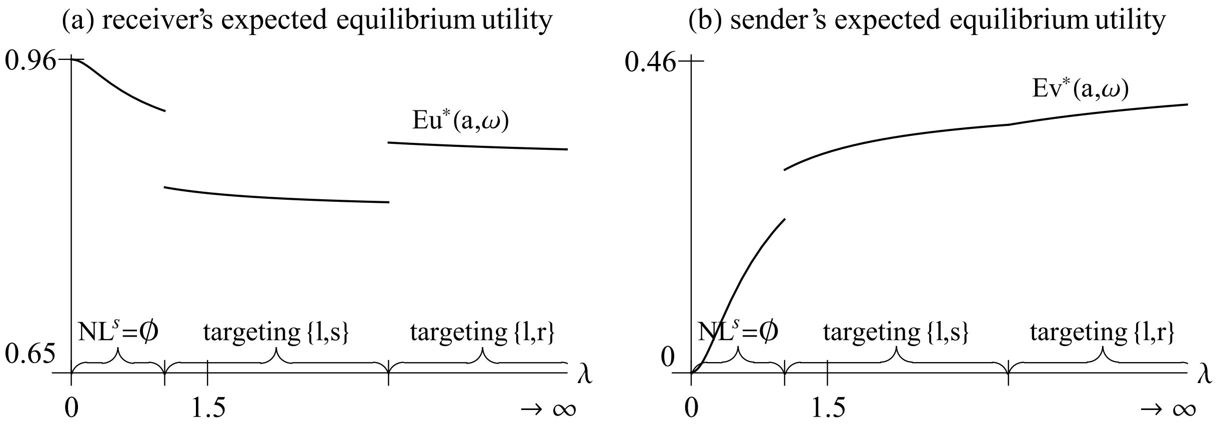


Example 2. $\Omega = \{0, 1\}$, $A = \{l, r, s\}$, *the prior belief* $\mu_0 := \Pr[\omega = 1] = 0.1$, *and*

$$u(a, \omega) = \begin{cases} 0.9 & a = l \wedge \omega = 0 \\ 1.5 & a = r \wedge \omega = 1 \\ 0.7 & a = s \\ 0 & otherwise \end{cases}, \quad v(a, \omega) = \begin{cases} 4 & a = r \wedge \omega = 0 \\ 0.9 & a = s \\ 0 & otherwise \end{cases}.$$

Consider a receiver with three actions, two risky (l, r) and one safe (s). At her prior, she would choose l, the sender's least preferred action. The sender can design an informative experiment inducing one other action to be chosen (upon favorable realization), where the 'amount' determines which one. With enough information, a sender's most preferred action, r, is chosen upon favorable realization. With less information, s is chosen upon favorable realization, but that happens with higher probability. Figure 2.5 depicts the sender's optimal choice; μ_0 and μ denote the prior and interim probabilities of $\omega = 1$, respectively. When $\lambda \to \infty$, the sender targets r. However, when $\lambda = 1.5$, too much information is now needed to target r and the sender finds it optimal to give less information and to be satisfied with targeting s, but with higher probability. Figure 2.6 depicts the agents' equilibrium expected utilities. For intermediate values of λ, the sender finds it optimal to target action s.

Figure 2.6: Example 2—Equilibrium expected utilities as a function of λ.



Example 3. *$\Omega = \{0, 1\}$, $A = \{L, R, l, r\}$. Let $u(L, 0) = u(R, 1) = 1$, $u(l, 0) = u(r, 1) = 0.8$, $u(l, 1) = u(r, 0) = 0.2$, and $u(a, \omega) = 0$ otherwise. Let $v(a, \omega) = u(a, \omega)$ if $a \in \{l, r\}$ and $v(a, \omega) = 0$ otherwise. Let the prior $\mu_0 := \Pr[\omega = 1] = 1/2$. Consider $\lambda_1 = 1.25$ and $\lambda_2 = 1250$.*

A receiver has four risky actions (L, R, l, r) and a prior belief at which she learns and possibly takes either l or r. A sender dislikes actions L, R, but cares about determining which of the two actions l, r is optimal for the receiver. He wants to give as much information as possible to distinguish which of the actions l, r is better, under the constraint that neither L nor R is chosen, occurring if too much information is given. As the receiver's information becomes more expensive, this constraint is less restrictive and the sender is able to give 'more' information (Blackwell sense) than before, see Figure 2.7; μ_0 and μ denote the prior and interim probabilities of $\omega = 1$, respectively.

2.7 Discussion of the cost function

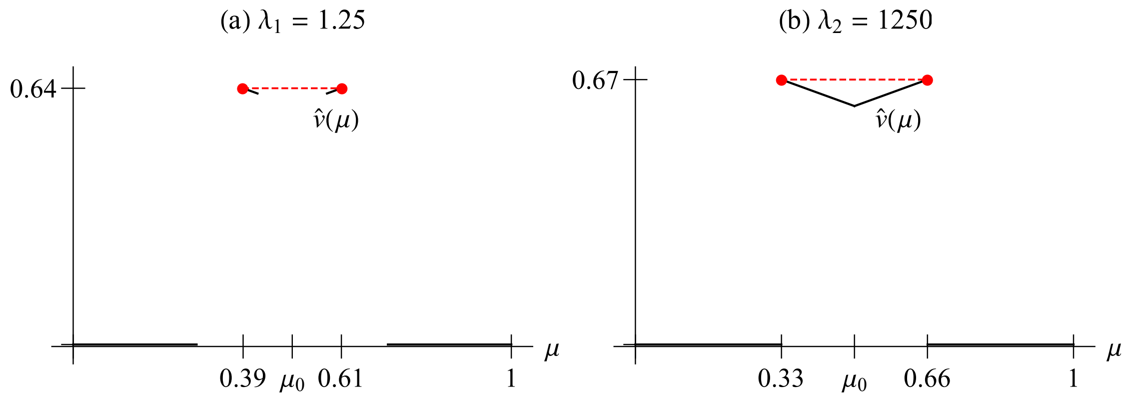
The assumed cost is a *posterior-separable* cost (Caplin, Dean, and Leahy 2017):

Definition 5. Given μ, a *posterior-separable* cost function is

$$c(\phi; \omega|\mu) = F(\omega|\mu) - \mathbb{E}_\phi[F(\omega|\gamma)] \tag{2.10}$$

where $F : \Delta(\Omega) \longrightarrow \mathbb{R}_+$ is a concave function and the expectation is over posteriors γ

Figure 2.7: Example 3—$\hat{v}(\mu)$ over non-learning regions and an optimal sender's strategy. The sender provides more information when $\lambda = 1250$ than when $\lambda = 1.25$.



induced by ϕ.

Lemma 2 and the Never-Learning Lemma hold under any posterior-separable cost because of the following properties. The cost is invariant to intermediate stages: the cost of achieving a particular distribution of posterior beliefs is the same if the learning occurs in one or more stages. The cost also does not impose any restriction on the set of feasible information strategies. The marginal cost of any receiver's information strategy is independent of the interim belief, i.e., of the starting point of the receiver's RI problem, and it is increasing in Blackwell informativeness. These properties guarantee Lemma 2, which is the core of the Never-Learning Lemma. Proposition 2 applies as well when the finiteness of $\cup_a EP^a$ is omitted from the statement. Lemmas 3 and 4 are specific for Shannon entropy, $F(\cdot) = \lambda H(\cdot)$.

The Never-Learning Lemma does not hold for all possible cost functions. For instance, a receiver choosing a precision of a normally distributed signal at some cost can wish to costly learn once more upon some signal realizations. However, in such a setting, if the receiver were allowed to engage in as many learning rounds as she wanted, an analogy of the Never-Learning Lemma would be obtained.

The possibility that the receiver can be hurt by having access to better information technology is not unique to posterior-separable costs. In Appendix 2.B, we solve the introductory example under a different cost function: by paying $c \geq 0$, the buyer obtains a binary signal $s \in \{good, bad\}$ of fixed precision $p := \Pr[s = \omega|\omega] > 0.5$. The seller exploits the restrictiveness of the buyer's set of signals, adding an additional effect in play. *First*, *the key simplification step*–the Never-Learning Lemma–*fails*. As the buyer cannot vary the precision of her signal with interim beliefs, she may wish to engage in more than one round of learning. The *seller* can take advantage of this invariance and *can strictly prefer to target buying through learning. Second, the seller's expected equilibrium utility is non-monotone in c.* Whenever he induces learning, he prefers a receiver with lower cost, because then inducing learning requires less provided information. *Third, the buyer's expected equilibrium utility is non-monotone in c.* The buyer prefers *intermediate* to low or high cost. Whenever the seller induces learning, he gives just enough information so that the buyer is indifferent between learning and not. Hence, any benefit from costly learning is exactly offset by paying cost c. With sufficiently high cost, however, the seller targets *buying* directly without the buyer's learning, who thus obtains valuable information without paying c.

2.8 Conclusion

We extend a model of Bayesian persuasion to a possibility of additional costly information acquisition by the receiver, modeled as in rational inattention. We exploit common features of Bayesian persuasion and rational inattention, resulting in a tractable model which can be used as a building block for applied problems. Based on the characterization of an optimal receiver's strategy, we offer an alternative solving algorithm characterized by a series of linear conditions. The new algorithm, which does not rely on standard concavification, is also applicable to a standard Bayesian persuasion model and can simplify, sometimes dramatically, the search for an optimal sender's strategy. We further show that the receiver does not necessarily benefit from having additional sources of information and can prefer commitment to not having any.

2.A Receiver's RI problem

2.A.1 Geometric interpretation—Concavification

The receiver's RI problem (2.5) can be rewritten as

$$\begin{aligned} \max_{\phi\in\Delta(\Delta(\Omega))} \quad & \mathbb{E}_\phi[\hat{u}(\gamma)] - \underbrace{\lambda H(\omega|\mu)}_{=const.} \\ s.t. \quad & \mathbb{E}_\phi[\gamma] = \mu \end{aligned} \tag{2.11}$$

where $\hat{u}(\gamma) := B(\gamma) + \lambda H(\omega|\gamma)$ is a receiver's value function at posterior γ and $B(\gamma) := \mathbb{E}_\gamma[u(\sigma^*(\gamma), \omega)]$ is a receiver's expected utility at posterior γ under her optimal action strategy $\sigma^*(\gamma) \in \arg\max_{a\in A} \mathbb{E}_\gamma[u(a, \omega)]$. The problem (2.11) has a geometric interpretation. Let

$$U(\gamma) := \sup\{z \in \mathbb{R}|(\gamma, z) \in co(\hat{u})\}, \tag{2.12}$$

where sup denotes supremum and $co(\hat{u})$ denotes the convex hull[20] of the graph $\hat{u}$, be the *concavification* of $\hat{u}$. U is the smallest concave function that is everywhere weakly greater than $\hat{u}$. CD showed that $U(\mu) - \lambda H(\omega|\mu)$ is the receiver's expected utility under her optimal behavior, the receiver learns at μ if and only if $U(\mu) > \hat{u}(\mu)$, and the support of the optimal information strategy, $\text{supp}(\phi^*_\mu)$, are the posterior beliefs that support the tangent hyperplane to the lower epigraph of the concavification U above μ. See Figure 2.1 in Sec. 2.3 for this interpretation in the introductory example.

2.A.2 Characterization of the solution

Convexity of the entropy-based cost function implies that strictly more informative strategies are strictly more costly than less informative such strategies. Hence optimization is inconsistent with the choice of the same action in two distinct posteriors (receiving distinct signals that lead to the same action is inefficient as information is acquired but not acted upon). This implies that given μ, for purposes of optimization, an optimal receiver's strategy (ϕ^*_μ, σ^*) can be specified as a subset of available actions $C_\mu \subseteq A$ (a consideration set)

[20] A convex hull of a set X is the smallest convex set that contains X.

chosen with strictly positive unconditional probabilities $\mathcal{P}_\mu^a := \mathbb{E}_{\phi_\mu^*}(\Pr[\sigma^*(\gamma) = a]) > 0$ and corresponding act-specific posteriors $\gamma_\mu^a := \{\gamma \in \Delta(\Omega) : \gamma \in \text{supp}(\phi_\mu^*) \wedge \sigma^*(\gamma) = a\}$, see Matějka and McKay (2015). Caplin, Dean, and Leahy (2016) provide characterization of the receiver's optimal strategy, which is captured in Definition 6.

Definition 6. Given interim belief μ, a *rational inattentive strategy* at μ (henceforth, RI strategy)—a solution to problem (2.5)—consists of tuples $\{\mathcal{P}_\mu^a\}_{a\in A}$ and $\{\gamma_\mu^a\}_{a\in C_\mu}$, where each action is chosen in at most one posterior, such that $\forall\omega \in \Omega : \mu(\omega) = \sum_{a\in A} \mathcal{P}_\mu^a \gamma_\mu^a(\omega)$, and:

1. *Invariant Likelihood Ratio Equations for Chosen Actions:* given $a, a' \in C_\mu$, and $\omega \in \Omega$,

$$\frac{\gamma_\mu^a(\omega)}{e^{\frac{u(a,\omega)}{\lambda}}} = \frac{\gamma_\mu^{a'}(\omega)}{e^{\frac{u(a',\omega)}{\lambda}}} \tag{2.13}$$

2. *Likelihood Ratio Inequalities for Unchosen Actions:* given $a \in C_\mu$ and $a'' \in A \setminus C_\mu$,

$$\sum_{\omega\in\Omega} \gamma_\mu^a(\omega) \left(\frac{e^{\frac{u(a'',\omega)}{\lambda}}}{e^{\frac{u(a,\omega)}{\lambda}}} \right) \leq 1. \tag{2.14}$$

Applying Lemma 4 to the motivating example in Section 2.3, we can find the (unknown) extreme points of non-learning regions $\underline{\mu}, \overline{\mu}$ by solving $(1 - \underline{\mu})e^{\frac{-1}{\lambda}} + \underline{\mu}e^{\frac{1}{\lambda}} = 1$ and $(1 - \overline{\mu})\frac{1}{e^{\frac{-1}{\lambda}}} + \overline{\mu}\frac{1}{e^{\frac{1}{\lambda}}} = 1$, respectively. Hence, $\underline{\mu} = \frac{1}{1+e^{\frac{1}{\lambda}}}$ and $\overline{\mu} = \frac{e^{\frac{1}{\lambda}}}{1+e^{\frac{1}{\lambda}}}$.

2.A.3 Uniqueness

Generally, the receiver's RI strategy may not be unique at all $\mu \in \Delta(\Omega)$.

Assumption 3. $\{e^{\frac{u(a,\omega)}{\lambda}}, a \in A\}$ *are affine independent. That is, one cannot find scalars* α_a*, not all zero, such that* $\sum_{a\in A} \alpha_a = 0$ *and* $\sum_{a\in A} \alpha_a e^{\frac{u(a,\omega)}{\lambda}} = 0$.

Matějka and McKay (2015) and CD show that A3 is a sufficient condition for uniqueness.

Lemma 6. *If A3 holds, then the receiver's optimal strategy is always unique.*

A3 rules out cases such as a receiver with two duplicate actions giving her the same state-dependent payoffs. It is not very restrictive: when A3 fails, it holds under a slight perturbation of $u(a, \omega)$. When A3 holds, all equilibria are sender-preferred since the receiver is never indifferent between two strategies in an equilibrium.

2.B Different cost function assumption

Here, we solve the introductory example under a different cost function. We assume the receiver can obtain a partially revealing binary signal at a fixed cost $c \geq 0$.

Let $\Omega = \{0, 1\}$, $A = \{0, 1\}$, $v(a, \omega) = a$, $u(a, \omega) = 1$ if $a = \omega = 1$, $u(a, \omega) = -1$ if $a = 1$ and $\omega = 0$ and 0 otherwise. For the purpose of this part of the appendix, we identify all the beliefs with the probability of state $\omega = 1$. Let $\mu_0, \mu, \gamma \in [0, 1]$ be the probability of $\omega = 1$ at a prior, interim, and posterior belief, respectively.

Given μ, the receiver can obtain a binary signal $s \in \{0, 1\}$ of precision $p := \Pr[s = \omega | \omega] > 0.5$ by paying $c \geq 0$. Say *the receiver learns* if she pays c and gets the signal.

Receiver's maximization problem

Given μ, if the receiver learns, she updates her beliefs to a posterior $\gamma_s(\mu) := \Pr[\omega = 1 | s, \mu]$ with probability $\phi_s(\mu) := \Pr[s | \mu]$, where

$$\gamma_1(\mu) = \frac{p\mu}{\phi_1}, \qquad \phi_1(\mu) = p\mu + (1-p)(1-\mu),$$
$$\gamma_0(\mu) = \frac{(1-p)\mu}{\phi_0}, \qquad \phi_0(\mu) = (1-p)\mu + p(1-\mu).$$

The receiver takes action $a = 1$ if and only if $\gamma_s \geq 1/2$. Her expected utility from learning is

$$U^L(\mu) = \max\{0, 2\gamma_1(\mu) - 1\}\phi_1(\mu) + \max\{0, 2\gamma_0(\mu) - 1\}\phi_0(\mu) - c.$$

If she does not learn, she takes action $a = 1$ if and only if $\mu \geq 1/2$, obtaining expected utility

$$U^{NL}(\mu) = \max\{0, 2\mu - 1\}.$$

For sufficiently low cost c, there are two interim beliefs at which the receiver is indifferent between learning and not: $\underline{\mu} < 1/2$ such that upon $s = 1$, the receiver switches to action $a = 1$, but the expected marginal benefit is exactly c: $U^{NL}(\underline{\mu}|\underline{\mu} < 1/2) = U^{L}(\underline{\mu}|\underline{\mu} < 1/2, \gamma_1(\underline{\mu}) \geq 1/2)$; and $\overline{\mu} \geq 1/2$ such that upon $s = 0$, the receiver switches to action $a = 0$, but the expected marginal benefit is exactly c: $U^{NL}(\overline{\mu}|\overline{\mu} \geq 1/2) = U^{L}(\overline{\mu}|\overline{\mu} > 1/2, \gamma_0(\overline{\mu}) < 1/2)$. The first equation is $0 = \left(2\gamma_1(\underline{\mu}) - 1\right)\phi_1(\underline{\mu}) - c$ and the second equation is $2\overline{\mu} - 1 = (2\gamma_1(\overline{\mu}) - 1)\,\phi_1(\overline{\mu}) - c$, yielding

$$\underline{\mu} = 1 - p + c, \quad \overline{\mu} = p - c,$$

where $p - c \geq 1/2$ must hold.

Hence, if $c \leq p - 1/2$, there are two non-learning, $[0, \underline{\mu})$, $[\overline{\mu}, 1]$, and one learning, $[\underline{\mu}, \overline{\mu})$, regions. In contrast to the original model, a non-learning region need not be closed, as the sender-preferred equilibrium assumption puts the belief $\underline{\mu}$ to the learning region. If $c > p - 1/2$, the receiver never learns for any μ.

Sender's maximization problem

Suppose $c \leq p - 1/2$. A seller's conditional expected utility $\hat{v}(\mu)$ is

$$\hat{v}(\mu) = \begin{cases} 0 & 0 \leq \mu < \underline{\mu} \\ p\mu + (1-p)(1-\mu) & \underline{\mu} \leq \mu < \overline{\mu} \\ 1 & \overline{\mu} \leq \mu \leq 1 \end{cases}.$$

A sender's optimal strategy can be found by concavification V of $\hat{v}$. Figure 2.8 depicts $\hat{v}(\mu)$ and an optimal sender's strategy with $p = 0.8$ when $c \to 0$ and $c = 0.2$.

Figure 2.8: $\hat{v}(\mu)$ and the sender's optimal strategy with $p = 0.8$. The sender targets (a) learning (less information) and (b) no learning (more information).

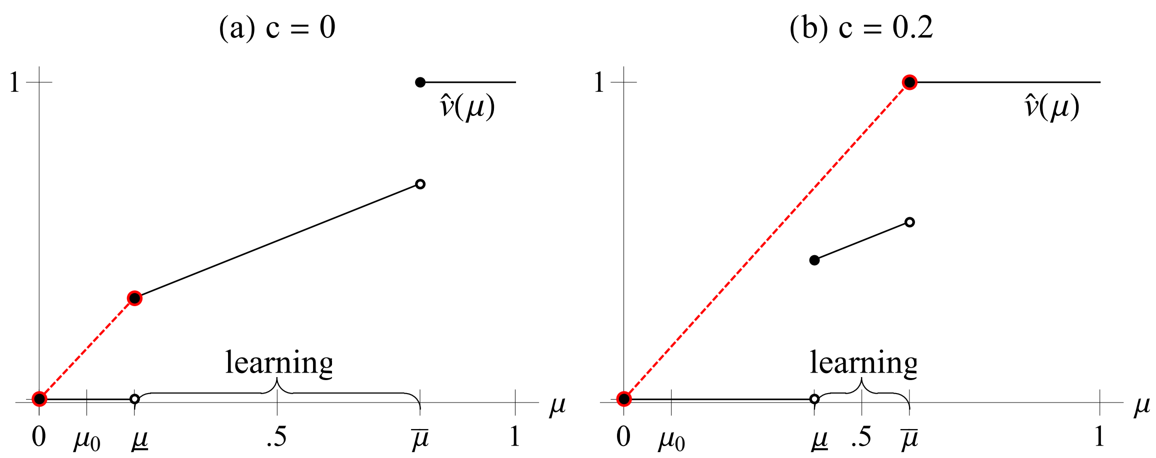
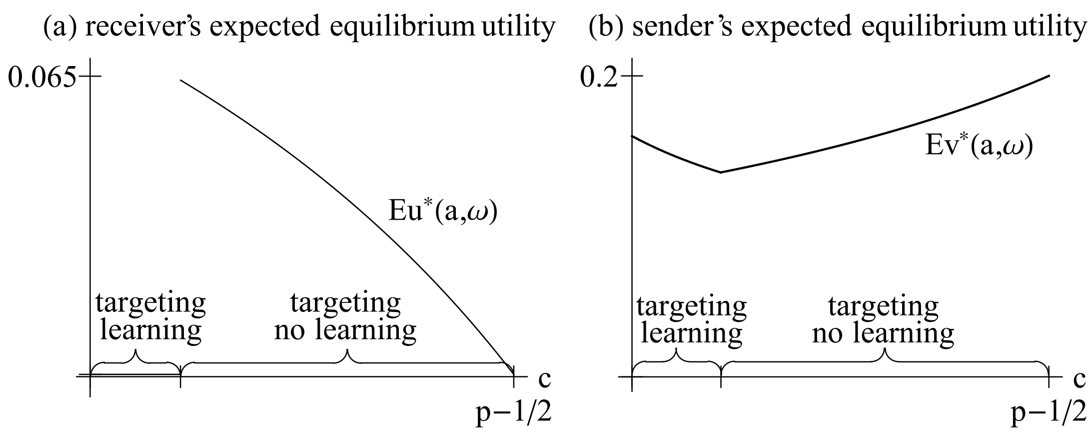


Figure 2.9: Equilibrium expected utilities as a function of c with $p = 0.8$.



Comparison with the original model

First, the key simplification step–the Never-Learning Lemma–*does not hold.* The posterior γ_s can fall into a learning region. If the sender then 'sent' the receiver to γ_s directly, the receiver would learn instead of acting right away, thus changing the outcome of the game. *Second, $\hat{v}(\mu)$ is discontinuous* at the indifference points of learning/not, i.e., at $\{\underline{\mu}, \overline{\mu}\}$. In the original model, such discontinuities do not exist in the introductory example as the threshold beliefs $\{\underline{\mu}, \overline{\mu}\}$ have a different interpretation. They are points toward which the optimal amount of learning gradually shrinks, but at which it is *strictly* optimal not to learn. The sender-preferred assumption is not needed there, because the RI strategy is always unique. *Third, the sender can strictly prefer to induce $a = 1$ indirectly through* *the receiver's learning* and provide her with just enough information so that she learns ($\text{supp}(\tau^*) = \{0, \underline{\mu}\}$ in Figure 2.8 (a)). Under entropy-based cost, the receiver can vary the precision of her information, but here it is fixed, which is taken advantage of. *Fourth, the sender's expected equilibrium utility* $\mathbb{E}v^*(a, \omega)$ *is non-monotone in* c: he strictly prefers low to intermediate cost, see Figure 2.9 (b). When the sender targets $a = 1$ indirectly through the receiver's learning, the required amount of information to induce learning increases with c. *Fifth, the receiver strictly prefers intermediate to high and low cost.* Under low cost, the sender targets the receiver's learning. He provides just enough information so that the receiver is indifferent between learning and not; any benefit is exactly offset by paying c. With sufficiently high cost, however, the sender induces $a = 1$ directly without the receiver's learning; the receiver obtains valuable information without paying any cost. In contrast, under entropy-based cost, the receiver's expected equilibrium utility is always the highest when her information cost parameter $\lambda = 0$. As she can decide on the amount of her information, she always becomes fully informed when $\lambda = 0$, not leaving any room for the sender's manipulation.

2.C Proofs

Proof of Lemma 1

Proof. Let τ be the sender's strategy of value v. Suppose $\exists \mu' \in \text{supp}(\tau)$ and $\mu' \neq \cup_a NL^a$. Then the receiver's optimal information strategy at μ', $\phi^*_{\mu'}$, satisfies: $\mu' \notin \text{supp}(\phi^*_{\mu'})$, μ' lies in the convex hull of $\text{supp}(\phi^*_{\mu'})$ and $\text{supp}(\phi^*_{\mu'}) \subseteq \cup_a NL^a$ (Lemma 2). Then there exists the sender's strategy τ' where $\text{supp}(\tau') = (\text{supp}(\tau)) \setminus \mu') \cup \text{supp}(\phi^*_{\mu'})$, which does not change the distribution of the receiver's actions conditional on the state (since $\text{supp}(\phi^*_{\mu'}) \subseteq \cup_a NL^a$). Hence, the value of τ' is also v. Formally,

$$\tau'(\mu) = \begin{cases} \tau(\mu) & \mu \notin \{\mu'\} \cup \text{supp}(\phi^*_{\mu'}) \\ 0 & \mu = \mu' \\ \tau(\mu) + \tau(\mu')\phi^*_{\mu'}(\mu) & \mu \in \text{supp}(\phi^*_{\mu'}) \end{cases}.$$

□

Proof of Lemma 2

Proof. Note that the functions used here are defined in Appendix 2.A. Let ϕ^*_μ be the receiver's optimal information strategy at μ and suppose that $\exists\gamma' \in \text{supp}(\phi^*_\mu)$ for which $\gamma' \notin \cup_a NL^a$. That is, there exists distribution $\phi_{\gamma'}$ of posterior beliefs γ, $\mathbb{E}_{\phi_{\gamma'}}[\gamma] = \gamma'$, such that

$$\begin{aligned} \mathbb{E}_{\phi_{\gamma'}}[B(\gamma) + \lambda H(\gamma)] - \lambda H(\gamma') &> B(\gamma'), \\ \mathbb{E}_{\phi_{\gamma'}}[B(\gamma) + \lambda H(\gamma)] &> B(\gamma') + \lambda H(\gamma'), \\ \mathbb{E}_{\phi_{\gamma'}}[\hat{u}(\gamma)] &> \hat{u}(\gamma'). \end{aligned}$$

Consider a different receiver's information strategy ϕ', where

$$\phi'(\gamma) = \begin{cases} \phi^*_\mu(\gamma) & \gamma \notin \{\gamma'\} \cup \text{supp}(\phi_{\gamma'}) \\ 0 & \gamma = \gamma' \\ \phi^*_\mu(\gamma) + \phi^*_\mu(\gamma')\phi_{\gamma'}(\gamma) & \gamma \in \text{supp}(\phi_{\gamma'}) \end{cases}.$$

But then ϕ' gives the receiver strictly higher expected utility than ϕ^*_μ, contradicting the optimality of ϕ^*_μ. □

Proof of Lemma 3

Proof. For $\lambda > 0$, the receiver's marginal cost of becoming fully informed is infinity (property of Shannon entropy). Hence, for any $\lambda > 0$, the receiver never decides to become fully informed, implying $\cup_a NL^a \neq \emptyset$ (if $\lambda \to 0$, the non-learning regions are the generic interim beliefs at which the state is fully revealed). The rest of the statement follows from the equation (2.14) of the solution to the receiver's RI problem. □

Let us further state another Lemma that will be used throughout the following proofs.

Lemma 7. *Let $V(\mu)$ be concavification of $\hat{v}(\mu)$ defined as the smallest concave function that is everywhere weakly greater than $\hat{v}$.*

i) If $\mu \in supp(\tau^*)$, then $V(\mu) = \hat{v}(\mu)$.

ii) The sender benefits from persuasion if and only if $\hat{v}(\mu_0) < V(\mu_0)$.

Lemma 7 is an analogy of Lemma 2 and Corollary 2 from KG, which is applicable to our setting when $\hat{v}(\mu)$ modified to our setting is considered.

Proof of Proposition 2

Proof. First, let us show that the set $\cup_a EP^a$ is non-empty and finite. Lemma 3 shows $\cup_a NL^a \neq \emptyset$. Let a be an action for which $NL^a \neq \emptyset$. By the Krein-Milman Theorem and (2.9)—showing convexity of NL^a—NL^a is a closed convex hull of its extreme points; $EP^a \neq \emptyset$. As NL^a is an intersection of the simplex $\Delta(\Omega)$ and a collection of half-spaces $\left\{\mu \in \mathbb{R}^\Omega : \sum_{\omega\in\Omega} \mu(\omega) \left(\frac{e^{\frac{u(a',\omega)}{\lambda}}}{e^{\frac{u(a,\omega)}{\lambda}}}\right) \leq 1 \;\; \forall a' \neq a\right\}$, the set EP^a is finite. As the action space A is finite, then $\cup_a EP^a \neq \emptyset$ and is also finite.

Second, let τ^* be the sender's optimal strategy under which the receiver never learns. Suppose $\exists \mu' \in \text{supp}(\tau^*)$: $\mu' \in NL^a \setminus EP^a$. Then there exists a subset $X \subseteq EP^a$ where μ' lies in the convex hull of X. Hence, there exists another sender's strategy τ' where $\text{supp}(\tau') = (\text{supp}(\tau') \setminus \mu') \cup X$. If the chosen action at some belief of X is different from $\sigma^*(\mu')$, the action chosen at μ', it can only lead to the sender's higher expected utility by the sender-preferred assumption.

Third, let τ^* be the sender's optimal strategy and suppose $|\text{supp}(\tau^*)| > |\Omega|$. Then $\text{supp}(\tau^*)$ supports the tangent hyperplane to the lower epigraph of the concavification above prior. Such hyperplane is defined by any $|\Omega|$ different points it contains. By Carathéodory's theorem, there exists a subset $C \subset |\text{supp}(\tau^*)|$ with $|C| \leq |\Omega|$ such that the prior belief μ_0 lies in the convex hull of C. Hence, there exists a sender's strategy τ' with $\text{supp}(\tau') = C$. As $\text{supp}(\tau')$ supports the tangent hyperplane to the lower epigraph of the concavification above prior, τ' is thus also optimal. □

Proof of Proposition 3

Proof. Let A2 hold, $A = \Omega = \{0, 1\}$, and suppose the sender benefits from persuasion. Let $\mu_0, \mu \in [0, 1]$ be the probability of $\omega = 1$ at the prior and interim belief, respectively. For any $\lambda > 0$, there are two non-learning regions with $EP^0 = \{0, \underline{\mu}\}$, $EP^1 = \{\overline{\mu}, 1\}$

where $0 < \underline{\mu} \leq \overline{\mu} < 1$. Note that $\hat{v}(\mu)$ is a piecewise-linear function, with linear segments over $[0, \underline{\mu}]$, $[\underline{\mu}, \overline{\mu}]$, and $[\overline{\mu}, 1]$. Without the loss of generality, we can consider the sender's strategies that induce at most 2 different interim beliefs.

Part i)

The concavification $V(\mu)$ of $\hat{v}(\mu)$ can have four forms:

i) $v(\mu) = V(\mu)$ if $\mu \in [\overline{\mu}, 1] \cup \{0\}$ and $v(\mu) < V(\mu)$ otherwise,

ii) $v(\mu) = V(\mu)$ if $\mu \in [0, \underline{\mu}] \cup \{1\}$ and $v(\mu) < V(\mu)$ otherwise,

iii) $v(\mu) = V(\mu)$ if $\mu \in \{0\} \cup \{1\}$ and $v(\mu) < V(\mu)$ otherwise,

iv) $v(\mu) = V(\mu)$ if $\mu \in [0, 1]$.

Note that the in iv), the sender does not benefit from persuasion for any prior $\mu_0 \in [0, 1]$. Suppose, contrary to the proposition, that there exists a sender's optimal strategy τ^* with $\tilde{\mu} \in \text{supp}(\tau^*)$ and $\tilde{\mu} \in (\underline{\mu}, \overline{\mu})$—the receiver learns at $\tilde{\mu}$. Then $\hat{v}(\tilde{\mu}) = V(\tilde{\mu})$ (Lemma 7). Since $\hat{v}(\mu)$ is linear over $[\underline{\mu}, \overline{\mu}]$, this implies that $\hat{v}(\mu) = V(\mu)$ for all $\mu \in [\underline{\mu}, \overline{\mu}]$. But then, only case iv) can happen. In particular, $\hat{v}(\mu_0) = V(\mu_0)$, which contradicts with the sender benefitting from persuasion.

Part ii)

Let us prove the *only if* part. From part i) the receiver never learns in an equilibrium. Let A2 hold and let τ^* be an optimal sender's strategy. From Proposition 5, the shape of $\hat{v}(\mu)$ and the fact that the sender benefits from persuasion, we have $\text{supp}(\tau^*) = \{\mu_l, \mu_r\} \in \{\{0, \overline{\mu}\}, \{\underline{\mu}, 1\}, \{0, 1\}\}$. Note that each pair is a point on the frontier of NL^0 and NL^1. Suppose, contrary to the proposition, there are two different optimal sender's strategies. Then there is a non-learning region of one of the actions such that one strategy induces belief on one frontier and the other strategy induces belief on the other frontier of that non-learning region. But then a new strategy that would instead, ceteris paribus, induce a convex combination of these two beliefs would also be optimal (since the convex combination still leads to the same action). However, such a new belief lies inside the non-learning region, which contradicts Proposition 5.

Let us prove the *if* part. Suppose there is a unique equilibrium, let τ^* be the optimal sender's strategy, but, contrary to the proposition, A2 does not hold. Without the loss of generality, let $a = 0$ be the action that does not satisfy A2. That is, $\forall \mu \in [0, 1]$: $\hat{v}(\mu) \leq (1 - \mu)v(0, 0) + \mu v(0, 1)$ and $\exists \mu \in [\overline{\mu}, 1]$: $\hat{v}(\mu) = (1 - \mu)v(0, 0) + \mu v(0, 1)$. Then it is either $\hat{v}(\overline{\mu}) = (1 - \overline{\mu})v(0, 0) + \overline{\mu}v(0, 1)$, $\hat{v}(1) = v(0, 1)$, or both. Since the equilibrium is unique and the sender benefits from persuasion, it must be $\text{supp}(\tau^*) = \{\mu_l, \mu_r\} \in \{\{0, \overline{\mu}\}, \{\underline{\mu}, 1\}, \{0, 1\}\}$ (from Proposition 2).

Suppose $\hat{v}(\overline{\mu}) = (1 - \overline{\mu})v(0, 0) + \overline{\mu}v(0, 1)$. Then $\hat{v}(\mu)$ is linear over the whole $[0, \overline{\mu}]$ and $\text{supp}(\tau^*) \neq \{0, \overline{\mu}\}$ (since then the sender would not benefit from persuasion from Lemma 7). Furthermore, as $\hat{v}(1) \leq v(0, 1)$, then $\text{supp}(\tau^*) \notin \{\{\underline{\mu}, 1\}, \{0, 1\}\}$, because under such strategies he cannot benefit from persuasion.

Suppose $\hat{v}(\overline{\mu}) < (1 - \overline{\mu})v(0, 0) + \overline{\mu}v(0, 1)$ and $\hat{v}(1) = v(0, 1)$. Then the values $\hat{v}(0)$, $\hat{v}(\underline{\mu})$, $\hat{v}(1)$ lie on the same line, which is thus the concavification V. Therefore, $\hat{v}(0) = V(0)$, $\hat{v}(\underline{\mu}) = V(\underline{\mu})$, $\hat{v}(1) = V(1)$ and $\hat{v}(\overline{\mu}) < V(\overline{\mu})$. Hence, $\text{supp}(\tau^*) \neq \{0, \overline{\mu}\}$ from Lemma 7. But then $\text{supp}(\tau^*) = \{\underline{\mu}, 1\}$ if and only if $\text{supp}(\tau^*) = \{0, 1\}$, contradicting the uniqueness of an equilibrium. $\square$

Proof of Proposition 4

Proof. We can follow the proof of Proposition 4 in KG applied to our setting. $\square$

Proof of Proposition 5

Proof. Let A2 hold and suppose the sender benefits from persuasion. Proposition 5 of KG, applied to our setting, implies that for any optimal sender's strategy τ^* we have $\text{supp}(\tau^*) \cap \cup_{a \in A} \text{int}(NL^a) = \emptyset$, where $\text{int}(NL^a)$ denotes the interior of NL^a. $\square$

Proof of Proposition 6

Proof. Focusing on the sender's strategies under which the receiver never learns is sufficient (Lemma 1). We show that the non-learning regions do not shrink as λ increases. As the sender can choose from the same (or possibly even bigger) set of strategies, he never becomes strictly worse off as λ increases.

Given μ, let s be a particular receiver's strategy at μ, $EV(s)$ be a gross expected receiver's utility under s, and $I(s;\omega|\mu)$ be mutual information based on Shannon entropy associated with s. Let $\lambda \geq 0$. Suppose the receiver does not learn at μ. Let s_{NL} be the receiver's optimal non-learning strategy at μ. Let s_L be an arbitrary strategy with strictly positive learning at μ. We have

$$EV(s_{NL}) - \lambda I(s_{NL};\omega|\mu) \geq EV(s_L) - \lambda I(s_L;\omega|\mu).$$

Let $\lambda' > \lambda$. Then $\lambda I(s_{NL};\omega|\mu) = \lambda' I(s_{NL};\omega|\mu) = 0$ (no-learning costs zero) and $\lambda I(s_L;\omega|\mu) < \lambda' I(s_L;\omega|\mu)$. Hence, $EV(s_{NL}) - \lambda' I(s_{NL};\omega|\mu) > EV(s_L) - \lambda' I(s_L;\omega|\mu)$, showing that no-learning strategy remains optimal at μ. $\square$

Proof of Proposition 7

Proof. Let A2 hold, $A = \Omega = \{0, 1\}$, and suppose the sender benefits from persuasion $\forall\lambda > 0$. Let $\mu_0, \mu \in [0, 1]$ be the probability of $\omega = 1$ at the prior and interim belief, respectively. For $\lambda > 0$, there are two non-learning regions with $EP^0 = \{0, \underline{\mu}(\lambda)\}$, $EP^1 = \{\overline{\mu}(\lambda), 1\}$ where $0 < \underline{\mu}(\lambda) \leq \overline{\mu}(\lambda) < 1$. For each $\lambda > 0$ there is a unique sender's optimal strategy τ^*_λ, where $\text{supp}(\tau^*_\lambda) = \{\mu_l, \mu_r\} \in \{\{0, \overline{\mu}(\lambda)\}, \{\underline{\mu}(\lambda), 1\}, \{0, 1\}\}$ (Proposition's 2 and 3). The receiver never learns under τ^*_λ.

1. If $\text{supp}(\tau^*_\lambda) = \{0, 1\}$, the receiver obtains complete information. If there is any change in the sender's strategy as a result of an increase in λ, the receiver can only be worse off.

2. Since the non-learning regions do not shrink as λ increases (proof of Proposition 6), we have $\frac{\partial \underline{\mu}(\lambda)}{\partial \lambda} \geq 0$ and $\frac{\partial \overline{\mu}(\lambda)}{\partial \lambda} \leq 0$. Therefore, the sender's strategies inducing either always $\{0, \overline{\mu}(\lambda)\}$ or always $\{\underline{\mu}(\lambda), 1\}$ are (weakly) Blackwell less informative as λ increases. Without the loss of generality, assume $\text{supp}(\tau^*_{\lambda_1}) = \{0, \overline{\mu}(\lambda_1)\}$. Consider any $\lambda_2 > \lambda_1$. We show that inducing $\{0, \overline{\mu}(\lambda_2)\}$ remains optimal when the receiver's marginal cost of information is λ_2. Let $\hat{v}_{\lambda_i}, V_{\lambda_i}$ denote $\hat{v}$ and its concavification V, respectively, when the marginal cost of information is λ_i. We have $0 < \underline{\mu}(\lambda_1) \leq \underline{\mu}(\lambda_2) \leq \overline{\mu}(\lambda_2) \leq \overline{\mu}(\lambda_1) < 1$. From Lemma 7 and the uniqueness of the sender's strategy, it must be the case that the line connecting $\hat{v}_{\lambda_1}(0)$ and

$\hat{v}_{\lambda_1}(\overline{\mu}(\lambda_1))$ is strictly above the line connecting $\hat{v}_{\lambda_1}(\underline{\mu}(\lambda_1))$ and $\hat{v}_{\lambda_1}(\overline{\mu}(\lambda_1))$. Based on the shape of $\hat{v}$ in this setting and the fact that $0 < \underline{\mu}(\lambda_2) \leq \overline{\mu}(\lambda_2)$, this then implies that the line connecting $\hat{v}_{\lambda_1}(0)$ and $\hat{v}_{\lambda_2}(\overline{\mu}(\lambda_2))$ will also be strictly above the line connecting $\hat{v}_{\lambda_2}(\underline{\mu}(\lambda_2))$ and $\hat{v}_{\lambda_2}(\overline{\mu}(\lambda_2))$. Hence $\hat{v}_{\lambda_2}(\underline{\mu}(\lambda_2)) \neq V_{\lambda_2}(\underline{\mu}(\lambda_2))$. A similar logic applies to showing that inducing $\{0, 1\}$ is not optimal either.

$\square$

Chapter 3

Habit Formation: An Experimental Study

Co-authored with Keh-Kuan Sun (Washington University in St. Louis)[1].

3.1 Introduction

When faced with the same choices repeatedly, people often form habits, i.e. tendencies to excessively repeat previous actions. In macroeconomics, a utility function that depends both on current and previous choices often explains empirical puzzles such as the equity premium puzzle (Constantinides 1990; Jermann 1998), and fits the data better than models with standard preferences (Carroll, Overland, and Weil 2000; Fuhrer 2000)[2].

In this paper, we ask why habits are formed. We consider a stochastic environment

[1]This study is a part of a larger project on the experimental testing of behavior in dynamic information acquisition tasks jointly with Brian Rogers (Washington University in St. Louis), Jakub Steinr (CERGE-EI), and Keh Sun (Washington University in St. Louis). This study provides some of the project's initial results.

[2]Other examples include a 'status quo bias', a tendency to maintain the status quo (Samuelson and Zeckhauser 1988), 'patient inertia', failure of patients to initiate treatment even after the diagnosis of a medical problem (Suri et al. 2013), 'clinical inertia', failure of health care providers to intensify therapy having unattained treatment goals (Okonofua et al. 2006; Phillips et al. 2001), or 'social inertia', the resistance to change in societies (Bourdieu 1985).

with imperfect stochastic information. We define habit as action inertia that is not explained by variations in optimal actions. One explanation argues that habits are driven by preferences, implicitly assuming people face some (mental) 'switching costs' (e.g. Abel (1990), Carroll, Overland, and Weil (2000), Constantinides (1990)). Alternatively, habits can arise as an adaptation rather than being a hard-wired feature. An agent who learns about a changing environment with positively serially correlated payoff states can save on learning costs by repeating their earlier actions. Such agents can then exhibit behavior *as if* they had switching costs (Steiner, Stewart, and Matějka 2017). Habit formation is then information driven and varies with specific features of the environment. However, it is invariant to these variations under fixed switching costs.

Recognizing whether habits are preference- or information-based is important for the evaluation of policy interventions. For instance, consider the impact of monetary policies, which operate through price stickiness (resulting from habits in pricing). If the state of the economy is more volatile (less serially correlated) during economic crises, then price-stickiness is low during crises under the information-based explanation. The monetary policy is less relevant during the crises when it is needed the most. However, this prediction does not hold for the preference-based explanation.

We test habit formation by observing how subjects take actions in binary perceptual decision problems over two periods. Past actions can have an informational value in these problems: if optimal actions are serially correlated, then past actions can contain useful information about a current optimal action. We vary this informational value. We test whether subjects form a habit only when there is such value, but not otherwise.

The main focus of our study is decision problems with positively serially correlated optimal actions. Two treatments are considered: (i) subjects freely observe feedback revealing past optimal actions; or (ii) no such feedback is provided. In (i), a past action does not contain useful information on a current optimal action since the feedback provides more precise information. Information-driven habits form only in (ii). Preference-driven habits, on the other hand, form in both (i) and (ii). We found that the subjects' behavior is not consistent with the preference-based explanation and it supports the information-based explanation—habits are formed only in (ii). The subjects also face two additional treatments (feedback and no feedback variations) with serially *independent* optimal actions. We examine whether subjects treat independent decision problems independently, which is supported by the data. We do not observe action inertia when the optimal actions are serially independent, regardless of the information feedback.

Our paper studies repeated behavior in dynamic information acquisition problems under no explicit cost of information. The paper closest to ours is by Khaw, Stevens, and Woodford (2017). Their subjects trace a state variable whose properties change at unknown random times. They focus on choice inertia and discreteness in adjustment. They argue that a rational inattention model (Sims 2003)—a model of costly information processing—provides a better fit than a model with constant switching costs. Our paper complements theirs by offering directly testable qualitative predictions for each theory instead of fitting different models to the data.

Closely related is an emerging literature on information acquisition under no explicit cost in *static* problems, which focuses on the rationalizability of subjects' behavior by a general model of costly information processing (Caplin and Dean 2015; Oliveira et al. 2017), by a model with more restricted information-processing costs (Dean and Neligh 2017; Dewan and Neligh 2017), and by a specific model with entropy-based information-processing costs (Caplin and Dean 2013; Cheremukhin, Popova, and Tutino 2015). Recently, Caplin, Csaba, and Leahy (2018) directly recover subjects' information-processing costs from the data instead of testing behavioral predictions implied by a certain model. Usually, the behavior can be rationalized by some model of costly information processing, but is not consistent with the entropy-based cost. We extend the design of Caplin and Dean (2013), Caplin and Dean (2015), and Dean and Neligh (2017) to a dynamic environment. We ask whether a dynamic feature of subjects' behavior, the *action inertia*, is consistent with a general model of costly information processing.

3.2 Hypotheses

Is habit a mechanical consequence of past decisions? Or, is habit a functional adaptation that is formed only if it is advantageous to the decision-maker? We hypothesize that two channels can result in habits: (i) a mechanical, preference-driven, channel; and/or (ii) an adaptive, information-driven, channel. We offer an experiment with four different treatments giving different behavioral predictions for (i) and (ii).

Table 3.1: Hypotheses about habit formation via preference-based (H_0, H_0') and information-based (A) channels.

H_0: A habit is formed in all four treatments.

H_0': A habit is formed in treatments FC and NC. A habit is not formed in treatments FI and NI.

A: A habit is formed in treatment NC. A habit is not formed in treatments FI, NI, and FC.

An agent faces the same binary task over two periods, $t = 1, 2$. At each t, a binary payoff state is realized. Correctly identifying the state yields a strictly positive reward. Before each action, the agent can acquire information about the realized state. Four treatments are considered. Both state realizations are equally likely at $t = 1$ in all treatments. In *independent* treatments (I), both state realizations are equally likely at $t = 2$, independently of the past state. In *correlated* treatments (C), the states are positively serially correlated. In *feedback* treatments (F), an additional free feedback revealing the past state before $t = 2$ is provided. In *no feedback* treatments (N), no such feedback is provided. This gives four treatments: FI, FC, NI, NC.

Denote a_t, θ_t the action and the state at t.

Definition 7. We say *a habit is formed* if, controlling for variations in the states, a_1 has a predictive value on a_2. That is, $\Pr[a_2|\theta_1, \theta_2, a_1] \neq \Pr[a_2|\theta_1, \theta_2]$. Otherwise, we say *a habit is not formed.*

Table 3.1 states behavioral predictions based on preference-driven (hypotheses H_0 and H_0'), and information-driven (hypothesis A) channels[3].

First, consider an agent with preference for action repetition. She incurs disutility, 'mental' switching costs, whenever $a_2 \neq a_1$. The agent has only imperfect stochastic information about the states, which results in stochastic choice. In all four treatments, the switching costs create asymmetry in payoffs at $t = 2$ with a bias toward action repetition, beyond what can be explained by imperfectly tracking the evolution of states (H_0). The hypothesis H_0' is a version for more flexible switching cost: the agent has zero switching

[3] An agent with both a preference- and information-driven channel behaves as in H_0 and H_0'.

costs for independent decision problems ("each day is a completely new day"), but positive switching cost otherwise.

Second, consider an agent with zero switching costs, but for whom it is costly to learn and process information. When her past behavior is informative about the current optimal action, she may decide to repeat past actions in order to save on learning costs. The behavior then resembles an agent with switching costs (Steiner, Stewart, and Matějka 2017). However, whether the agent's own past behavior is informative or not varies with treatments. For serially independent optimal actions (treatments FI, NI), the past action contains no information on the current optimal action. For positively serially correlated optimal actions under feedback (treatment FC), feedback reveals more precise information than past actions. Only for positively serially correlated optimal actions under no feedback (treatment NC), the agent's own past action is informative about the current optimal action.

3.3 Experimental design

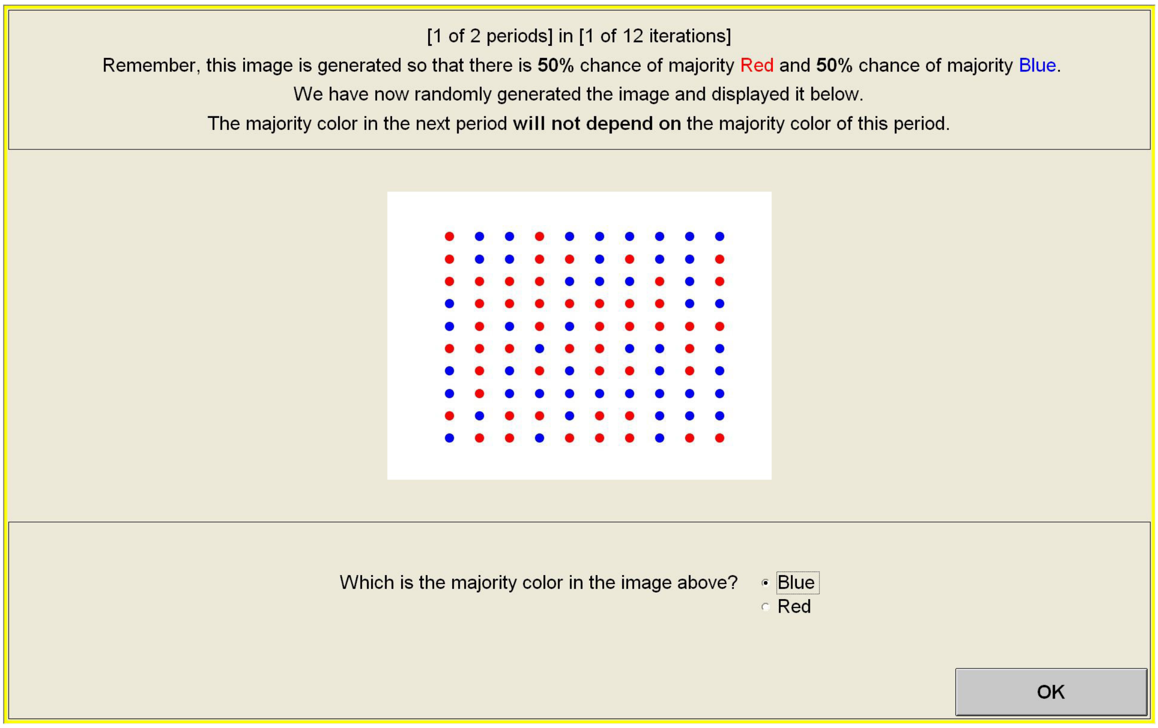
Our design follows Caplin and Dean (2015). Subjects are presented with a screen with 100 red and blue balls. There are two possibilities (states of the world): either 51 red and 49 blue or 51 blue and 49 red balls are displayed. The position of the balls is random. Subjects are asked to determine a prominent color, see Fig. 3.1.

The subjects do not face any explicit cost of information and can perfectly learn the realized state. Information cost stems from real cognitive effort and time. For technical reasons, a 45s time limit is imposed, after which the screen with the balls disappears and the decision is forced[4].

We recruited 41 subjects from the University of California, Santa Barbara over 2 sessions. Participants were randomly assigned to numbered stations in the laboratory. In each session, subjects faced 4 different treatments. Each treatment consisted of 12 iterations; each iteration consisted of 2 periods of ball-counting tasks. In each iteration, both state realizations were equally likely at $t = 1$. The treatments differed in: (i) a serial correlation between the states within an iteration, and (ii) free informational feedback.

[4] The time limit was set to ensure a reasonable ending time for the whole experiment, but so that the subjects do not feel pressured by the time limit.

Figure 3.1: A representative screenshot of a ball-counting decision problem



First, given an iteration, let $\rho = \Pr[\theta_1 = \theta_2|\theta_1]$. In *independent* treatments (I), $\rho = 1/2$—the states are serially independent. In *correlated* treatments (C), $\rho = 3/4$—the states are positively serially correlated. The states were drawn from a given distribution by a computer.

Second, in *feedback* treatments (F), the subjects were shown the realized state immediately after *each* taken *action*. Hence, they perfectly learned the realization θ_1 before entering $t = 2$, independently of their learning effort at $t = 1$. In *no feedback* treatments (N), the subjects were shown the realized states only after *actions* in *both* periods were taken. Hence, at the beginning of $t = 2$ they could draw inference only from information *actively* gathered at $t = 2$. The four treatments are captured in Table 3.2. Prior to starting each treatment, subjects were informed about the state generating process and feedback specification. The order of the treatments was: FI, FC, NI, NC (session 1) and NI, NC, FI, FC (session 2). Within a treatment, each subject faced the same sequence of images.

Prior to making their decisions, subjects were informed that their payoff would proba-

Table 3.2: Four different treatments

	feedback	no feedback
$\rho = 1/2$ (independent states)	FI	NI
$\rho = 3/4$ (correlated states)	FC	NC

bilistically depend on their choices. At the end of the experiment, for each subject one decision problem was selected at random and the reward based on the subject's answer to the selected decision problem was added to a \$10 show-up fee. The additional reward was \$10 for a correct answer and \$0 otherwise. Hence, the expected additional payoff for a correct answer was $\$\frac{10}{96} \approx \$0.1$ in each decision problem[5]. Instructions are in the Appendix.

3.4 A Theoretical example

In this section, we derive the above hypotheses in a theoretical model. We consider Sims (2003) model of information acquisition with entropy based information-processing costs and (mental) switching costs.

An agent chooses an action $a_t \in \{0, 1\}$ at $t = 1, 2$. Let $\theta_t \in \{0, 1\}$ denote a payoff-relevant state. The states are symmetrically distributed: θ_1 is equally likely to be 0 or 1, and, whatever the realized value of θ_1, the probability that $\theta_2 = \theta_1$ is $\rho \geq 1/2$. When $\rho = 1/2$, the states are serially independent (FI and NI treatments); when $\rho > 1/2$, they are positively serially correlated (FC and NC treatments).

Before choosing an action at each t, the agent can acquire costly information about the realized state θ_t. There is a fixed realization signal space X satisfying $2 \leq |X| < \infty$. A signal π_t consists of a family of distributions $\{\pi_t(\cdot|\theta_t)\}_{\theta_t \in \{0,1\}}$ over X, such that $0 \leq \pi_t(x_t|\theta_t) \leq 1$ for each x_t, θ_t and $\sum_{x_t \in X} \pi_t(x_t|\theta_t) = 1$ for each θ_t. The agent chooses a signal π_t and observes a signal realization $x_t \in X$. In addition, after choosing a_1, she observes a costless signal realization y from a finite signal space Y distributed according to a given $\{g(\cdot|\theta_1)\}_{\theta_1 \in \{0,1\}}$, which is assumed to be independent of a_1 and x_1. Two cases

[5] In Caplin and Dean (2013), the expected payoff per correct answer in each decision problem varied from \$0.01 to \$0.15.

are considered: (i) there is no costless information, corresponding to $|Y| = 1$ (treatments NC, NI); or (ii) the costless information perfectly reveals the realization θ_1, corresponding to $|Y| = \{0, 1\}$ and $g(y = \theta_1|\theta_1) = 1$ for all θ_1 (treatments FC, FI). Let $s^1 = \{x_1\}$ and $s^2 = \{x_1, x_2, y\}$. We refer to s^t as a *signal history* at t. At each t, the agent forms a posterior $\mu_t^{s^t} = \Pr[\theta_t = 1|s^t]$ using Bayes' law and then she takes an action a_t. At each t, the agent's strategy consists of an *information* strategy π_t, and an *action* strategy $\sigma_t : [0, 1] \to \{0, 1\}$, where $\sigma_t(\mu_t^{s^t})$ states a chosen action a_t at a posterior $\mu_t^{s^t}$.

In each period in which the agent matches the action to the state, she obtains a reward $R > 0$, and 0 otherwise. Further, whenever $a_2 \neq a_1$, she has to pay a 'mental' switching cost $0 \leq c < R$. There is no discounting. Her payoffs u_t are thus

$$u_1(a_1, \theta_1) = \begin{cases} R & \text{if } a_1 = \theta_1 \\ 0 & \text{if } a_1 \neq \theta_1 \end{cases}, \quad u_2(a_1, a_2, \theta_2, c) = \begin{cases} R - c|a_1 - a_2| & \text{if } a_2 = \theta_2 \\ -c|a_1 - a_2| & \text{if } a_2 \neq \theta_2 \end{cases}.$$

As is standard in the literature on rational inattention, we assume entropy-based information-processing cost. For a binary random variable $T \in \{0, 1\}$ distributed according to $p = \Pr[T = 1] \in [0, 1]$, the entropy is

$$H_T(p) = -(p \ln p + (1 - p) \ln(1 - p)), \tag{3.1}$$

where $0 \ln 0 = 0$ by convention. It is a measure of uncertainty about T at distribution p. Let μ_t^0 denote the prior belief at the beginning of period t about θ_t. We have $\mu_1^0 = \Pr[\theta_1 = 1] = 1/2$ in all treatments, and

$$\mu_2^0 = \Pr[\theta_2 = 1|s^1, y] = \begin{cases} \rho y + (1 - \rho)(1 - y) & \text{in FC and FI treatments} \\ \rho \mu_1^{s^1} + (1 - \rho)(1 - \mu_1^{s^1}) & \text{in NC and NI treatments} \end{cases}. \tag{3.2}$$

where y equals the realization θ_1 in FC and FI treatments. At any prior μ_t^0, we assume that the cost of signal π_t is proportional to the conditional mutual information[6]

$$I(\theta_t; \pi_t|\mu_t^0) = H_{\theta_t}(\mu_t^0) - \mathbb{E}_{x_t \sim \pi_t}[H_{\theta_t}(\mu_t^{s^t})] \tag{3.3}$$

between the signal π_t and the state θ_t. Before observing x_t, the agent's level of uncer-

[6] For more on Shannon entropy and mutual information, see Cover and Thomas (2006).

tainty about θ_t is given by $H_{\theta_t}(\mu_t^0)$. After observing x_t, it is $H_{\theta_t}(\mu_t^{s^t})$. Given μ_t^0, the conditional mutual information thus captures how much uncertainty about θ_t is expected to be reduced after observing a signal realization x_t drawn from π_t.

The agent solves the following problem[7]:

Definition 8. The agent's maximization problem is

$$\max_{\sigma_1,\sigma_2,\pi_1,\pi_2} \mathbb{E}\left[u_1(\sigma_1(\mu_1^{s^1}),\theta_1) + u_2(\sigma_1(\mu_1^{s^1}),\sigma_2(\mu_2^{s^2}),\theta_2,c) - \lambda\left(\sum_{t=1,2} I(\theta_t;\pi_t|\mu_t^0)\right)\right] \quad (3.4)$$

where $\lambda > 0$ is an information cost parameter, and the expectation is taken with respect to the distribution of the sequences $(\theta_t, \mu_t^{s^t})_t$ and μ_2^0 induced by the prior μ_1^0 together with the strategies $(\sigma_1, \sigma_2, \pi_1, \pi_2)$ and the distributions g of a costless signal.

We impose a regularity condition that the agent learns strictly a positive amount of information in both periods: $I(\theta_t; \pi_t^*|\mu_t^0) > 0$ for μ_1^0 and for each realized μ_t^0 induced by π_1^* or feedback, where π_t^* is part of the solution to (3.4)[8].

Under perfect information, $a_t = \theta_t$ at each t. However, a marginal cost of perfect information is infinite (a property of the cost function) and a marginal benefit is finite. Hence, the agent never chooses to become perfectly informed for any t. The agent's behavior is thus stochastic reflecting randomness in the obtained signal realizations and she sometimes makes mistakes. Recall Definition 7. When a habit is formed, the mistakes at $t = 2$ are not fully random and can be partly predicted by past action.

Proposition 8.

i) Let $c > 0$ for all $\rho \geq 1/2$. Then a habit is formed in all treatments. That is, $\Pr[a_2|\theta_1, \theta_2, a_1] \neq \Pr[a_2|\theta_1, \theta_2]$ in all treatments.

ii) Let $c > 0$ for all $\rho > 1/2$ and $c = 0$ for $\rho = 1/2$. Then a habit is formed in treatments NC and FC, and a habit is not formed in treatments NI and FI. More specifically, $\Pr[a_2|\theta_1, \theta_2, a_1] \neq \Pr[a_2|\theta_1, \theta_2]$ in treatments NC and FC, and $\Pr[a_2|\theta_1, \theta_2, a_1] = \Pr[a_2|\theta_2]$ in treatments NI and FI.

[7] Alternatively, we can assume a myopic agent who does not internalize the continuation value of information. She chooses $\{\pi_t, \sigma_t\}$ in order to maximize the static expected gross payoffs net of any costs in that period only.

[8] The data shows that variations in a_2 are predominantly explained by variations in θ_2, suggesting subjects were learning in period 2 (see Section 3.5.2).

iii) Let $c = 0$ for all $\rho \geq 1/2$. Then a habit is formed in treatment NC, and a habit is not formed in treatments NI, FI, and FC. More specifically, $\Pr[a_2|\theta_1, \theta_2, a_1] = \Pr[a_2|\theta_2, a_1] \neq \Pr[a_2|\theta_2]$ *in NC treatment,* $\Pr[a_2|\theta_1, \theta_2, a_1] = \Pr[a_2|\theta_1, \theta_2] \neq \Pr[a_2|\theta_2]$ *in FC treatment, and* $\Pr[a_2|\theta_1, \theta_2, a_1] = \Pr[a_2|\theta_2]$ *in NI and FI treatments.*

Proposition 8 specifies when a habit is formed under optimal behavior. Part (i) considers an agent with inflexible switching costs: $c > 0$ for all $\rho \geq 1/2$. Part (ii) considers an agent with more flexible switching costs: $c > 0$ only when the states are positively serially correlated. Part (iii) considers an agent with zero switching costs independently of serial correlation in states. Under strictly positive switching costs, a_1 can always explain some variations in a_2 after controlling for the states. Under zero switching costs, this remains to hold only in NC treatment. More specifically, a_2 is then predicted only by variations in θ_2 in NI and FI treatments, and it is predicted by variations in both θ_1 and θ_2 in FC treatment.

Part (i) follows from asymmetry in u_2 depending on a past action with a bias toward action repetition, ceteris paribus. Part (iii) follows from (3.2) describing the relationship between μ_2^0 and $\mu_1^{s^1}$ (associated with optimal action $\sigma_1^*(\mu_1^{s^1})$) and μ_2^0 and the costless signal y (when it is informative). When $\rho = 1/2$, then $\mu_2^0 = 1/2$ regardless of the distribution specification of the costless signal: the states are serially independent and μ_2^0 is independent of θ_1 and $\sigma_1^*(\mu_1^{s^1})$. When $\rho > 1/2$, then μ_2^0 depends on θ_1 but not on $\sigma_1^*(\mu_1^{s^1})$ when the past state is freely observed and on $\sigma_1^*(\mu_1^{s^1})$ but not on θ_1 when the past state is not freely observed. Part (ii) is then implied by (i) and (iii). See the Appendix for the proof.

3.5 Results

3.5.1 Data summary

We recruited 41 subjects over two sessions; 20 subjects in session 1 and 21 subjects in session 2. There were 4 treatments with 12 iterations of 2-period decision problems. Each subject answered 96 decision problems. For each treatment, we thus have 492 observations per period about the subjects' choices.

Figure 3.2: Aggregate precisions in period 1 (left) and period 2 (right).

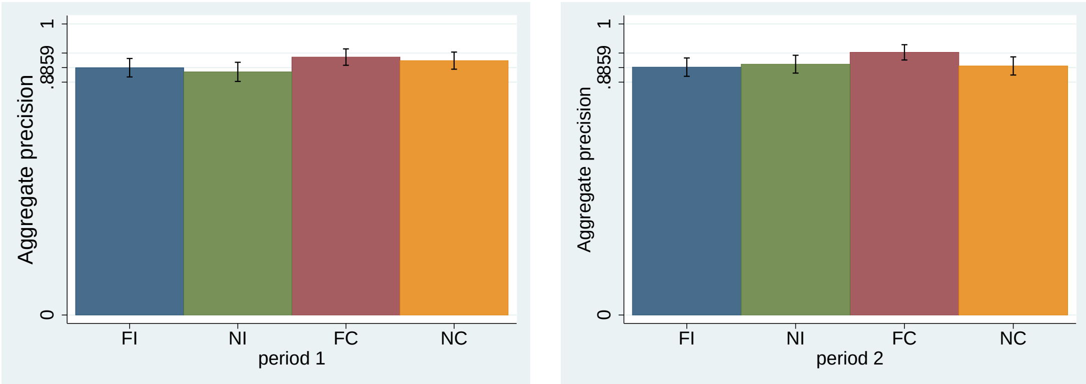
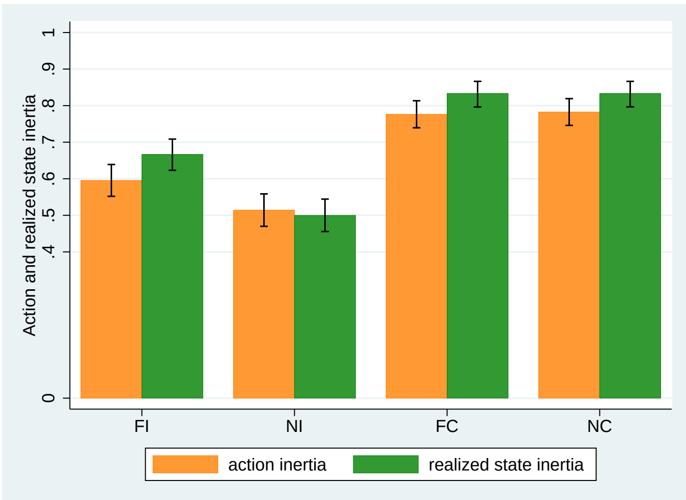


Figure 3.3: Proportion of observations with $a_1 = a_2$ and $\theta_1 = \theta_2$ pooled across subjects for each treatment.



The error rate was low and similar across treatments. Subjects answered correctly in both periods in 1,529 cases, only in period 2 in 179 cases, only in period 1 in 166 cases, and in neither period in 94 cases. Fig. 3.2 captures the precision defined as a proportion of correct choices pooled across subjects in each period of each treatment. It ranges from 0.84 to 0.9.

Subjects exhibited serial correlation in actions. Fig. 3.3 depicts the total proportion of observations with $a_1 = a_2$ (action inertia) and with $\theta_1 = \theta_2$ (realized state inertia) for each treatment. Action inertia equal to 1/2 means that $\Pr[a_2 = a_1] = \Pr[a_2 \neq a_1]$: unconditionally, the actions are serially independent. Action inertia higher than 1/2 means that $\Pr[a_1 = a_1] > \Pr[a_2 \neq a_1]$: unconditionally, the actions are positively serially correlated. In all treatments, the action inertia is in between the theoretical state inertia (ρ)[9] and the realized state inertia. The factors underlying action inertia, the focus of our study, are discussed in Section 3.5.2.

3.5.2 Logit regressions

We run four logit regressions separately for each treatment for a latent variable form specification

$$a_{2,i}^n = \begin{cases} 1 & \beta_0 + \beta_{a_1} a_{1,i}^n + \beta_{\theta_1}\theta_1^n + \beta_{\theta_2}\theta_2^n + \beta_{session_2} session_2 + \beta_{score} score_i^n \theta_2^n + \varepsilon_i^n > 0 \\ 0 & otherwise \end{cases} \tag{3.5}$$

with robust standard errors clustered at the subject level, where $a_{t,i}^n$ is an action taken by subject i in iteration $n = 1, \ldots, 12$ at $t = 1, 2$; θ_t^n is the realized state in iteration n at t; $session_2$ is a dummy variable equal to 1 if $session = 2$ and 0 otherwise; and $score_i^n$ is a subject-specific proxy for her ability added to control for subjects' heterogeneity.

We define $score_i^n$ as the total number of correct answers of subject i in all treatments excluding answers in iteration n of the considered treatment to avoid the endogeneity problem. Hence, $score_i^n$ can vary from 0 (all included answers were incorrect) to 94 (all included answers were correct). The minimum was 49, the maximum was 94, with mean 81.3 and standard deviation 13.2. We use $score_i^n \theta_2^n$ in the regressions in order to capture the sensitivity/ability with which subjects recognize the realized state (when $\theta_2^n = 1$, then a correct answer is $a_{2,i}^n = 1$). Not accounting for subjects' heterogeneity could result in omitted variable bias.

Table 3.3 reports the estimated average marginal effects and p-values of the explanatory variables based on the estimated regressions (3.5), taking into account that all variables but *score* are binary. The regression outputs are in the Appendix.

The results are not consistent with the null hypotheses H_0 and H_0', but they support the alternative A stated in Section 3.2. In all four treatments, variations in a_2 are predominantly explained by variations in θ_2: subjects were learning at $t = 2$. When states

[9] Recall that $\rho = 1/2$ in NI and FI treatments, and $\rho = 3/4$ in FC and NC treatments.

Table 3.3: The estimated average marginal effects and p-values of explanatory variables in regressions (3.5) for each treatment.

	a_1	θ_1	θ_2	$session_2$	$score$
FI	-.021 (.548)	.071 (.092)	**.681** (**.000***)	-.005 (.862)	**.002** (**.034***)
NI	.034 (.406)	-.026 (.596)	**.692** (**.000***)	-.011 (.749)	**.004** (**.000***)
FC	.017 (.603)	**.258** (**.000***)	**.611** (**.000***)	-.038 (.050)	.001 (.120)
NC	**.191** (**.000***)	.002 (.948)	**.629** (**.000***)	.026 (.327)	**.004** (**.000***)

are serially independent (treatments FI, NI), θ_2 is the *only* statistically significant explanatory variable apart of the score: subjects treated the decision problems as unrelated. The average marginal effect of θ_2 decreases with positive serial correlation in states: subjects were offsetting part of their learning at $t = 2$ by drawing inference from the past period. Consistently with the hypothesis A, they partly followed θ_1 when the feedback was provided (FC treatment), and a_1 when no feedback was provided (NC treatment).

3.6 Conclusion

We design an experiment in order to test habit formation. Using a simple perceptual binary choice task, we aim to answer two questions. First, are habits formed as a mechanical consequence of past actions (preference-based explanation)? Or, is habit a functional adaptation that is formed only if it is advantageous to the decision-maker (information-based explanation)? Our data support the hypotheses that habits are not formed when the repeated decision problems are independent, and whenever they are observed, they follow the predictions laid out by information frictions rather than intrinsic preferences.

In our experiment, the incentives are either serially independent or positively serially correlated. An interesting direction for future research is a similar exercise with negatively serially correlated incentives. An implication of the information-based explanation is that we should observe excessive switching rather than excessive action inertia.

3.A Proof of Proposition 8

To solve the model, we will use a posterior-based approach (Caplin and Dean 2013; Kamenica and Gentzkow 2011). Given a signal π_t, each signal realization x_t leads to a posterior $\mu_t^{s^t}$. Accordingly, each π_t leads to a distribution over posterior beliefs. Let the distribution denote by $\phi_t \in \Delta([0,1])$. A signal π_t induces ϕ_t if $\text{supp}(\phi_t) = \{\mu_t^{s^t}\}_{x_t \in X}$ and

$$\begin{aligned}
\mu_t^{s^t} &= \frac{\pi_t(x_t|\theta_t = 1)\mu_t^0}{\pi_t(x_t|\theta_t = 1)\mu_t^0 + \pi_t(x_t|\theta_t = 0)(1-\mu_t^0)} \quad \text{for all } x_t \\
\phi_t(\mu_t) &= \sum_{x_t:\ \mu_t^{s^t} = \mu_t} \big(\pi_t(x_t|\theta_t = 1)\mu_t^0 + \pi_t(x_t|\theta_t = 0)(1-\mu_t^0)\big) \quad \text{for all } \mu_t.
\end{aligned}$$

A distribution ϕ_t satisfies the martingale property if the expected posterior probability equals the prior:

$$\sum_{\text{supp}(\phi_t)} \mu_t \phi_t(\mu_t) = \mu_t^0. \tag{3.6}$$

Caplin and Dean (2013) and Kamenica and Gentzkow (2011) show that the choice of a signal π_t can be equivalently expresses as a choice of a distribution over posteriors ϕ_t subject to (3.6).

For the purpose of our discussion, we only need to focus on the optimal behavior at $t = 2$ without the need to solve the optimal behavior per se at $t = 1$.

Optimal behavior at $t = 2$

No switching costs: $c = 0$

Given μ_2^0, the agent maximizes

$$\max_{\phi_2 \in \Delta([0,1])} \quad \mathbb{E}_{\phi_2}[\max\{\mu_2 R, (1-\mu_2)R\} + \lambda H(\mu_2)] - \underbrace{\lambda H(\mu_2^0)}_{=const.} \tag{3.7}$$

$$s.t. \qquad \mathbb{E}_{\phi_2}[\mu_2] = \mu_2^0 \tag{3.8}$$

where the expectation is with respect to the posteriors μ_2 induced by ϕ_2. The first term is the agent's expected gross utility at μ_2 under her optimal action strategy: $\sigma_2^*(\mu_2) = 1$

if and only if $\mu_2 \geq 1/2$.

Matysková (2018), Lemma 4, shows that the optimal information strategy satisfies $\text{supp}(\phi_2^*) = \{\underline{\mu_2}, \overline{\mu_2}\}$, where $\underline{\mu_2} e^{\frac{R}{\lambda}} + (1-\underline{\mu_2})\frac{1}{e^{\frac{R}{\lambda}}} = 1$ and $\overline{\mu_2}\frac{1}{e^{\frac{R}{\lambda}}} + (1-\overline{\mu_2})e^{\frac{R}{\lambda}} = 1$ yielding

$$\underline{\mu_2} = \frac{1}{1+e^{\frac{R}{\lambda}}} < 1/2, \quad \overline{\mu_2} = \frac{e^{\frac{R}{\lambda}}}{1+e^{\frac{R}{\lambda}}} = 1 - \underline{\mu_2} > 1/2. \tag{3.9}$$

The agent's optimal action strategy satisfies $\sigma_2^*(\underline{\mu_2}) = 0$ and $\sigma_2^*(\overline{\mu_2}) = 1$. Hence, the conditional probabilities with which action $a_2 = 1$ is chosen is the probability with which the posterior $\overline{\mu_2}$ is realized conditional on the prior μ_2^0, which is determined by the martingale property (3.8): $\Pr[a_2 = 0|\mu_2^0]\underline{\mu_2} + \Pr[a_2 = 1|\mu_2^0]\overline{\mu_2} = \mu_2^0$. Hence,

$$\Pr[a_2 = 1|\mu_2^0, a_1] = \frac{\mu_2^0 - \underline{\mu_2}}{\overline{\mu_2} - \underline{\mu_2}}, \tag{3.10}$$

where $\underline{\mu_2}$, $\overline{\mu_2}$ are given by (3.9). Note that this probability is independent of action a_1 once controlled for a possible effect of a_1 on μ_2^0.

Switching costs: $c > 0$

The switching costs create asymmetry in the payoffs from taking action a_2 depending on a_1. Consider $a_1 = 0$. The agent maximizes

$$\max_{\phi_2 \in \Delta([0,1])} \quad \mathbb{E}_{\phi_2}[\max\{\mu_2 R - c, (1-\mu_2)R\} + \lambda H(\mu_2)] - \underbrace{\lambda H(\mu_2^0)}_{=const.} \tag{3.11}$$

$$s.t. \qquad \mathbb{E}_{\phi_2}[\mu_2] = \mu_2^0 \tag{3.12}$$

where the expectation is with respect to the posteriors μ_2 induced by ϕ_2. The first term is the agent's expected gross utility at a posterior μ_2 under her optimal action strategy: $\sigma_2^*(\mu_2) = 1$ if and only if $\mu_2 \geq \frac{R+c}{2R}$.

Matysková (2018), Lemma 4, shows that the optimal information strategy satisfies $\text{supp}(\phi_2^*) = \{\underline{r}, \overline{r}\}$, where the values $\underline{r}$, $\overline{r}$ are given by $\underline{r} e^{\frac{R-c}{\lambda}} + (1-\underline{r})\frac{e^{\frac{-c}{\lambda}}}{e^{\frac{R}{\lambda}}} = 1$, and $\overline{r}\frac{1}{e^{\frac{R-c}{\lambda}}} + (1-\overline{r})\frac{e^{\frac{R}{\lambda}}}{e^{\frac{-c}{\lambda}}} = 1$, yielding

$$\underline{r} \;=\; \frac{e^{\frac{c+R}{\lambda}}-1}{e^{\frac{2R}{\lambda}}-1}, \quad \overline{r} = e^{\frac{R-c}{\lambda}}\underline{r}. \tag{3.13}$$

The agent's optimal action strategy satisfies: $\sigma_2^*(\underline{r}) = 0$ and $\sigma_2^*(\overline{r}) = 1$. As before, the conditional probability that $a_2 = 1$ is the probability that a posterior $\overline{r}$ is realized conditional on the prior μ_2^0, which is determined by the martingale property (3.12). Hence,

$$\Pr[a_2 = 1|\mu_2^0, a_1 = 0] = \frac{\mu_2^0 - \underline{r}}{\overline{r} - \underline{r}}. \tag{3.14}$$

Consider $a_1 = 1$. The agent's optimal action strategy is then $\sigma_2^*(\mu_2) = 1$ if and only if $\mu_2 \geq \frac{R-c}{2R}$. Let $\mathrm{supp}(\phi_2^*) = \{\underline{s}, \overline{s}\}$ be the support of the optimal strategy. Analogically, we get

$$\underline{s} \;=\; 1 - \overline{r}, \quad \overline{s} = 1 - \underline{r}. \tag{3.15}$$

The agent's optimal action then satisfies $\sigma_2^*(\underline{s}) = 0$ and $\sigma_2^*(\overline{s}) = 1$. As before, the conditional probability that $a_2 = 1$ is the probability with which posterior $\overline{s}$ is realized conditional on the prior μ_2^0, which is determined by the martingale property. Hence,

$$\Pr[a_2 = 1|\mu_2^0, a_1 = 1] = \frac{\mu_2^0 - \underline{s}}{\overline{s} - \underline{s}}. \tag{3.16}$$

Action inertia

NI and FI treatments

The prior belief $\mu_2^0 = 1/2$, i.e., it is independent of a_1 and θ_1 regardless of the distribution specification of the costless signal. Then a habit is not formed when $c = 0$ (eq. (3.10) is independent of a_1), while a habit is formed when $c > 0$ (eq.'s (3.14) and (3.16) are not equal even when $\mu_2^0 = 1/2$).

NC treatment

Note that in period $t = 1$, the (myopic and a fully rational) agent always learns[10]. Therefore, it must be $\mu_1^{s^1} \neq 1/2$. From the optimality of action strategy we have: if $\sigma_1^*(\mu_1^{s^1}) = 1$, then $\mu_1^{s^1} > 1/2$ and if $\sigma_1^*(\mu_1^{s^1}) = 0$, then $\mu_1^{s^1} < 1/2$. Therefore, a habit is formed for all $c \geq 0$: eq.'s (3.10), (3.14) and (3.16) are strictly increasing in μ_2^0 where μ_2^0 under $a_1 = 1$ is higher than μ_0^2 under $a_1 = 0$.

FC treatment

The agent perfectly learns the realization θ_1 from the costless signal. Then the prior μ_2^0 depends on θ_1, but is independent of a_1 (see eq. (3.2)). As the eq. (3.10) depends on μ_2^0 but not explicitly on a_1, a habit is not formed when $c = 0$. However, a habit is formed when $c > 0$ as eq.'s (3.14) and (3.16) are not equal, keeping μ_0^2 fixed.

[10] At the prior $\mu_1^0 = 1/2$, he is exactly indifferent between the actions. The marginal value of information is thus infinite while the marginal cost of information is finite.

3.B Regression outputs

Note that we use a command that exports the estimated coefficients already as an estimated odds ratio instead of a log odds ratio.

FI TREATMENT

```
Logistic regression                               Number of obs   =        492
                                                  Wald chi2(6)    =     142.93
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -163.10254                 Pseudo R2       =     0.5150

                                (Std. Err. adjusted for 41 clusters in subject)
--------------------------------------------------------------------------------
               |               Robust
       choice2 | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
     1.choice1 |   .8083733   .2882132    -0.60   0.551     .4019102    1.625904
      1.state1 |    1.96372   .7578438     1.75   0.080      .921683    4.183862
      1.state2 |   .0000563   .0000942    -5.85   0.000     2.12e-06    .0014955
    1.session2 |   .9545169   .2568266    -0.17   0.863     .5633204    1.617379
               |
state2#c.score |
            0  |   .9325278   .0140836    -4.63   0.000      .905329    .9605438
            1  |   1.108449   .0179016     6.38   0.000     1.073912    1.144097
               |
         _cons |   35.59983   41.01105     3.10   0.002     3.722741    340.4341
--------------------------------------------------------------------------------
```

```
Average marginal effects                          Number of obs   =        492
Model VCE    : Robust

Expression   : Pr(choice2), predict()
dy/dx w.r.t. : 1.choice1 1.state1 1.state2 1.session2 score

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   1.choice1 |  -.0213619   .0355646    -0.60   0.548    -.0910673    .0483434
    1.state1 |   .0717546    .042604     1.68   0.092    -.0117478    .1552569
    1.state2 |    .680717   .0323603    21.04   0.000      .617292    .7441419
  1.session2 |  -.0046997   .0270881    -0.17   0.862    -.0577913     .048392
```

score	.0021162	.0009961	2.12	0.034	.000164	.0040684

Note: dy/dx for factor levels is the discrete change from the base level.

NI TREATMENT

Logistic regression   Number of obs = 492
Wald chi2(6) = 133.54
Prob > chi2 = 0.0000
Log pseudolikelihood = -169.51864   Pseudo R2 = 0.3798

(Std. Err. adjusted for 41 clusters in subject)

choice2	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.choice1	1.389138	.5456704	0.84	0.403	.6432551	2.999907
1.state1	.7762836	.3740146	-0.53	0.599	.3019328	1.995863
1.state2	.0013902	.0030984	-2.95	0.003	.0000176	.109694
1.session2	.8971487	.3004906	-0.32	0.746	.4653324	1.729679
state2#c.score						
0	.9428402	.0218128	-2.54	0.011	.9010427	.9865765
1	1.073454	.0132159	5.76	0.000	1.047861	1.099671
_cons	21.84399	41.73286	1.61	0.106	.5165522	923.7395

Average marginal effects   Number of obs = 492
Model VCE : Robust

Expression : Pr(choice2), predict()
dy/dx w.r.t. : 1.choice1 1.state1 1.state2 1.session2 score

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.choice1	.0341974	.0411712	0.83	0.406	-.0464966	.1148914
1.state1	-.0262127	.0494824	-0.53	0.596	-.1231964	.070771
1.state2	.6918165	.0542746	12.75	0.000	.5854404	.7981927
1.session2	-.0112297	.0351127	-0.32	0.749	-.0800493	.0575899
score	.0044963	.0010308	4.36	0.000	.002476	.0065166

Note: dy/dx for factor levels is the discrete change from the base level.

FC TREATMENT

```
Logistic regression                               Number of obs     =        492
                                                  Wald chi2(6)      =     102.34
                                                  Prob > chi2       =     0.0000
Log pseudolikelihood = -108.77105                 Pseudo R2         =     0.6757

                                  (Std. Err. adjusted for 41 clusters in subject)
---------------------------------------------------------------------------------
               |               Robust
       choice2 | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+-----------------------------------------------------------------
     1.choice1 |   1.292079   .6226411     0.53   0.595     .5024617    3.322579
      1.state1 |   13.99275   7.383847     5.00   0.000     4.974295    39.36175
      1.state2 |   .0000638   .0001345    -4.58   0.000     1.02e-06    .0039746
    1.session2 |   .5520429   .1616916    -2.03   0.043     .3109271    .9801375
               |
state2#c.score |
            0  |   .9254083   .0185939    -3.86   0.000     .8896732    .9625788
            1  |    1.10163   .0182308     5.85   0.000     1.066471    1.137947
               |
         _cons |   20.33895   28.34802     2.16   0.031     1.324162    312.4036
---------------------------------------------------------------------------------


Average marginal effects                          Number of obs     =        492
Model VCE    : Robust

Expression   : Pr(choice2), predict()
dy/dx w.r.t. : 1.choice1 1.state1 1.state2 1.session2 score

-------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
   1.choice1 |    .016663   .0320507     0.52   0.603    -.0461552    .0794813
    1.state1 |    .257624   .0581595     4.43   0.000     .1436334    .3716147
    1.state2 |   .6112043    .045856    13.33   0.000     .5213282    .7010805
  1.session2 |  -.0378612   .0193417    -1.96   0.050    -.0757701    .0000478
       score |   .0010713   .0006895     1.55   0.120      -.00028    .0024226
-------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

NC TREATMENT

```
Logistic regression                               Number of obs   =        492
                                                  Wald chi2(6)    =     195.03
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -130.91083                 Pseudo R2       =     0.5968

                                 (Std. Err. adjusted for 41 clusters in subject)
--------------------------------------------------------------------------------
               |               Robust
       choice2 | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
     1.choice1 |   6.528634   3.008268     4.07   0.000     2.646093    16.10792
      1.state1 |   1.029548   .4622255     0.06   0.948     .4270629       2.482
      1.state2 |    .000018   .0000266    -7.37   0.000     9.84e-07    .0003276
    1.session2 |    1.38445   .4490417     1.00   0.316     .7331505    2.614338
               |
state2#c.score |
            0  |   .9290356   .0170773    -4.00   0.000     .8961604    .9631167
            1  |   1.120191   .0174057     7.30   0.000     1.086591    1.154831
               |
         _cons |    20.5216   27.46066     2.26   0.024     1.490011    282.6396
--------------------------------------------------------------------------------


Average marginal effects                          Number of obs   =        492
Model VCE    : Robust

Expression   : Pr(choice2), predict()
dy/dx w.r.t. : 1.choice1 1.state1 1.state2 1.session2 score

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   1.choice1 |   .1910301   .0509049     3.75   0.000     .0912583    .2908018
    1.state1 |   .0023206   .0358441     0.06   0.948    -.0679325    .0725737
    1.state2 |   .6287987   .0674834     9.32   0.000     .4965337    .7610637
  1.session2 |   .0259479   .0264879     0.98   0.327    -.0259674    .0778633
       score |   .0035593   .0007534     4.72   0.000     .0020827    .0050358
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

3.C Instructions

Welcome to the experiment! Please take a record sheet at the front if you don't have one already. Please do not use the computers during the instructions. When it is time to use the computer, please follow the instructions precisely.(Repeat if necessary.)

Please raise your hand if you need a pencil. Please put away and silence all your personal belongings, especially your phone. We need your full attention during the experiment.

Raise your hand at any point if you cannot see or hear well.

The experiment you will be participating in today is an experiment in decision making. At the end of the experiment, you will be paid for your participation in cash. The amount you earn depends on your decisions and on chance. You will be using the computer for the experiment, and all decisions will be made through the computer. DO NOT socialize or talk during the experiment.

All instructions and descriptions that you will be given in the experiment are accurate and true. In accordance with the policy of this lab, at no point will we attempt to deceive you in any way.

If you have any questions, raise your hand and your question will be answered out loud so everyone can hear.

After you have completed all the tasks, please wait while everyone else finishes his or her tasks. Once everyone has completed the experiment, I will ask you to fill in the questionnaire. After the questionnaire you will collect your earnings and leave.

///

I will now describe the main features of the experiment and show you how to use the software. Again, if you have any questions during this period, please raise your hand.

You will be presented with a series of choices to make. There will be four SETS of choices in today's experiment. Each set contains twelve ITERATIONS, and each iteration has two PERIODS. In each period, you will be shown a picture of 100 dots. Each dot will be either RED or BLUE. We have displayed an example of such a screen on your computer monitor. (show an example screen)

This is an example of the screens you will see during the experiment. In every period, the picture will contain either 51 red dots and 49 blue dots, or instead, 51 blue dots and 49 red dots. We will call these two cases MAJORITY RED and MAJORITY BLUE, respectively. In each case, the dots are randomly allocated to the positions in the matrix. In each period the computer will choose randomly between MAJORITY RED and MAJORITY BLUE. You will be told in advance how likely each case is to happen.

In each period, you will be asked to determine if the image is MAJORITY RED or MAJORITY BLUE. While you may take as much time as you need to make your choice, the image will disappear after 45 seconds.

I am now going to describe the details of the experiment.

The experiment is divided into four SETS. In each set, you will be presented with twelve iterations, and each iteration consists of two periods, each with its own image. The rules for the 12 iterations within each set are identical, but the rules are different in different sets.

In PERIOD 1 of each iteration, the image is always generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, meaning that there is a 50% chance of MAJORITY RED and a 50% chance of MAJORITY BLUE.

In period 2 of each iteration, the image will be generated in a way that differs across sets. In some sets, the majority color for period 2 is chosen in a way that is completely separate from the period 1 image, and is randomly generates so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, just like the period 1 image. But in other sets, the period 2 image depends on the majority color of the period 1 image. In these sets, the computer generates the period 2 image so that there is a 75% chance that the majority color matches the period 1 majority color, and a 25% chance that the majority color is different from the period 1 majority color.

It is important to remember that while the periods within each iteration may be related to each other, the periods across iterations are never related.

After making your choices, you will always be told what the majority color was, but the timing of this differs from set to set. In some sets, the majority colors will be revealed after every period. In other sets, the majority colors for an iteration will not be revealed until you complete both periods. Before each set, you will be told about the timing of the feedback you will receive.

The amount of money you will receive at the end of the experiment depends on your choices. After we have completed all four sets, you will have made 96 choices (4 sets times 12 iterations times 2 periods). The computer software will randomly select one of these 96 periods. Your payment will be determined by your choice in that single period. If your choices in the randomly chosen period matches the majority color, you will earn an additional $5 dollars on top of the $15 show-up fee. Otherwise, you will receive no additional payment, but you will still receive the show-up fee.

After you complete the last set, please wait until we start the questionnaire part. After you finish the questionnaire, please fill your record sheet on the desk. I will pay one by one to keep everyone’s privacy.

To summarize, remember that we have four sets in the experiment today. Each set consists of 12 iterations, and each iteration consists of two periods. The sets will vary in how likely it is that the majority colors are the same for both periods within an iteration, and in the timing that the majority colors are revealed. Please raise your hand if you have any questions.

(1) FI/FC/NI/NC

Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the majority color from period 1, and a 25% chance that the majority color is different from period 1.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

No Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

No Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the period 1 majority color, and a 25% chance that the majority color is different from the period 1 majority color.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

(2) NI/NC/FI/FC

No Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

No Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the period 1 majority color, and a 25% chance that the majority color is different from the period 1 majority color.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the majority color from period 1, and a 25% chance that the majority color is different from period 1.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

Bibliography

Abel, Andrew B. 1990. "Asset prices under habit formation and catching up with the Joneses." Technical Report, National Bureau of Economic Research.

Alonso, Ricardo, and Odilon Câmara. 2016. "Bayesian persuasion with heterogeneous priors." *Journal of Economic Theory* 165:672–706.

———. 2018. "On the value of persuasion by experts." *Journal of Economic Theory* 174:103–123.

Aoyagi, Masaki. 2014. "Strategic obscurity in the forecasting of disasters." *Games and Economic Behavior* 87:485–496.

Au, Pak Hung. 2015. "Dynamic information disclosure." *The RAND Journal of Economics* 46 (4): 791–823.

Bergemann, Dirk, and Stephen Morris. 2016. "Information design, bayesian persuasion, and bayes correlated equilibrium." *American Economic Review* 106 (5): 586–91.

Bhalla, Manaswini. 2013. "Waterfall versus sprinkler product launch strategy: Influencing the herd." *The Journal of Industrial Economics* 61 (1): 138–165.

Bizzotto, Jacopo, Jesper Rüdiger, and Adrien Vigier. 2015. "The Optimal Timing of Persuasion."

Blackwell, David A, and Meyer A Girshick. 1954. *Theory of games and statistical decisions.* John Wiley and sons.

Bose, Subir, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. 2006. "Dynamic monopoly pricing and herding." *The RAND Journal of Economics* 37 (4): 910–928.

———. 2008. "Monopoly pricing in the binary herding model." *Economic Theory* 37 (2): 203–241.

Bourdieu, Pierre. 1985. "The social space and the genesis of groups." *Information (International Social Science Council)* 24 (2): 195–220.

Brown, Alexander L, Colin F Camerer, and Dan Lovallo. 2012. "To review or not to review? Limited strategic thinking at the movie box office." *American Economic Journal: Microeconomics* 4 (2): 1–26.

Caplin, Andrew, Dániel Csaba, and John Leahy. 2018. "Rational Inattention and Psychometrics."

Caplin, Andrew, and Mark Dean. 2013. "Behavioral implications of rational inattention with shannon entropy." Technical Report, National Bureau of Economic Research.

———. 2015. "Revealed preference, rational inattention, and costly information acquisition." *American Economic Review* 105 (7): 2183–2203.

Caplin, Andrew, Mark Dean, and John Leahy. 2016. "Rational Inattention, Optimal consideration sets and stochastic choice." Technical Report, Working paper.

———. 2017. "Rationally inattentive behavior: Characterizing and generalizing Shannon entropy." Technical Report, National Bureau of Economic Research.

Carroll, Christopher D, Jody Overland, and David N Weil. 2000. "Saving and growth with habit formation." *American Economic Review* 90 (3): 341–355.

Cheremukhin, Anton, Anna Popova, and Antonella Tutino. 2015. "A theory of discrete choice with information costs." *Journal of Economic Behavior & Organization* 113:34–50.

Constantinides, George M. 1990. "Habit formation: A resolution of the equity premium puzzle." *Journal of political Economy* 98 (3): 519–543.

Cover, Thomas M, and Joy A Thomas. 2006. *Elements of information theory 2nd edition.* John Wiley & Sons.

Crapis, Davide, Bar Ifrach, Costis Maglaras, and Marco Scarsini. 2016. "Monopoly pricing in the presence of social learning." *Management Science* 63 (11): 3586–3608.

Dean, Mark, and Nathaniel Neligh. 2017. "Experimental Tests of Rational Inattention." Technical Report, Working Paper, Columbia University.

Dellarocas, Chrysanthos. 2006. "Strategic manipulation of internet opinion forums: Implications for consumers and firms." *Management science* 52 (10): 1577–1593.

Dewan, Ambuj, and Nate Neligh. 2017. "Estimating information cost functions in models of rational inattention." *unpublished, Columbia University.*

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica. 2015. "Suspense and surprise." *Journal of Political Economy* 123 (1): 215–260.

Ely, Jeffrey C. 2017. "Beeps." *American Economic Review* 107 (1): 31–53.

Eyster, Erik, and Matthew Rabin. 2005. "Cursed equilibrium." *Econometrica* 73 (5): 1623–1672.

Fuhrer, Jeffrey C. 2000. "Habit formation in consumption and its implications for monetary-policy models." *American Economic Review* 90 (3): 367–390.

Gehlbach, Scott, and Alberto Simpser. 2015. "Electoral manipulation as bureaucratic control." *American Journal of Political Science* 59 (1): 212–224.

Gentzkow, Matthew, and Emir Kamenica. 2014. "Costly persuasion." *American Economic Review* 104 (5): 457–62.

———. 2016. "Competition in persuasion." *The Review of Economic Studies* 84 (1): 300–322.

———. 2017. "Bayesian persuasion with multiple senders and rich signal spaces." *Games and Economic Behavior* 104:411–429.

Gick, Wolfgang, and Thilo Pausch. 2012. "Persuasion by stress testing: Optimal disclosure of supervisory information in the banking sector."

Hanoch, Giora, and Haim Levy. 1969. "The Efficiency Analysis of Choices Involving Risk." *Review of Economic Studies* 36 (3): 335–346.

Hedlund, Jonas. 2017. "Bayesian persuasion by a privately informed sender." *Journal of Economic Theory* 167:229–268.

Ifrach, Bar, Costis Maglaras, and Marco Scarsini. 2014. "Bayesian social learning with consumer reviews." *ACM SIGMETRICS Performance Evaluation Review* 41 (4): 28–28.

Jermann, Urban J. 1998. "Asset pricing in production economies." *Journal of monetary Economics* 41 (2): 257–275.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101 (6): 2590–2615.

Kessler, Anke S. 1998. "The value of ignorance." *The Rand Journal of Economics*, pp. 339–354.

Khaw, Mel Win, Luminita Stevens, and Michael Woodford. 2017. "Discrete adjustment to a changing environment: Experimental evidence." *Journal of Monetary Economics* 91:88–103.

Kolotilin, Anton. 2015. "Experimental design to persuade." *Games and Economic Behavior* 90:215–226.

———. 2018. "Optimal information disclosure: A linear programming approach." *Theoretical Economics*, vol. forthcoming.

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. 2017. "Persuasion of a privately informed receiver." *Econometrica* 85 (6): 1949–1964.

Levy, Haim. 2015. *Stochastic dominance: Investment decision making under uncertainty.* Springer.

Li, Fei, and Peter Norman. 2017. "On Bayesian persuasion with multiple senders."

Li, Xinxin, and Lorin M Hitt. 2008. "Self-selection and information role of online product reviews." *Information Systems Research* 19 (4): 456–474.

Lipnowski, Elliot, and Laurent Mathevet. 2017. "Simplifying Bayesian Persuasion." Technical Report, Working Paper, University of Chicago and New York University.

Liu, Ting, and Pasquale Schiraldi. 2012. "New product launch: herd seeking or herd preventing?" *Economic Theory* 51 (3): 627–648.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2015. "Inattention to rare events."

Martin, Daniel. 2017. "Strategic pricing with rational inattention to quality." *Games and Economic Behavior* 104:131–145.

Matějka, Filip, and Alisdair McKay. 2015. "Rational inattention to discrete choices: A new foundation for the multinomial logit model." *American Economic Review* 105 (1): 272–98.

Matysková, Ludmila. 2018. "Bayesian Persuasion with Costly Information Acquisition."

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier. 2014. "Promotional reviews: An empirical investigation of online review manipulation." *American Economic Review* 104 (8): 2421–55.

Okonofua, Eni C, Kit N Simpson, Ammar Jesri, Shakaib U Rehman, Valerie L Durkalski, and Brent M Egan. 2006. "Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals." *Hypertension* 47 (3): 345–351.

Oliveira, Henrique, Tommaso Denti, Maximilian Mihm, and Kemal Ozbek. 2017. "Rationally inattentive preferences and hidden information costs." *Theoretical Economics* 12 (2): 621–654.

Papanastasiou, Yiangos, Nitin Bakshi, and Nicos Savva. 2014. "Scarcity strategies under quasi-bayesian social learning." *History.*

Parlour, Christine A, Vesna Prasnikar, and Uday Rajan. 2007. "Compensating for the winner's curse: Experimental evidence." *Games and Economic Behavior* 60 (2): 339–356.

Perez-Richet, Eduardo, and Delphine Prady. 2011. "Complicating to persuade."

Phillips, Lawrence S, William T Branch, Curtiss B Cook, Joyce P Doyle, Imad M El-Kebbi, Daniel L Gallina, Christopher D Miller, David C Ziemer, and Catherine S Barnes. 2001. "Clinical inertia." *Annals of internal medicine* 135 (9): 825–834.

Ravid, Doron. 2017. "Bargaining with rational inattention."

Rayo, Luis, and Ilya Segal. 2010. "Optimal information disclosure." *Journal of political Economy* 118 (5): 949–987.

Roesler, Anne-Katrin, and Balázs Szentes. 2017. "Buyer-optimal learning and monopoly pricing." *American Economic Review* 107 (7): 2072–80.

Samuelson, William, and Richard Zeckhauser. 1988. "Status quo bias in decision making." *Journal of risk and uncertainty* 1 (1): 7–59.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of monetary Economics* 50 (3): 665–690.

Steiner, Jakub, Colin Stewart, and Filip Matějka. 2017. "Rational Inattention Dynamics: Inertia and Delay in Decision-Making." *Econometrica* 85 (2): 521–553.

Suri, Gaurav, Gal Sheppes, Carey Schwartz, and James J Gross. 2013. "Patient inertia and the status quo bias: when an inferior option is preferred." *Psychological science* 24 (9): 1763–1769.

Van Nieuwerburgh, Stijn, and Laura Veldkamp. 2009. "Information immobility and the home bias puzzle." *The Journal of Finance* 64 (3): 1187–1215.

Yang, Ming. 2011. "Coordination with rational inattention."