INATTENTION AND BELIEF POLARIZATION

KRISTOFFER P. NIMARK AND SAVITAR SUNDARESAN

Abstract. Disagreement persists over issues that have objective truths. In the presence of increasing amounts of data, such disagreement should vanish, but it is nonetheless observable. This paper studies persistent disagreement in a model where rational Bayesian agents learn about an unobservable state of the world through noisy signals. We show that agents (i) choose signal structures that are more likely to reinforce their prior beliefs and (ii) choose less informative signals when their prior beliefs are more precise. For sufficiently precise beliefs, agents choose completely uninformative signals. We call the former the *confirmation effect* and the latter the *complacency effect*. Taken together, the two effects imply that the beliefs of ex ante identical agents over time can cluster in two distinct groups at opposite ends of the belief space. The complacency effect holds uniformly when information cost is proportional to channel capacity, but not when cost is proportional to reduction in entropy.

1. Introduction

Many countries have experienced increases in political polarization and in disagreement about objective facts. For example, disagreement in the US about whether climate change is real and caused by human activities has increased and the views on what is essentially an empirical scientific question is well predicted by party affiliation.[1] Political polarization and disagreement about facts clearly has many causes and politically motivated disinformation, is likely to be one of them. Stating a particular belief about a given fact may for many people also be more an expression of group belonging, rather than an expression of a sincerely held belief about the true nature of the world. However, what we show in this paper is that even ex ante identical, rational agents may self-sort into different informational bubbles, where agents within one group permanently hold beliefs about a fact that are the opposite of the beliefs of the members of the other group.[2]

Date: November 19, 2018, Nimark: Economics Department, Cornell University. *e-mail*: pkn8@cornell.edu *webpage*: www.kris-nimark.net. Sundaresan: Imperial College London Business School. *e-mail*: s.sundaresan@ic.ac.uk, *webpage*: www.savitars.com. The authors thank Alessandra Casella, Tommaso Denti, Sylverie Herbert, Navin Karthik, Alex Kohlhas, Bartosz Mackowiak, Thomas Palfrey, Ricardo Reis, Tomasz Strzalecki, Mathieu Taschereau-Dumouchel, Sebastien Turban, Laura Veldkamp, Mirko Wiederholt, Michael Woodford, participants at seminars at NYU, Columbia, Caltech, SITE, SED and the Conference on Rational Inattention and Related Theories, as well as an anonymous referee, for helpful comments and suggestions.

[1] See for instance McCright and Dunlap (2011). Another example is Italy where negative views about the safety of vaccines appear to be closely correlated with support for the populist 5 Star Movement, e.g. *Rise of Italian populist parties buoys anti-vaccine movement* , The Guardian, 11 July 2018 available at https://www.theguardian.com/world/2018/jul/11/rise-of-italian-populist-parties-buoys-anti-vaccine-movement.

[2] See Pariser (2011) for a popular treatment of filter bubbles and Resnick et al. (2013) for a brief overview of some of the related scholarly literature.

This result is driven by two effects, both of which are consequences of agents' endogenous information choices. The first effect, which we call the *confirmation effect*, causes agents to choose to observe signals that are more precise in states they believe to be more likely. Those signals are therefore more likely to confirm their prior beliefs. If two agents initially observe different realizations of signals drawn from the same distribution, the confirmation effect then makes it less likely that agents' beliefs will converge over time. The second effect, which we call the *complacency effect*, causes agents to choose less precise channels as their uncertainty decreases. For sufficiently precise beliefs, this effect causes agents to choose completely uninformative signals. Combined, the two effects imply that the beliefs of ex ante identical agents over time will cluster in two distinct groups on opposite ends of the belief space.

A basic premise of this paper is that there exists an objective reality. However, agents cannot observe this reality directly and instead have to choose a noisy information channel. Less noisy channels are more costly and agents can choose channels that have different precisions in different states of the world. One way to interpret this setup is that agents choose the medium through which facts about the world are channeled. For example, we may all be trying to verify whether global warming exists. Informative sources require a lot of attention, while uninformative sources require very little: it may be harder to extract a signal from an article in *Nature* than from a sound bite on cable television. Some channels may also be more accurate in one state of the world, while other channels may be more accurate in other states. If precision is costly, we show that agents allocate more precision to those states they find a priori more likely. That means that agents endogenously choose channels that are less likely to prove their priors wrong. The endogenous allocation of precision may then perpetuate differences in beliefs and lead to permanent disagreement about the true state of the world.

The early literature on costly information acquisition treated information as a scarce resource, e.g. Grossman and Stiglitz (1980). The key conceptual shift introduced by Sims (1998) and Sims (2003) was that information may be plentiful, but people's *attention* is a scarce resource. Beyond this basic change of perspective, the rational inattention literature inspired by Sims early work makes specific assumptions about how to model the cost of attention. The most common formulations either put a constraint on the reduction in entropy that agents experience by observing a signal, or a utility or pecuniary cost that is increasing in the reduction in entropy, e.g. Sims (1998), Sims (2003), Maćkowiak and Wiederholt (2009), Woodford (2009), Matějka (2015) and Matejka and McKay (2015).[3]

In the analysis below, we study the optimal information choices under two different cost functions for information acquisition. An alternative to the standard approach to model information cost as proportional to the expected reduction in entropy, is to model information cost as proportional to channel capacity. Like mutual information, channel capacity is based on 1948's entropy concept and it can be interpreted as the maximum possible entropy reduction that any agent, could achieve by observing a given signal structure. Woodford (2012) was the first paper to propose using channel capacity as a measure of information cost and

[3]It is common in the rational inattention literature to not model the signals explicitly but instead to focus on the mutual information between actions and latent states. This is without loss of generality and is inconsequential for the argument here.

used it to explain several choice anomalies. Here, we show that the modeling information cost as mutual information or channel capacity lead to substantively different predictions about how agents choose signal structures as functions of their prior beliefs.

Woodford (2012) argued that specifying information cost as proportional to channel capacity rather than to entropy reduction allows for better explanations of experimental evidence on visual perception. We interpret the different specifications of information costs as reflecting differences in terms of how much control an agent has over the information generating process. If an agent can control the information generating process so that he only spends resources on acquiring new information, it is reasonable to model the cost of information as being related to how much he learns, i.e. how large the reduction in entropy is between his prior beliefs and his posterior. Examples of such behavior are surveys of market participants, direct measurements of some aspect of reality, or experiments designed to answer a particular question.

On the other hand, consider an agent who does not have complete control over how the information he observes is generated, but instead uses secondary sources such as newspapers, radio broadcasts or TV shows to acquire more information. It is then more natural to think of the cost of information in terms of time spent reading, listening or watching, and it is less clear that the cost of information should be measured relative to the prior knowledge of the agent. For instance, a longer broadcast that contains more information is more costly to watch, but also potentially more informative. It is possible that some information in a broadcast is already known to a relatively informed agent, but unless the agent knows exactly at which points in the broadcast already known information will be revealed, it is as costly for the well-informed as for the uninformed to watch the broadcast. In such settings it may be more natural to think of the cost of information as being determined by the maximum any person could learn from a text or broadcast and being independent of that agent's prior knowledge. This latter scenario corresponds to modeling the cost of information in terms of channel capacity.

Using mutual information as a measure of informativeness may thus be more reasonable in settings where agents can design their signal structure to completely avoid repetition of already known information. However, using channel capacity may be more reasonable in settings where agents can choose the precision of information but cannot avoid repetition of already known information. Both approaches have appealing features and we will not argue that one is always a better description of human behavior than the other. Instead, our focus is on demonstrating the different predictions that the two specifications imply about how agents form beliefs and how these evolve over time. In particular, we demonstrate that modeling information cost in terms of channel capacity implies that when agents' priors are sufficiently precise, they are more likely to choose completely uninformative signals and thus stop updating their beliefs.

To understand why the two cost functions imply different behavior, consider the decision problem of an agent who wants to decide whether a given signal is worth paying attention to or not. When the precision of the agent's prior is higher, the marginal value of observing an additional signal decreases since the agent is already pretty sure about the state of the world. The expected reduction in uncertainty from observing the signal is then small. But if the cost of the signal is measured in terms of entropy reduction, a given signal also becomes

cheaper as the prior precision of an agent's beliefs increases. In the limit with perfectly precise priors, any signal, regardless of its precision, can be observed for free. On the other hand, if the cost of paying attention to a given signal depends only on the precision of the signal, agents demand less and less precise signals as the precision of their beliefs increase, since the marginal usefulness of the signal then decreases. This implies that when an agent's beliefs are precise enough, he will choose to observe completely uninformative signals and not update his beliefs further. Importantly, an agent may stop updating his beliefs before they become degenerate. In fact, even an agent that attaches a higher probability to the incorrect state than to the correct state of the world may stop updating his beliefs. The beliefs of different agents may then cluster permanently in two distinct groups, where one group is almost certain that one state has occurred and the other group is almost certain that it has not.

Our paper is not the first to propose a theory that can explain persistent disagreement among agents. The starting point of many studies is a well-known result by Savage (1954). He argued that repeated observation of signals will lead a Bayesian agent to assign probability 1 to a true event almost surely as his length of experience increases. Even if different agents start off with different priors, their beliefs should then converge over time. Underlying this result is an assumption that the true state of nature is assigned a positive probability a priori. Blackwell and Dubins (1962) built on Savage's result and showed that as long as the two agents' priors are absolutely continuous, then their beliefs will converge over time.

One way to break the result that beliefs converge over time is to assume that agents assign a zero probability to some states that may in fact occur. For instance, Freedman (1963) and Miller and Sanchirico (1999) show that relaxing the assumption of absolute continuity between the subjective and the true distribution weakens the result. In related work, Berk (1966) shows that if an agent uses an incorrect model, his beliefs may not converge to a single point. Our setup generates persistent disagreement without such restrictions. Our agents are rational Bayesians and the support of their prior beliefs contain the true state of nature, which conforms to the assumptions of Blackwell and Dubins (1962).

The agents in our model solve a dynamic information choice problem in a simple binary state setting and our paper contributes to a growing literature studying the optimal allocation of attention in dynamic settings. Examples include Steiner et al. (2017) who propose a framework to study richer discrete state models of inattention, Maćkowiak et al. (2018) who propose analytical methods to study dynamic attention problems in linear Gaussian settings and Afrouzi and Yang (2016) who study how inflation dynamics and forward guidance are affected by rational inattention. Sundaresan (2017) shows that inattention to ex-ante low-probability events can generate endogenously persistent increases in uncertainty. Ilut and Valchev (2017) propose a dynamic framework in which agents can pay attention to learn about a policy function, rather than about an exogenous random variable.

One way to interpret the choice of channels in our model is to think of it as a choice about which sources to get information from, or which news papers to read or what TV channels to watch. Gentzkow and Shapiro (2006) and Besley and Prat (2006) show that competition among information providers makes it difficult to hide information, and Hong and Kacperczyk (2010) shows it can decrease reporting bias. Mullainathan and Shleifer (2005) show that if behavioral agents prefer news that favors their beliefs, media sources

will tend to be biased. This paper can deliver the same demand story while using rational agents. Additionally, this paper provides two dimensions over which channels are graded: overall informativeness and asymmetric precision across states.

Gentzkow and Shapiro (2006) present a model that, like our model, is populated with agents that are rational Bayesians. They posit that if newspapers are rewarded for the perceived accuracy of their reporting, they will bias their reporting to conform to agents' (possibly incorrect) priors. The mechanism behind their result is that the perceived accuracy of an information source is decreasing in the distance between signals and priors. Rational agents then perceive information sources that confirm their priors to be more accurate. In our model, agents know how precise their information is, and yet, they still choose channels that are more likely to confirm their prior beliefs.

Perego and Yuksel (2017) study ideological slant in news media markets where agents have heterogenous preferences over both what the political agenda should be and how issues should be addressed. In their model, increased competition leads to news outlets to provide more specialized content, making agents disagree more about the desirability of a given policy. However, the agents that disagree the most about the desirability of a policy, agree the most about the state, i.e. the consequences of the policy. In contrast, the agents in our model disagree about objective reality.

The confirmation effect that makes agents choose channels that are more likely to confirm their priors is related to other forms of confirmation biases that have been studied in the literature. Nickerson (1998) defines confirmation bias as "the unwitting selectivity in the acquisition and use of evidence". Models of confirmation bias include Suen (2004), Cukierman and Tommasi (1998), Rabin and Schrag (1999), Koszegi and Rabin (2006). Baliga et al. (2013) show that divergence cannot occur in a Bayesian updating framework. Fryer et al. (2013) assume that agents receive ambiguous signals which they interpret as signals in favor of their prior, and keep only this interpretation in memory. In the present paper, the selectivity is not unwitting but intentional, and thus provides a theory to explain the observed behavior without relying on behavioral biases.

Maćkowiak and Wiederholt (2018) studies a model in which agents can choose how much information to acquire about their optimal actions in different states of the world. They show that agents will allocate more attention to learn about optimal behavior in states that are ex ante more likely to occur, and that the expected loss in a given state is inversely proportional to how likely that state is. As in our model, corner solutions may occur where agents may choose to not learn anything about the optimal behavior in one state and instead allocate all their attention to learn about what to do in the other state. However, their model differs from ours in that agents do not allocate attention to learn about which state they are in. In Maćkowiak and Wiederholt (2018), agents do not allocate attention to learn about objective reality, but to prepare for different contingencies that once they occur, are known with certainty.

Some papers such as Lord et al. (1979), Baumeister and Newman (1994), show that people pay less attention to information confirming their prior and evaluate "disconfirming evidence" more thoroughly. This paper's model finds that if agents receive disconfirming signals, they update their beliefs more strongly, which is consistent with this result. Benoît

and Dubra (2017) propose a model that can explain why rational agents may interpret the same evidence differently.

Our model, as well as the rational inattention literature more broadly, presumes that agents' information gathering behavior responds to incentives. A parallel experimental literature provides supporting evidence for this assumption, e.g. Ambuehl (2016) and Bartoš et al. (2016).

The next two sections present the basic set up and describe an agent's optimal channel choice. There, we formally derive the confirmation and the complacency effects and show that for sufficiently precise priors, agents will choose a completely uninformative channel. We then generalize these results to a dynamic setting where agents choose the precision of the current channel while taking into account that more precise information today will also increase future utility. Using the dynamic model, we demonstrate that the combination of the confirmation and complacency effects endogenously generate permanent disagreement in a population of ex ante identical agents. We also extend some of the results to allow for asymmetric preferences across states and irreducible noise.

2. A model of hidden states and costly channels

We start by presenting a simple setup in which a single agent wants to determine the state of the world. Different states of the world cannot be distinguished by direct observation and the agent instead needs to rely on noisy channels. The agent can choose how informative the channel is subject to a cost, and more informative channels are more costly.

Costly information acquisition can be modeled in several ways. We use the set up presented here to study two cost of information functions that both build on 1948's entropy concept. *Mutual information* is a measure of how much information an agent with a specific prior learns on average from observing a given signal. *Channel capacity* is a measure of how much any agent can learn from a given signal. It is defined as the maximum mutual information of a given signal over all priors. Stated differently, channel capacity measures how much a hypothetical minimally informed agent on average learns from a given signal.

To date, most of the literature has followed Sims (2003) and used mutual information to specify the cost of information, e.g. Mackowiak and Wiederholt (2009), Woodford (2009), Matějka (2015, 2016) and Matějka and McKay (2016). An exception is Woodford (2012) who uses channel capacity to model consumer choices. In this section, we formally define mutual information and channel capacity and discuss some of their properties that are important for understanding the results in the sections that follow.

2.1. States and agent utility.

Nature determines the state of the world $\omega \in \{0, 1\} \equiv \Omega$. An agent cannot observe the state of nature directly and we denote an agent's prior as $\pi \equiv p(\omega = 1)$.

An agent can pay a cost to observe a binary signal $s \in \{0, 1\}$ about the state of the world. We will refer to the information structure defined by the error probabilities ϵ_0 and ϵ_1 that determine the probability of observing a signal $s \neq \omega$ as a *channel* through which the agent gets information about the true state of nature.

Definition 1. *A channel* $S \in (0, 1) \times (0, 1)$ *is defined by the error probabilities* $\epsilon_0 \equiv p(s = 1 \mid \omega = 0)$ *and* $\epsilon_1 \equiv p(s = 0 \mid \omega = 1)$.

Conditional on the signal s, agents use Bayes rule to update their beliefs and we denote the posterior belief conditional on the signal s as $p(\omega \mid s)$. We assume, without loss of generality, that $\epsilon_0 + \epsilon_1 \leq 1$ so that $p(\omega = 1 \mid s = 1) \geq p(\omega = 1)$.

The agents make their decisions in two stages. In the first, agents choose the optimal precision of the channel. That is, they choose error probabilities ϵ_0 and ϵ_1 in order to maximize the objective function Φ

$$\max_S \Phi(S, \pi) = E[U(a^*, \omega) \mid S, \pi] - \theta\Gamma(S, \pi) \tag{2.1}$$

where Γ is the cost associated with the channel S. The parameter θ scales the cost function Γ.

In the second stage, after the signal s has been observed, agents choose their optimal action $a^* \in \{0, 1\}$ such that

$$a^* = \arg\max E[U(a, \omega) \mid s, \pi] \tag{2.2}$$

where

$$U(a, \omega) \equiv 1 - |\omega - a|. \tag{2.3}$$

An agent thus gets one unit of utility for taking the action $a = \omega$ and zero otherwise. Below, we will analyze the implications of two different specifications for the cost function Γ that both build on 1948's concept of entropy to quantify information.[4]

2.2. **Information, entropy and channel capacity.** 1948 defined the amount of information I contained in a message that is observed with probability p as

$$I \equiv \log \frac{1}{p} = -\log p. \tag{2.4}$$

This definition is appealing: Being told that something unexpected happened is more informative than being told about an event occurring that you already assigned a high probability to. Entropy $H(\Omega)$ of a source (or random variable) $\omega \in \Omega$ is defined as the expected amount of information

$$H(\Omega) = \sum_{\omega \in \Omega} p(\omega) \log \frac{1}{p(\omega)}. \tag{2.5}$$

In a binary state model, entropy is maximized when both states are equally likely, i.e. when $p(\omega = 1) = \frac{1}{2}$. Entropy is minimized when one state is certain, i.e when $p(\omega = 1)$ equals 0 or 1.

2.2.1. *Two candidate cost functions.* Conditional entropy $H(\Omega \mid S)$ quantifies how much uncertainty remains about Ω after observing S. It is defined as

$$H(\Omega \mid S) = \sum_{s \in S} \sum_{\omega \in \Omega} p(\omega, s) \log \frac{1}{p(\omega \mid s)}. \tag{2.6}$$

Conditional entropy (2.6) can be used to compute the mutual information $I(\Omega, S)$

$$I(\Omega, S) = H(\Omega) - H(\Omega \mid S) \tag{2.7}$$

[4]Alternatively, we could have let agents take an action in $(0, 1)$ with the aim of minimizing the variance $E(\omega - a)^2$. All the qualitative results of this paper would continue to hold, but we would loose tractability.

which quantifies how much information is revealed on average about Ω by observing S. In a binary state model, mutual information is maximized when $p(\omega = 1) = \frac{1}{2}$ and $p(\omega = 1 \mid s)$ equals 0 or 1.

An information channel $p(S \mid \Omega)$ is defined by a probability distribution of signals conditional on the state. The channel capacity C is the maximum reduction in entropy that can be achieved by applying the channel to a (source) distribution $p(\omega)$. In our binary state setting, it is thus defined as

$$C\left(p(S \mid \Omega)\right) = \max_{p(\omega=1)\in(0,1)} I(\Omega, S) \tag{2.8}$$

where the maximum is taken over all possible distributions $p(\omega = 1) \in (0, 1)$.[5] Below, we will study the optimal channel choices when the cost of information function Γ is proportional to mutual information as well as when it is proportional to channel capacity. To ease notation, we will denote the cost functions as, respectively $\Gamma_I \equiv I(\Omega, S)$ and $\Gamma_C \equiv C\left(p(S \mid \Omega)\right)$.

2.3. **Properties of mutual information and channel capacity.** Mutual information and channel capacity share many characteristics that make both of them attractive to use as information cost functions. However, as cost functions they also differ in ways that affect agents' optimal choice of signal structure. The properties of the two functions that are most important for our purposes are summarized by the following four lemmas.

Lemma 1. *Mutual information $I(\Omega, S)$ and channel capacity $C(\Omega, S)$ are both convex in the precision of signals $(1 - \epsilon_0)$ and $(1 - \epsilon_1)$.*

Proof. That mutual information is convex in the precision of signals is a standard result, e.g. Theorem 2.7.4 in Cover and Thomas (2006). That channel capacity is convex follows from that the (point-wise) maximum of convex functions is a convex function. □

Lemma 1 follows directly from the definitions of mutual information and channel capacity. It implies that regardless of whether we model the cost of information using mutual information or channel capacity, the marginal cost of information is increasing in the signal precisions $(1 - \epsilon_0)$ and $(1 - \epsilon_1)$. The lemma also implies that the marginal cost of precision in one state is increasing in the precision of the channel in the other state.

Mutual information $I(\Omega, S)$ and channel capacity $C(\Omega, S)$ thus share some properties, but they differ in at least one aspect that is important for agents' decisions in our setting. Mutual information depends directly on an agent's prior beliefs while channel capacity does not. The choice of cost function will thus affect how the optimal information channel depends on an agent's prior beliefs. The following lemma states a general property of the relationship between mutual information and the prior distribution that is useful for understanding an agent's choice of channel precision.

[5]The closed form expression for the channel capacity of a binary asymmetric channel is given by

$$\Gamma_C = \frac{\epsilon_0}{1 - \epsilon_0 - \epsilon_1} H(\epsilon_1) - \frac{1 - \epsilon_1}{1 - \epsilon_0 - \epsilon_1} H(\epsilon_0) + \log_2\left(1 + 2^{\frac{H(\epsilon_0) - H(\epsilon_1)}{1 - \epsilon_0 - \epsilon_1}}\right)$$

where $H(\epsilon)$ is the entropy of a binary random variable with probabilities given by ϵ and $1 - \epsilon$. See Moser et al. (2009).

Lemma 2. *The mutual information of the channel S is concave in the prior π everywhere on $(\epsilon_0, \epsilon_1) \in (0,1) \times (0,1)$*

Proof. See Theorem 2.7.4 in Cover and Thomas (2006). □

Again, Lemma 2 is a standard result and follows directly from manipulation of the definition (2.7). Importantly, the lemma implies that the marginal cost of information is decreasing as an agent's prior π tends to either 0 or 1. As mentioned above, the same is not true for channel capacity. Since channel capacity is defined as the maximum mutual information that can be achieved between a given channel S for any prior π, channel capacity cost is by construction independent of the precision of an agent's prior.

Lemma 3. *The channel capacity of a given channel S is independent of an agent's prior.*

Proof. The lemma follows immediately from the definition (2.8) of channel capacity. □

It is perhaps useful to point out here that, while channel capacity is defined as the maximum entropy reduction that can be achieved for any prior distribution, below agents condition only on their specific (and unique) prior π when choosing ϵ_0 and ϵ_1 in the signal structure. For our purposes, the relevance of the max operator in the definition of channel capacity is simply to make information cost a function only of channel precision and not depend on an agent's prior beliefs.

The last properties of mutual information and channel capacity that we discuss here relates to how these functions behave at the boundaries of the precision of the channel S and prior beliefs π.

Lemma 4. *Mutual information $I(\Omega, S)$ tends to zero when π tends to either 0 or 1 or when the channel S tends towards being completely uninformative. Channel capacity $C(\Omega, S)$ tend to zero only when the channel tend towards being completely uninformative.*

Proof. The mutual information between the signal S and the state Ω can be expressed as

$$I(\Omega, S) = \sum_{s \in \{0,1\}} p(s) \sum_{\omega \in \{0,1\}} p(\omega \mid s) \log \frac{p(\omega \mid s)}{p(\omega)}. \tag{2.9}$$

Clearly, (2.9) is only zero if and only if $p(\omega \mid s) = p(\omega)$ for every ω and s. Mutual information is therefore zero either when the prior $p(\omega)$ is dogmatic, i.e. when $p(\omega = 1)$ equals either 0 or 1, or when the signal s is independent of ω. Since channel capacity is defined as the maximum of $I(\Omega, S)$ over all priors $p(\omega)$, channel capacity is zero only when the signal s is independent of ω. □

Lemma 1 ensures that a solution to the agent's optimization problem exists, is unique and that an interior solution can be characterized by standard first order conditions. The properties described by Lemma 2, 3 and 4 are important for understanding why the two cost functions have different predictions about an agent's channel choice, and in extension, his beliefs. In particular, the fact that channel capacity is independent of an agent's prior while mutual information is not, is important for the results that follow.

3. Optimal channels and beliefs

In this section we study the optimal choice of information channels while taking an agent's prior belief π as given. We show that an agent chooses a channel that (i) is more accurate in the state that is a priori more likely, (ii) is less informative the more precise his prior is, and (iii) at some sufficiently precise prior, the chosen channel is completely uninformative. In the next section we generalize these results to a dynamic setting where an agent's prior in period $t+1$ is the posterior beliefs inherited from period t. In what follows, we denote the parameters of an optimally chosen channel S^* as ϵ_0^* and ϵ_1^*.

3.1. Priors and optimal channels.

Agents maximize the objective function (2.1) by choosing ϵ_0 and ϵ_1 while taking their prior belief π as given. Optimally chosen informative signals imply that it is always optimal for an agent to take the action $a = s$. This is a general property of choices subject to a binding entropy constraint: If an agent takes the same action in response to both signals, he can achieve the same utility at lower information cost by always observing the same signal. The expected utility of a channel is therefore equal to the expected probability of observing a correct signal

$$E\left(U \mid \epsilon_0, \epsilon_1, \pi\right) = \left(1-\pi\right)\left(1-\epsilon_0\right) + \pi\left(1-\epsilon_1\right). \tag{3.1}$$

An optimally chosen channel equates the expected marginal benefit of increasing the precision in a given state with the marginal cost of doing so. At the optimum, agents trade off a lower error probability in one state against a lower error probability in the other state. The first order conditions that describe the optimal choice of error probabilities at an interior solution are given by

$$\left(1-\pi\right) = -\theta\frac{\partial\Gamma}{\partial\epsilon_0}, \qquad \pi = -\theta\frac{\partial\Gamma}{\partial\epsilon_1}. \tag{3.2}$$

The relative expected marginal benefit of a more precise signal is higher in the state that the agent believes is more likely to occur. Because of this, agents display what we call the *confirmation effect*.

Proposition 1. *(Confirmation effect) The optimal channel is more precise in the state that is a priori more likely, i.e.* $\epsilon_0^* > \epsilon_1^*$ *if* $\pi > \frac{1}{2}$ *and* $\epsilon_0^* < \epsilon_1^*$ *if* $\pi < \frac{1}{2}$.

Proof. In the Appendix. □

Because the confirmation effect makes agents choose a signal that is more accurate in the a priori more likely state, the agent is also more likely to observe a signal that reinforces his prior beliefs about which state is more likely. That is, $\epsilon_0 > \epsilon_1$ implies that $p\left(s=1\right) > p\left(s=0\right)$ and vice versa. The confirmation effect holds regardless of whether information cost is proportional to mutual information or channel capacity.

A corollary of the confirmation effect is that when agents receive a disconfirming signal, i.e. the signal realization that is less likely given their prior beliefs, they will update their beliefs further. To see this, note that for a Bayesian agent the expected posterior must equal the prior. The expected posterior is equal to the sum of the probability of observing each signal times the posterior conditional on the signal in question. We thus have that

$$\pi = p\left(\omega=1 \mid s=1\right)p\left(s=1\right) + p\left(\omega=1 \mid s=0\right)p\left(s=0\right). \tag{3.3}$$

Rearranging this expression gives

$$\frac{p(s=1)}{p(s=0)} = \frac{\pi - p(\omega = 1 \mid s = 0)}{p(\omega = 1 \mid s = 1) - \pi} \tag{3.4}$$

so that $p(s=1) > p(s=0)$ implies that the belief revision $\pi - p(\omega = 1 \mid s = 0)$ is larger than $p(\omega = 1 \mid s = 1) - \pi$.

In the dynamic model studied in the next section, the confirmation effect reinforces differences in beliefs that may arise from different agents observing different signal realizations. Therefore, the confirmation effect makes it more likely that ex ante identical agents will disagree permanently. To see why, consider two agents that both assign probability $1/2$ to each state. Because the agents' priors are the same, they will therefore also choose the same signal structures. However, even with the same signal structures, the two agents may by chance observe different signal realizations and their posterior beliefs will then be different. Because of the confirmation effect, each agent then chooses a signal structure in the next period that makes it relatively more likely that their beliefs will continue to diverge.

The confirmation effect relates how the relative precision of signals in different states depends on the prior beliefs π. The optimal channel capacity also depends on the prior, but the two cost functions Γ_C and Γ_I imply qualitatively different relationships.

Proposition 2. *(Complacency effect) Under the cost function Γ_C, but not under Γ_I, an agent always chooses a less informative channel with lower capacity as the precision of his prior is increased.*

Proof. In the Appendix. □

To understand the intuition behind the proof of Proposition 2, note that by (3.1), the expected utility of a channel's precision in a given state is increasing linearly in the prior probability of the state in question. Under Γ_C, the cost of a channel is independent of the prior π and convex in the precision of the signals. The optimal channel then has lower capacity the more precise the prior is. We call the fact that agents choose less informative channels the more precise their prior is the *complacency effect* and it holds monotonically when cost is proportional to channel capacity.

Under Γ_I, the relationship between precision of the prior and the capacity of the optimal channel is not monotonic. Mutual information is bounded above by prior entropy, so the cost of any signal tends to zero as the prior become more precise. With a perfectly precise prior, any signal structure, including a perfectly revealing one, is thus consistent with optimizing behavior.

The next result shows that under Γ_C, the complacency effect is powerful enough so that for sufficiently precise priors, an agent will choose a completely uninformative channel.

Proposition 3. *(Information shutdown) When information cost is proportional to Γ_C, there exist sufficiently precise priors $\overline{\pi} \in [\frac{1}{2}, 1)$ and $\underline{\pi} \in (0, \frac{1}{2}]$ so that an agent will choose a completely uninformative channel for any $\theta > 0$.*

Proof. By symmetry, we only need to prove the statement for $\overline{\pi} \in [\frac{1}{2}, 1)$. The marginal utility of more precise signals in state $\omega = 0$ is equal to $(1 - \pi)$. As $\pi \to 1$ the marginal utility thus tends to zero. However, because Γ_C is non-negative and convex in the precision of signals,

the first order condition is slack at the limit. For some $\overline{\pi} < \pi$ it is thus optimal to choose $\epsilon_0 = 1$ and $\epsilon_1 = 0$. Since an agent then observes $s = 1$ with probability 1 in both states, the channel is uninformative (and has zero cost). □

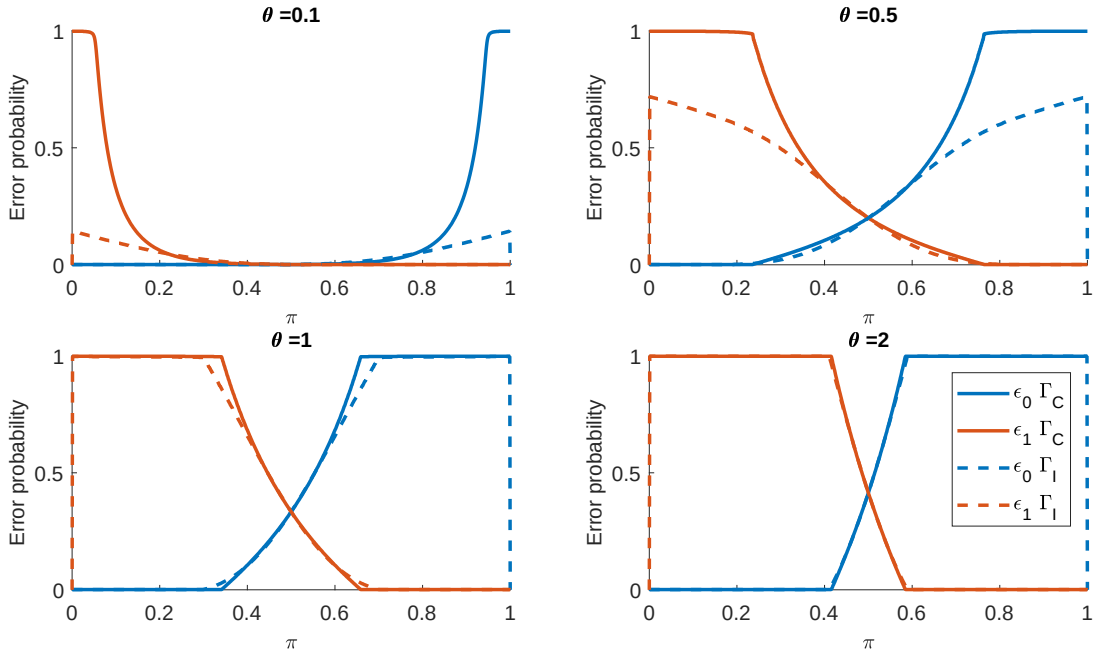

FIGURE 1. Optimal channel precision as a function of prior beliefs π for different values of the cost parameter θ. Solid lines indicate optimal error probabilities under the cost function Γ_C and dashed lines indicate optimal error probabilities under the cost function Γ_I.

The optimal channel choices are illustrated in Figure 1 where we have plotted the optimal error probabilities ϵ_0^* and ϵ_1^* under both Γ_C and Γ_I as functions of the prior π and for different values of θ. The confirmation effect states that as an agent becomes more certain about the state, he will increase the precision of his channel in the state that he thinks is more likely and decrease the precision in the other. In the figure, this is manifested by that ϵ_0^* (blue lines) is increasing in π and that ϵ_1^* (red lines) is decreasing in π. We can see that this effect is stronger under Γ_C (solid lines) than under Γ_I (dashed lines) by the fact that the ratio $\frac{\epsilon_0^*}{\epsilon_1^*}$ is lower under Γ_C than under Γ_I for all $\pi < \frac{1}{2}$, and symmetrically, that $\frac{\epsilon_1^*}{\epsilon_0^*}$ is lower under Γ_C than under Γ_I for all $\pi > \frac{1}{2}$.

Figure 1 also shows that under Γ_C, for sufficiently precise priors agents choose a channel that always report the signal that indicates that the more likely state is true. This holds for all values of θ and the channel is then completely uninformative. For large values of θ the optimal signal choices under Γ_I coincide with the optimal choices under Γ_C for all beliefs away from the boundaries. However, for lower values of θ, there are no prior beliefs for which agents choose uninformative signals under Γ_I. At the boundary, i.e. for π equal to 0 or 1,

the optimal signal structure is indeterminate under Γ_I since all signal structures are then costless.

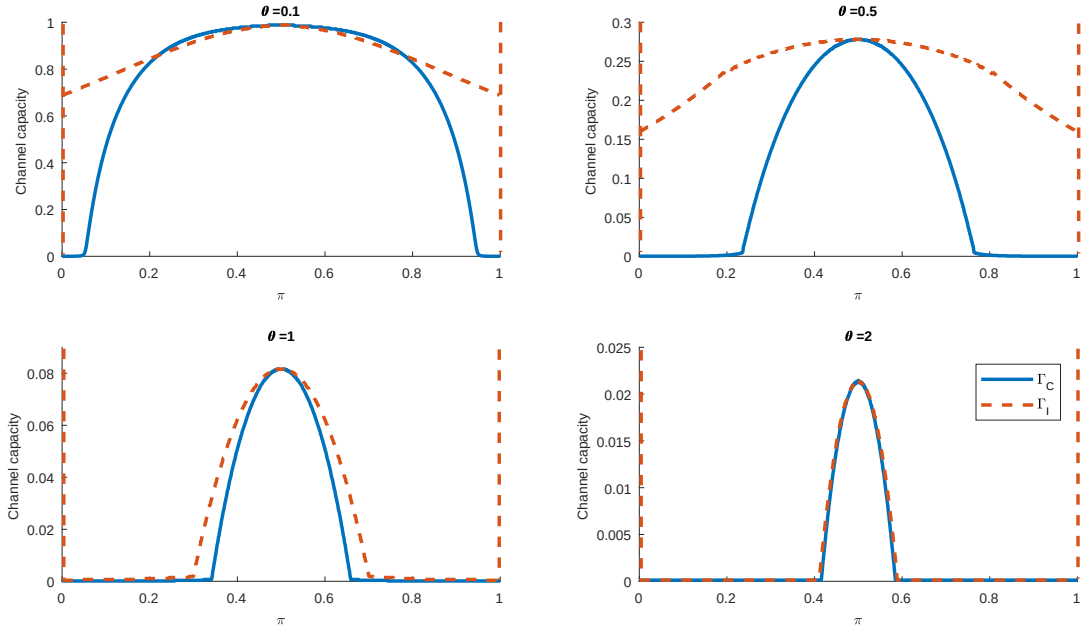

FIGURE 2. Channel capacity Γ_C as a function of prior beliefs π for different values of the cost parameter θ. Solid lines indicate the optimal channel capacity under the cost function Γ_C and dashed lines indicate optimal channel capacity under the cost function Γ_I.

In Figure 1 we could see that agents generally choose lower error probabilities under Γ_I than under Γ_C. This is so because the complacency effect is weaker (and not-necessarily monotonic in the precision of the prior) under Γ_I. Figure 2 illustrates the complacency effect more directly. There, we have plotted the optimal channel capacity under both Γ_C and Γ_I as functions of the prior π and for different values of θ. Regardless of cost function and the value of θ, an agent chooses the most informative channel when his prior is most uncertain, i.e. when $\pi = 1/2$. When an agent becomes more certain about the state, he will decrease the informativeness (and the cost) of the channel. Under Γ_C, there always exists sufficiently precise priors so that an agent chooses a completely uninformative channel with zero capacity for any $\theta > 0$. However, under Γ_I, this is only true for higher values of θ. For lower values of θ, there are no intervals of beliefs for which agents choose uninformative channels.

In this section we have shown that regardless of which cost function we use, an agent will choose a channel that is more precise in the state of the world that is a priori more likely. As the precision of the prior is increased, an agent whose information cost is proportional to channel capacity always chooses a less informative channel. We call this the complacency effect, and we showed that under Γ_C, it leads an agent to choose a completely uninformative channel when his prior becomes sufficiently precise. In the next section, where we study

a dynamic setting, we show that the confirmation and complacency effects make ex ante identical agents choose channels that lead to permanent disagreement about he state ω.

4. Optimal channels and divergence of beliefs

The analysis above treated the agent's prior π as given. Here, we study how an agent's beliefs evolve over time as he recursively chooses the precision of his channel. We also study how the distribution of beliefs across a population changes over time. In particular, we show that beliefs of different agents will endogenously diverge to a time-invariant distribution in which some agents permanently attach a higher probability to $\omega = 1$ while other agents permanently attach a higher probability to $\omega = 0$.

4.1. **The dynamic channel choice.** Time is discrete and indexed by $t = 0, 1, 2, ...$ The state ω is determined by nature at time zero. Agents are indexed by $i \in N$ and agent i enters period t with prior beliefs $p\left(\omega = 1 \mid s_i^{t-1}\right) \equiv \pi_{i,t-1}$, where s_i^{t-1} is the history of signals observed by agent i. In each period, agent i chooses the precision of the channel $S_{i,t}$, defined by the error probabilities $\epsilon_{0i,t} \equiv p(s_{i,t} = 1 \mid \omega = 0)$ and $\epsilon_{1i,t} \equiv p(s_{i,t} = 0 \mid \omega = 1)$. Agent i chooses his channel to maximize the sum of the within-period objective function Φ and the discounted expected continuation value. The optimization problem can be expressed recursively as a Bellman equation with the prior $\pi_{i,t}$ as the state variable

$$\begin{aligned}\mathbb{V}\left(\pi_{i,t-1}\right) &= \max_{\epsilon_{0i,t},\epsilon_{1i,t}} \{U\left(\epsilon_{0i,t}, \epsilon_{1i,t}, \pi_{i,t-1}\right) - \theta\Gamma(\epsilon_{0i,t}, \epsilon_{1i,t}, \pi_{i,t-1}) \\ &\quad + \delta E\left[\mathbb{V}\left(\pi_{i,t}\right) \mid \pi_{i,t-1}, \epsilon_{0i,t}, \epsilon_{1i,t}\right]\}. \end{aligned} \tag{4.1}$$

The parameter $\delta \in (0, 1)$ is the rate at which an agent discounts the future. $U\left(\epsilon_{0i,t}, \epsilon_{1i,t}, \pi_{i,t-1}\right)$ denotes the period t expected utility of the channel $S_{i,t}(\epsilon_{0i,t}, \epsilon_{1i,t})$ conditional on the prior $\pi_{i,t-1}$. The function $\theta\Gamma(\epsilon_{0i,t}, \epsilon_{1i,t}, \pi_{i,t-1})$ denotes the associated period t cost.

Proposition 4. *(Contraction mapping) The Bellman equation (4.1) is a contraction mapping.*

Proof. In the Appendix. □

Proposition 4 is proven using standard methods, e.g. Theorem 3.3 in Stokey and Lucas (1989). That the Bellman equation is a contraction mapping implies that there exists a unique solution to the optimal channel choice that can be found by function iteration.

4.2. **Optimal dynamic channel precision.** To understand how the dynamic channel choice differs from the static case, consider an agent who is about to choose the period t precision of his channel. The intra-period trade-offs are the same as in the static model, but the agent now also needs to consider the impact of his choices on future utility. The continuation value $\mathbb{V}\left(\pi_{i,t}\right)$ depends on his beliefs in the next period and the agent thus needs to take into account how his choices today affect the distribution of his next period beliefs.

A rational Bayesian agent's expected posterior must coincide his prior. That is, because of the law of iterated expectations, the expected posterior must be the same as his prior. The choice of precisions thus cannot affect the conditional mean of $\pi_{i,t}$. However, the choice of precision does affect the conditional variance of the posterior. As shown in Figure 3, the value function $\mathbb{V}$ is convex in π and thus increasing in the conditional variance of beliefs

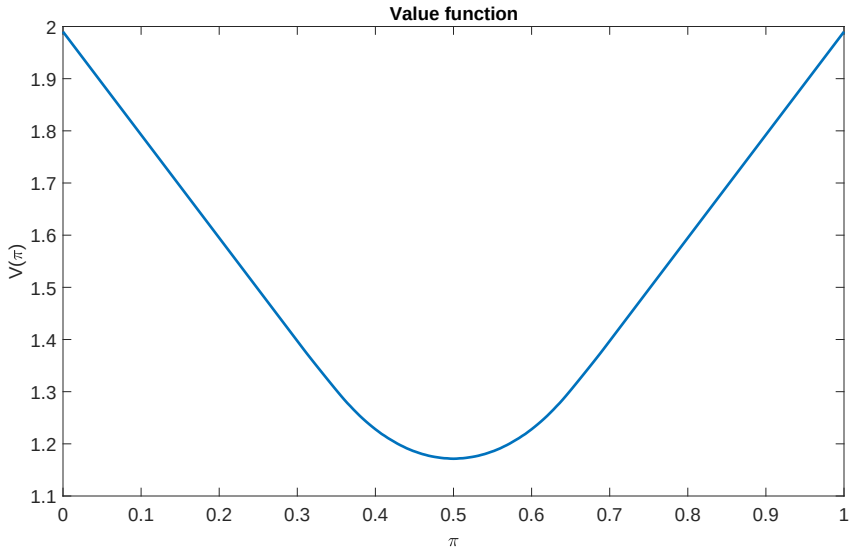

FIGURE 3. The value function $\mathbb{V}(\pi_{i,t})$

$\pi_{i,t}$. The conditional variance is in turn increasing in the precision of signals. To see why, note that with a perfectly uninformative channel, the prior will be unchanged, but with perfectly precise signals, the posterior will be at either boundary in $\{0, 1\}$. Choosing more precise signals today, thus increases the expected continuation value. Since the intra-period trade-offs are unchanged from the static model, the convexity of $\mathbb{V}$ implies that for a given prior, agents always choose a channel that is at least as informative period-by-period as the channel in a static model with the same cost parameter θ.[6] This is illustrated in Figure 4 and 5, where we have plotted, respectively, the optimal signal precisions and the optimal channel capacity for the model with parameters $\theta = 2$ and $\delta = 0.5$.

We can see from Figures 4 and 5 that when beliefs are sufficiently precise, with π_t either below 0.3 or above 0.7, agents choose completely uninformative signals also in the dynamic model. The intuition for the information shutdown in the dynamic model is the same as in the static set up. When the agent has a sufficiently precise prior entering a period, the marginal benefit of any informative channel is smaller than its marginal cost. For future reference, we will denote the information shut down points in the dynamic model as $\underline{\pi}^d$ and $\overline{\pi}^d$.

4.3. **Population dynamics of beliefs.** The dynamic model can be used to study how beliefs evolve over time. In particular, we will use it to show that agents endogenously sort into two distinct groups who permanently disagrees about which state is more likely. Figure 6 illustrates the implied belief dynamics for a population of agents under the assumption that the true state $\omega = 0$. The discounting and information cost parameters are set to $\delta = 0.5$ and $\theta = 2$ respectively. The dotted lines are the initial distribution of prior beliefs, the dashed lines are the distribution of priors in period 1 and the solid lines are the limiting distribution

[6]The distinction between a static and dynamic model is only relevant if $\delta \neq 0$.

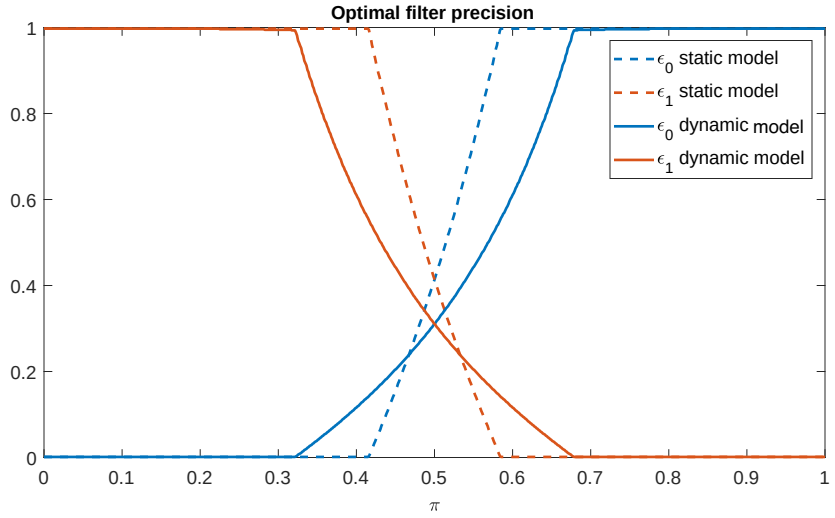

FIGURE 4. The optimal static (dashed lines) and dynamic (solid lines) signal precisions with $\theta = 2$ and $\delta = 0.5$ (for the dynamic model).

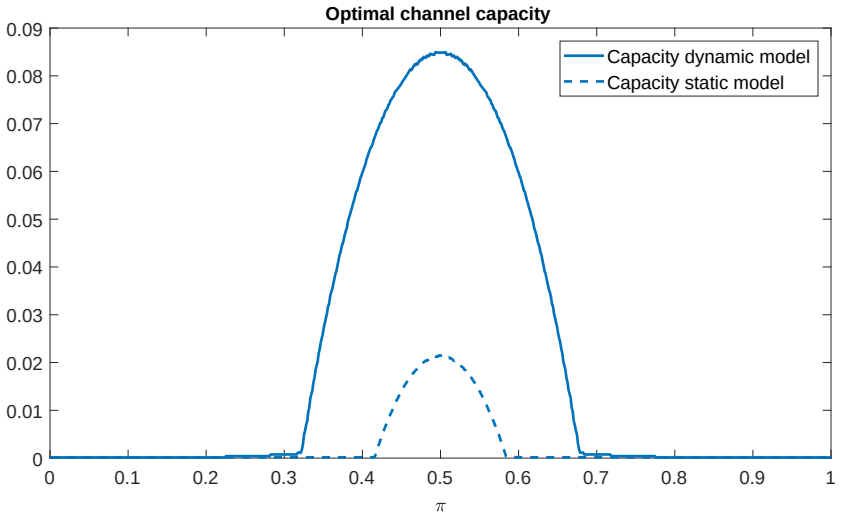

FIGURE 5. The optimal channel capacity in the static (dashed line) and dynamic (solid line) model with $\theta = 2$ and $\delta = 0.5$ (for the dynamic model).

of beliefs as $t \to \infty$. The top panel is initialized from a degenerate distribution where all agents starts from the same prior $\pi_{0,i} = \frac{1}{2}$ for all i. The bottom panel is initialized from a uniform distribution $\pi_{0,i} \sim U(0,1)$.

In the top panel, where the initial distribution put equal weight on both states, all agents choose the same error probabilities. A fraction $\left(1 - \epsilon^*_{0i,t}\right)$ then observes the signal $s = 0$ and the fraction $\epsilon^*_{0i,t}$ observes the signal $s = 1$, generating the two distinct groups of beliefs in

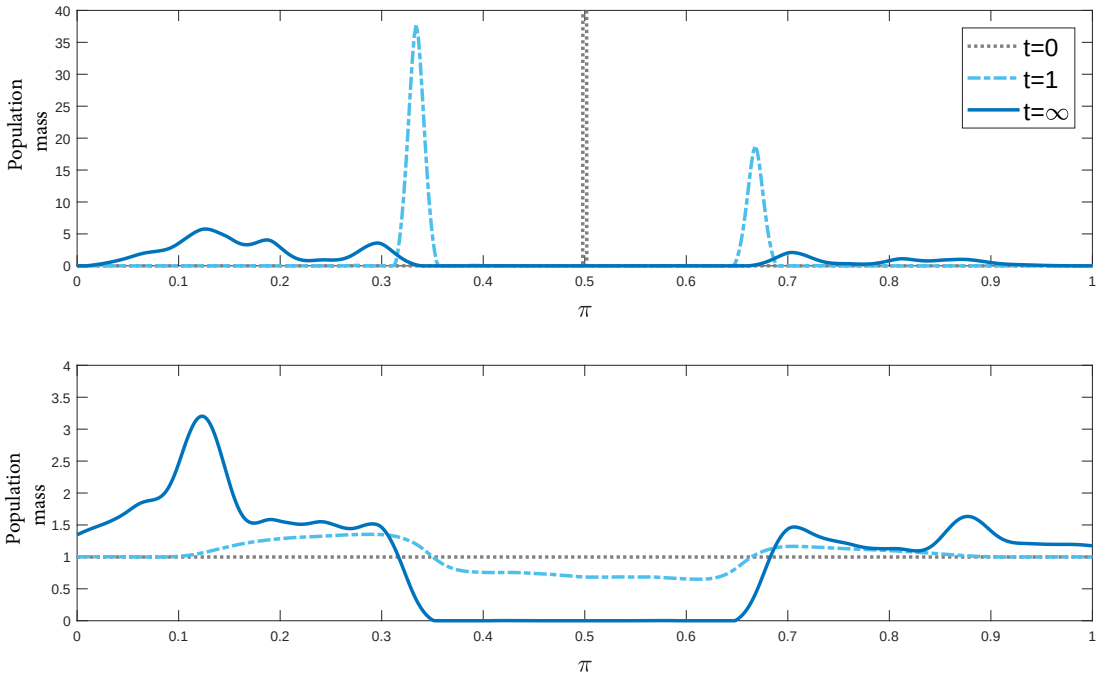

FIGURE 6. Belief dynamics with $\omega = 0$. The dotted lines are the initial distribution of beliefs, the dashed lines are the distribution of priors in period 1 and the solid lines are the limiting distribution as $t \to \infty$. In the top panel, the simulation is initialized from degenerate distribution of priors, i.e. $\pi_{i,0}(\omega) = 1/2 \forall i$. Bottom panel is initialized from a uniform distribution of priors, i.e. $\pi_{i,0}(\omega) \sim U(0,1)$.

period 1. Because signals are weakly accurate (that is, $\epsilon_0 + \epsilon_1 \neq 1$), a majority of agents update towards the correct belief that $\omega = 0$.

In the next period, agents re-optimize their channels, and choose more precise signals in the state that they believe is more likely. Once beliefs are precise enough, agents will choose completely uninformative channels and stop updating their beliefs. By chance, some agents will then have observed a sequence of incorrect signals and end up with beliefs $\pi_{i,t} > \overline{\pi}^d > \frac{1}{2}$ even though the true state is $\omega = 0$. These agents will choose a completely uninformative channel in the next period, and thus permanently believe that the incorrect state is more likely than the correct state. However, since agents are more likely to update their beliefs towards the correct state, this group is smaller than the group of agents with beliefs $\pi_{t,i} < \underline{\pi}^d < \frac{1}{2}$. Importantly, the limiting distribution is not degenerate and puts positive weight on both the intervals $\left(0, \underline{\pi}^d\right)$ and $\left(\overline{\pi}^d, 1\right)$. Ex ante identical agents may thus disagree permanently which state is more likely.

The bottom panel of Figure 6 illustrates the population belief dynamics starting from a uniform distribution on the interval $(0,1)$. After one period, we can see that the mass of beliefs have moved away from the midpoint toward more certain (but not necessarily more correct) beliefs. In the long run, the middle region of beliefs where agents choose informative

channels contains no mass at all. This is so because as long as an agent's beliefs are in the interval $\left(\underline{\pi}^d, \overline{\pi}^d\right)$, they will choose a weakly informative channel and therefore with positive probability move away from this region of beliefs and into the regions where they will choose uninformative signals. Once an agent holds beliefs that imply that it is optimal to choose uninformative signals, the agent never moves out of this region.

Again, the mass of beliefs in the region below the cut-off is larger, reflecting that any agent that choose a informative signal is more likely to have beliefs closer to the true state $\omega = 0$. However, the mass of agents with "wrong" beliefs above the cut-off $\overline{\pi}^d$ is sizeable. This group consists of agents with two different types of histories. Some agents' initial period zero priors were above $\overline{\pi}^d$ and these agents therefore never chose an informative channel and their beliefs therefore never changed. The second type had initial priors in the interval $\left(\underline{\pi}^d, \overline{\pi}^d\right)$, but by chance observed a sequence of incorrect signals and therefore ended up with beliefs above $\overline{\pi}^d$. Given the initial uniform distribution, the mass of agents that permanently assign a higher probability to the incorrect state is bounded from below by $1 - \overline{\pi}^d$. The net addition to this group of agents that are there because they observed incorrect signals is represented in the figure by the mass to the right of the cut-off $\overline{\pi}^d$ and between the dotted and solid lines.

The limiting distribution of the population beliefs have positive mass in the two discrete intervals near the boundaries of the belief space where agents choose to no longer acquire informative signals. Both the absolute and the relative size of these intervals depend on the cost parameter θ. For a lower value of θ, i.e. for lower cost of channel precision, the width of both intervals shrink as there is then a wider set of beliefs for which agents choose an informative channel. On the intensive margin, disagreement thus increases when information is cheaper, since the average distance between agents in the two intervals then increases. However, for a given belief, an agent also chooses a more precise channel when θ is lower. This means that fewer agents end up with incorrect beliefs when information is cheaper. Cheaper information thus decreases permanent disagreement on the extensive margin.

5. Extensions

In this section we briefly discuss two extensions of the set up described above where we allow for (i) asymmetric preferences and (ii) irreducible noise in the channels.

5.1. **Asymmetric preferences across states.**

In the benchmark model, we assumed that agents care equally about taking the right action in both states of the world. In some situations it may be more costly to be wrong in one state than in the other. To take such asymmetries into account, we can modify the utility function as follows

$$U(a, \omega) \equiv 1 - (1 - \omega) |\omega - a| - \gamma\omega |\omega - a| . \tag{5.1}$$

The expected utility is then given by

$$E(U \mid \epsilon_0, \epsilon_1, \pi) = (1 - \pi)(1 - \epsilon_0) + \gamma\pi(1 - \epsilon_1) . \tag{5.2}$$

For $\gamma > 1$, this implies that agents care more about taking the right action when $\omega = 1$ and that the relative marginal utility of more precision is tilted towards that state. Agents will then choose channels that are relatively more precise when $\omega = 1$. With $\gamma \neq 1$, the regions of beliefs where agents choose completely uninformative channels become then asymmetric,

and for $\gamma > 1$ we have that $1 - \overline{\pi} < \underline{\pi}$. Beliefs thus have to be more certain in the direction towards the state for which the preferences are tilted for information acquisition to shut down.

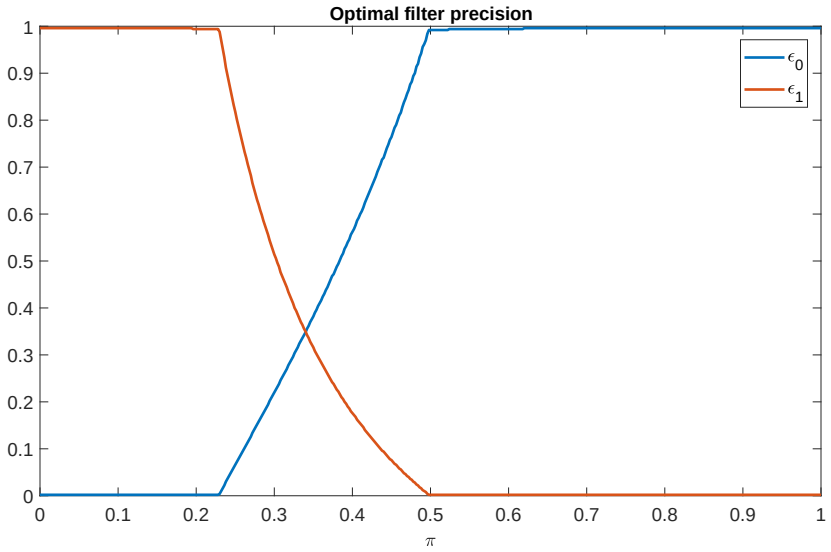

FIGURE 7. Optimal channel precision as a function of prior beliefs π and with cost parameter $\theta = 1$ when preferences are asymmetric $\gamma = 2$.

The consequences of asymmetric preferences for optimal signal precisions are illustrated in Figure 7. The asymmetry parameter is set to $\gamma = 2$ so that agents attach a higher value to taking the correct action when $\omega = 1$. We can see that the region of priors for which the agent is certain to observe the signal $s_{i,t} = 1$ then shifts left, relative to Figure 1. This means that for a larger region of priors, an agent observes a signal that confirms that the state where it is more costly to make a mistake has occurred. This is intuitive. With $\gamma > 1$, the cost of a observing an incorrect signal and taking the wrong action is higher when $\omega = 1$ than when $\omega = 0$. An agent is then willing to more often incorrectly believe that $\omega = 1$ since that is less costly than to incorrectly believe that $\omega = 0$.

Figure 8 illustrates the implications of asymmetric preferences for the belief dynamics. As in the Figure 3, the true state is $\omega = 0$ and the discounting and information cost parameters are set to $\delta = 0.5$ and $\theta = 2$, respectively. Compared to Figure 6, where agents put equal weight on taking the correct action in both states, more agents end up with beliefs that are permanently assigning a higher probability to the incorrect state. When it is more important for an agent to take the right action in one state, rather than the other, an agent is more likely to incorrectly attach a higher probability to the state where mistakes are more costly.

The main implication of asymmetric preferences is thus that an agent is more likely to believe that the state in which he is more concerned about taking the correct action is more likely than the state he is less concerned about. Asymmetric preferences also affect belief heterogeneity. If a majority of agents are more concerned about taking the right action in the state that is objectively not true, the mass of agents who attaches more weight to the ex

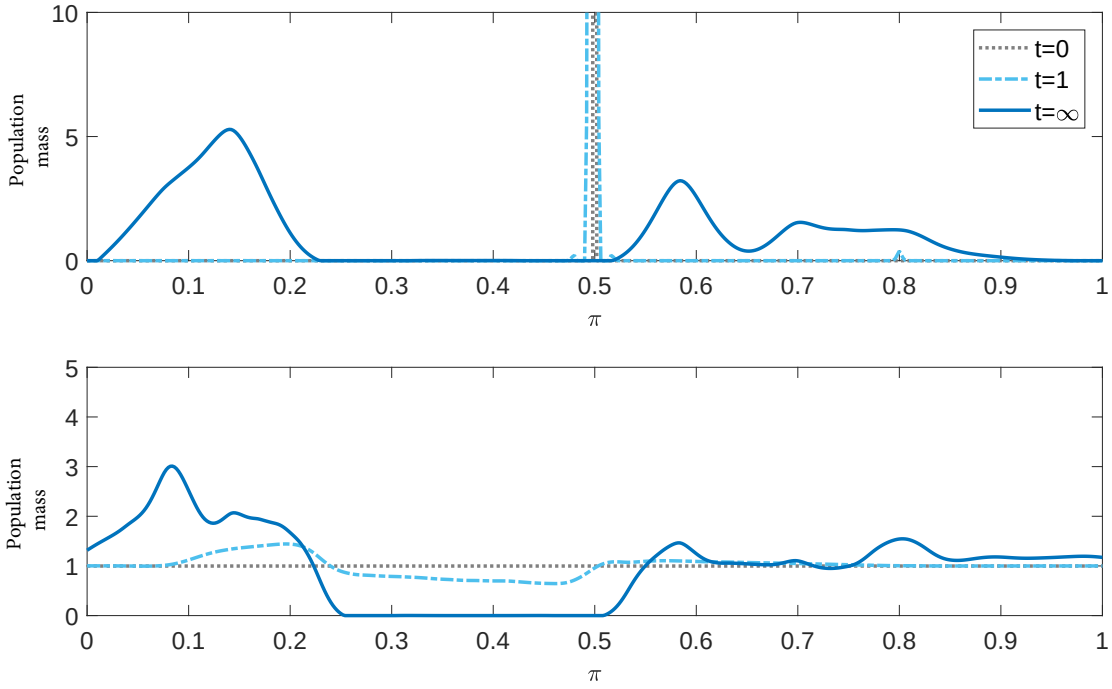

FIGURE 8. Belief dynamics with $\omega = 0$ and asymmetric preference parameter $\gamma = 2$. The dotted lines are the initial distribution of beliefs, the dashed lines are the distribution of priors in period 1 and the solid lines are the limiting distribution as $t \to \infty$. In the top panel, the simulation is initialized from degenerate distribution of priors, i.e. $\pi_{i,0}(\omega) = 1/2 \forall i$. Bottom panel is initialized from a uniform distribution of priors, i.e. $\pi_{i,0}(\omega) \sim U(0,1)$.

post incorrect state will increase and potentially increase belief heterogeneity. If the majority attaches more weight to the true state, disagreement decreases.

5.2. **Irreducible noise.** In some circumstances, it may be reasonable to think that the channels available to agents contain a noise component that is irreducible. To model such situations, we use the modified channel $\widetilde{S} \in (0,1) \times (0,1)$ defined by the probabilities $(1-\epsilon_0)(1-\widetilde{\epsilon}_0) \equiv p(s=0 \mid \omega = 0)$ and $(1-\epsilon_1) \times (1-\widetilde{\epsilon}_1) \equiv p(s=1 \mid \omega = 1)$ where $\widetilde{\epsilon}_0, \widetilde{\epsilon}_1 \in (0,1)$. The parameters $\widetilde{\epsilon}_0$ and $\widetilde{\epsilon}_1$ thus determine an exogenous upper bound on the precision of the channel $\widetilde{S}$. Expected utility is now given by

$$E(U \mid \epsilon_0, \epsilon_1, \pi) = (1-\pi)(1-\widetilde{\epsilon}_0)(1-\epsilon_0) + \pi(1-\widetilde{\epsilon}_1)(1-\epsilon_1). \tag{5.3}$$

The marginal utility of precision is thus decreasing in $\widetilde{\epsilon}_0$ and $\widetilde{\epsilon}_1$. If the cost of information is unchanged relative to Γ_C as in Section 2, introducing irreducible noise then makes agents choose less informative signals. The reason is that the value of noise reduction has decreased, so for the same cost function agents will choose a less informative channel. The range of beliefs for which agents choose completely uninformative channels expands and the complacency effect still holds.

If $\widetilde{\epsilon}_0 \neq \widetilde{\epsilon}_1$, the complacency effect affects the choice of ϵ_0 and ϵ_1 asymmetrically. If the irreducible noise when $\omega = 0$ is much higher than when $\omega = 1$, an agent will pay more attention to state 1 than state 0, as the marginal benefit of doing so will be higher. Conditional on a given level of irreducible noise, agents allocate more precision to the state that they believe is more likely. The confirmation effect thus still holds, conditional on fixed values of $\widetilde{\epsilon}_0$ and $\widetilde{\epsilon}_1$.

6. Conclusions

In this paper we have studied optimal information choice in a setting where agents can choose how precise their signals are in different states of the world. We showed that the optimal information choice displays what we call the *confirmation effect*, by which agents choose channels that are more precise in the state that they believe to be a priori more likely. When information cost is proportional to channel capacity, agents uniformly choose channels that are less informative the more precise their priors are. We call this effect the *complacency effect*. The complacency effect can be strong enough so that agents stop updating their beliefs entirely when their priors are sufficiently precise. This is not a limit result: Information shutdown can occur even when an agent still attaches a substantial probability to both states and it may also occur for agents that attached a larger probability to the incorrect state than to the correct state. Ex ante identical agents may then disagree permanently, and a population of agents may cluster in two distinct groups near the opposite boundaries of the belief space.

In the model, the true state of the world is not directly revealed to the agents through the utility they experience from their actions. The proposed model is thus suitable to study disagreement about facts that agents cannot evaluate the truth of immediately. Consider for instance the question of whether climate change caused by human activity is real. For most people, there is nothing in their everyday experience that would directly indicate whether the scientific consensus or the climate change skeptics are right. However, their optimal voting and consumption behavior may depend on whether global warming is real and man made. Individuals who initially attach a higher probability to climate change being real will seek out information from sources that are likely to be right, if in fact, climate change is real. However, individuals who initially believe that it is a Chinese hoax perpetrated in conspiracy with liberal elites and special interests at universities, will seek out information from sources that are more likely to be accurate if these beliefs are correct. This paper's mechanism can thus explain why even a rational population may self-sort into different media bubbles.

In this paper we used two different functions to model the cost of acquiring information. Both of them build on 1948's concept of entropy and both imply that attention is a scarce resource that should be allocated rationally. The main difference between them arises from the fact that observing a given signal leads to small reduction in entropy for a well-informed agent while the channel capacity of a given signal structure is independent of an agents prior. In this paper, we argued that modeling cost as proportional to channel capacity captures the time cost of reading articles or observing or listening to broadcasts for which the agent does not have complete control over the content. This may be of independent interest to some readers.

References

Afrouzi, H. and C. Yang (2016). Dynamic inattention, the phillips curve, and forward guidance. Technical report, Working Paper.

Ambuehl, S. (2016). An offer you can't refuse? incentives change how we think. Technical report, working paper, Stanford University.

Baliga, S., E. Hanany, and P. Klibanoff (2013). Polarization and ambiguity. *..., Kellogg CMS-EMS ...*.

Bartoš, V., M. Bauer, J. Chytilová, and F. Matějka (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review 106*(6), 1437–75.

Baumeister, R. F. and L. S. Newman (1994, February). Self-Regulation of Cognitive Inference and Decision Processes. *Personality and Social Psychology Bulletin 20*(1), 3–19.

Benoît, J.-P. and J. Dubra (2017). A theory of rational attitude polarization.

Berk, R. H. (1966, February). Limiting Behavior of Posterior Distributions when the Model is Incorrect. *The Annals of Mathematical Statistics 37*(1), 51–58.

Besley, T. and A. Prat (2006). Handcuffs for the grabbing hand? media capture and government accountability. *The American Economic Review 96*(3), 720–736.

Blackwell, D. and L. Dubins (1962, September). Merging of Opinions with Increasing Information. *The Annals of Mathematical Statistics 33*(3), 882–886.

Cover, T. M. and J. A. Thomas (2006). Elements of information theory 2nd edition.

Cukierman, A. and M. Tommasi (1998). When does it take a Nixon to go to China? *American Economic Review 88*(1), 180–197.

Freedman, D. A. (1963, December). On the Asymptotic Behavior of Bayes' Estimates in the Discrete Case. *The Annals of Mathematical Statistics 34*(4), 1386–1403.

Fryer, R., P. Harms, and M. Jackson (2013). Updating Beliefs with Ambiguous Evidence: Implications for Polarization. *Available at SSRN 2263504*.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of political Economy 114*(2), 280–316.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *The American economic review 70*(3), 393–408.

Hong, H. and M. Kacperczyk (2010). Competition and bias. *The Quarterly Journal of Economics 125*(4), 1683–1725.

Ilut, C. and R. Valchev (2017). Economic agents as imperfect problem solvers.

Koszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133.

Lord, C. G., L. Ross, and M. R. Lepper (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology 37*(11), 2098–2109.

Maćkowiak, B., F. Matějka, and M. Wiederholt (2018). Dynamic rational inattention: Analytical results. *Journal of Economic Theory 176*, 650–692.

Maćkowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *The American Economic Review 99*(3), 769–803.

Maćkowiak, B. and M. Wiederholt (2018). Lack of preparation for rare events. *forthcoming, Journal of Monetary Economics*.

Matějka, F. (2015). Rigid pricing and rationally inattentive consumer. *Journal of Economic Theory 158*, 656–678.

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

McCright, A. M. and R. E. Dunlap (2011). The politicization of climate change and polarization in the american public's views of global warming, 2001–2010. *The Sociological Quarterly 52*(2), 155–194.

Miller, R. and C. Sanchirico (1999). The Role of Absolute Continuity in Merging of Opinions and Rational Learning. *Games and Economic Behavior 29*(1-2), 170–190.

Moser, S. M., P. Chen, and H. Lin (2009). Error probability analysis of binary asymmetric channels. *Dept. El. & Comp. Eng., Nat. Chiao Tung Univ*.

Mullainathan, S. and A. Shleifer (2005). lthe market for newsm. *American Economic Review 95*(4), 1005–1030.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology 2*(2), 175–220.

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

Perego, J. and S. Yuksel (2017). Media competition and the source of disagreement. Technical report, mimeo, New York University, New York.

Rabin, M. and J. Schrag (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics* (February).

Resnick, P., R. K. Garrett, T. Kriplean, S. A. Munson, and N. J. Stroud (2013). Bursting your (filter) bubble: strategies for promoting diverse exposure. In *Proceedings of the 2013 conference on Computer supported cooperative work companion*, pp. 95–100. ACM.

Savage, L. J. (1954). *The foundations of statistics*. Dover Books.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal 27*(3), 379–423.

Sims, C. A. (1998). Stickiness. In *Carnegie-Rochester Conference Series on Public Policy*, Volume 49, pp. 317–356. Elsevier.

Sims, C. a. (2003, April). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Steiner, J., C. Stewart, and F. Matějka (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica 85*(2), 521–553.

Stokey, N. L. and R. Lucas (1989). *Recursive methods in economic dynamics*. Harvard University Press.

Suen, W. (2004). The Self Perpetuation of Biased Beliefs. *The Economic Journal 114*, 377–396.

Sundaresan, S. (2017). Rare events and the persistence of uncertainty. Technical report, working paper, Imperial College London.

Woodford, M. (2009). Information-constrained state-dependent pricing. *Journal of Monetary Economics 56*, S100–S124.

Woodford, M. (2012). Inattentive Valuation and Reference-Dependent Choice. *Working Paper*.

Appendix A. Proof of Proposition 1

We need to prove that $\pi > 1/2$ implies that $\epsilon_1 < \epsilon_0$ in equilibrium for both Γ_I and Γ_C. For both proofs, we will use that an agent always take the action $a = s$. This is a general property of choices subject to a binding entropy constraint: If an agent takes the same action in response to both signals, he can achieve the same utility at lower information cost by always observing the same signal.

When information cost is given by Γ_I the agent's decision problem is a special case of the discrete choice framework studied in McKay and Matejka (2015). From Proposition 3 (and Equation 1) of their paper it follows that the probability of taking the correct action in state 1 can be expressed as

$$\frac{p\left(a=1 \mid \omega=1\right)}{p\left(a=0 \mid \omega=1\right)} = e^{(1+\alpha_1)} \tag{A.1}$$

where α_1 is a term that is positive if $\pi > \frac{1}{2}$. Using that $p\left(a=1 \mid \omega=1\right) = 1-\epsilon_1$ and the symmetry of the problem, we get

$$\begin{aligned} \frac{1-\epsilon_1}{\epsilon_1}\frac{\epsilon_0}{1-\epsilon_0} &= \frac{e^{\left(\frac{1}{\theta}+\alpha_1\right)}}{e^{\left(\frac{1}{\theta}-\alpha_1\right)}} \\ &> 1 \end{aligned}$$

implying that $\epsilon_1 < \epsilon_0$ which completes the proof for Γ_I.

By symmetry and convexity of the problem, it follows that when $\pi = \frac{1}{2}$ we have $\epsilon_1 = \epsilon_0$. It is thus sufficient to show $\frac{\partial \epsilon_1}{\partial \pi} < 0$ everywhere. For Γ_C we can use that the first order conditions

$$\frac{\partial E\left(U\right)}{\partial \epsilon_0} - \theta\frac{\partial \Gamma}{\partial \epsilon_0} = 0 \tag{A.2}$$

$$\frac{\partial E\left(U\right)}{\partial \epsilon_1} - \theta\frac{\partial \Gamma}{\partial \epsilon_1} = 0 \tag{A.3}$$

together with the implicit function theorem implies that

$$\frac{\partial \epsilon_i}{\partial \pi} = -\frac{\frac{\partial^2 E(U)}{\partial \epsilon_i \partial \pi} - \theta\frac{\partial^2 \Gamma}{\partial \epsilon_i \partial \pi}}{\frac{\partial^2 E(U)}{\partial \epsilon_i^2} - \theta\frac{\partial^2 \Gamma}{\partial \epsilon_i^2}}.$$

Since the expected utility is therefore equal to the probability of receiving the correct signal, i.e.

$$E\left(U \mid \epsilon_0, \epsilon_1, \pi\right) = \left(1-\pi\right)\left(1-\epsilon_0\right) + \pi\left(1-\epsilon_1\right) \tag{A.4}$$

we then have that

$$\frac{\partial \epsilon_1}{\partial \pi} = -\frac{1}{-\theta\frac{\partial^2 \Gamma}{\partial \epsilon_i^2}} \tag{A.5}$$

$$< 0 \tag{A.6}$$

since $\frac{\partial^2 U}{\partial \epsilon_i^2} = 0$, $\frac{\partial^2 \Gamma_C}{\partial \epsilon_i \partial \pi} = 0$ and $\frac{\partial^2 \Gamma_C}{\partial \epsilon_i^2} < 0$ which completes the proof. For the final step we used that expected utility is linear in error probabilities, that channel capacity Γ_C is by definition independent of the prior π and convex in precision.

Appendix B. Proof of Proposition 2

We need to show that for a more precise prior, i.e. a π closer to 0 or 1, agents will choose a less informative channel with lower cost Γ_C. The first order conditions are

$$\frac{\partial \Gamma(\epsilon_0,\epsilon_1)}{\partial \epsilon_0} = -\frac{\pi}{\theta} \qquad \frac{\partial \Gamma(\epsilon_0,\epsilon_1)}{\partial \epsilon_1} = -\frac{1-\pi}{\theta} \tag{B.1}$$

These expressions are linear in π, and convex in ϵ_0 and ϵ_1, respectively. Consider the solution when $\pi = \frac{1}{2}$. By the symmetry of the problem, the optimal solution must entail that $\epsilon_0 = \epsilon_1 = \bar{\epsilon}$. Then consider any other value of $\pi = \pi'$. Without loss of generality, consider the solution at $\pi' > \frac{1}{2}$. The marginal utility of $\frac{\pi'}{\theta}$ and $\frac{1-\pi'}{\theta}$ will differ from the $\pi = \frac{1}{2}$ solution in equal and opposite measure due to the linearity of the first order conditions. That is:

$$\frac{1-\pi'}{\theta} < \frac{1}{2\theta} < \frac{\pi'}{\theta} \qquad \text{and} \qquad \frac{1-\pi'}{\theta} - \frac{1}{2\theta} = \frac{1}{2\theta} - \frac{\pi'}{\theta}$$

The marginal costs, however, are convex, and will change convexly in opposite directions. That is:

$$\frac{\partial \Gamma(\epsilon_0',\epsilon_1')}{\partial \epsilon_0'} < \frac{\partial \Gamma(\bar{\epsilon},\bar{\epsilon})}{\partial \epsilon_0} < \frac{\partial \Gamma(\epsilon_0',\epsilon_1')}{\partial \epsilon_1'} \qquad \text{and} \qquad \frac{\partial \Gamma(\bar{\epsilon},\bar{\epsilon})}{\partial \epsilon_0} - \frac{\partial \Gamma(\epsilon_0',\epsilon_1')}{\partial \epsilon_0'} < \frac{\partial \Gamma(\epsilon_0',\epsilon_1')}{\partial \epsilon_1'} - \frac{\partial \Gamma(\bar{\epsilon},\bar{\epsilon})}{\partial \epsilon_0}$$

By Jensen's inequality the new average marginal cost is higher, but the new average marginal benefit is the same as before. Therefore, the cost must adjust downward, which implies that the total amount of information is lower. Because Jensen's inequality increases in magnitude the larger the deviation, the lowering is monotonic in $|\pi' - \pi|$.

Appendix C. Proof of Proposition 4

Here we show that the Bellman equation

$$\begin{aligned}\mathbb{V}(\pi_{i,t-1}) = \max_{\epsilon_{0i,t},\epsilon_{1i,t}} \{ & U(\epsilon_{0i,t},\epsilon_{1i,t},\pi_{i,t-1}) - \theta\Gamma(\epsilon_{0i,t},\epsilon_{1i,t},\pi_{i,t-1}) \\ & + \delta E\left[\mathbb{V}(\pi_{i,t}) \mid \pi_{i,t-1},\epsilon_{0i,t},\epsilon_{1i,t}\right]\}. \end{aligned} \tag{C.1}$$

satisfies Blackwell's sufficient conditions and thus describes a contraction mapping, see Stokey and Lucas (1989). The value function $\mathbb{V}$ is bounded since the period utility cannot be lower than zero or larger than one. The value function thus describes a mapping from $(0,1)$ to the interval $\left(0, \frac{1}{1-\delta}\right)$, implying that the Bellman equation (C.1) describes a self-map on the space of bounded functions $\mathcal{B}(X)$. Blackwell's Theorem states that a mapping $T : \mathcal{B}(X) \to \mathcal{B}(X)$ is a contraction mapping with contractive constant β if the following two conditions are met:

1. Monotonicity: If $f, g \in \mathcal{B}(X)$ and $f(x) \leq g(x)$ for all $x \in X$ implies that $Tf(x) \leq Tg(x)$ for all $x \in X$.
2. Discounting: For $a \in \mathbb{R}_+$ there exists a β such that for all $f \in \mathcal{B}(X)$ and all $x \in X$

$$T(f+a)(x) \leq Tf(x) + \beta a. \tag{C.2}$$

To verify that the Bellman equation satisfies Blackwell's conditions, we introduce the following notation. Let $\Phi_f(\pi)$ denote the period value of the objective function Φ for $\epsilon_0, \epsilon_1 = \arg\max f$ and denote the probability distribution of π_t conditional on π_{t-1} and $\epsilon_0, \epsilon_1 =$

$\arg\max\ f$ as $p\left(\pi^{'} \mid \pi, \epsilon_0^f, \epsilon_1^f\right)$. The set of possible posteriors π_t given $\pi_{t-1}, \epsilon_0^f, \epsilon_1^f$ are denoted Π_f. We use analogous notation for the corresponding objects for g.

To show the monotonicity property, use that

$$\begin{aligned} Tf(\pi) &= \max_{\epsilon_0,\epsilon_1\in(0,1)} \left\{ \Phi_f(\pi) + \delta \sum_{\pi\in\Pi_f} f(\pi') p\left(\pi^{'} \mid \pi, \epsilon_0^f, \epsilon_1^f\right) \right\} && \text{(C.3)} \\ &\leq \max_{\epsilon_0,\epsilon_1\in(0,1)} \left\{ \Phi_f(\pi) + \delta \sum_{\pi\in\Pi_f} g(\pi') p\left(\pi^{'} \mid \pi, \epsilon_0^f, \epsilon_1^f\right) \right\} && \text{(C.4)} \\ &\leq \max_{\epsilon_0,\epsilon_1\in(0,1)} \left\{ \Phi_g(\pi) + \delta \sum_{\pi\in\Pi_g} g(\pi') p\left(\pi^{'} \mid \pi, \epsilon_0^g, \epsilon_1^g\right) \right\} && \text{(C.5)} \\ &= Tg(\pi) && \text{(C.6)} \end{aligned}$$

where the first equality follows from the definition of the Bellman equation, the second line follows from the condition that $f(x) \leq g(x)$ for all $x \in X$, the third line follows from that if agents maximize the value function taking $g(\pi')$ as given, they cannot do worse than the choice implied by $\Phi_f(\pi)$. The last equality follows from the definition of the Bellman equation and thus verifies the monotonicity conditions.

To show the discounting property, use

$$\begin{aligned} T(f+a)(x) &= \max_{\epsilon_0,\epsilon_1\in(0,1)} \left\{ \Phi_f(\pi) + \delta \left[\sum_{\pi\in\Pi_f} f(\pi') p\left(\pi^{'} \mid \pi, \epsilon_0^f, \epsilon_1^f\right) + a \right] \right\} && \text{(C.7)} \\ &= Tf(x) + \delta a && \text{(C.8)} \end{aligned}$$

The Bellman equation (C.1) is thus a contraction mapping, implying that there exists a unique fixed point that can be found by function iteration.

Appendix D. Deriving capacity of binary asymmetric channel

Mutual information between the state Ω and the signal S is defined as

$$I(\Omega; S) = H(\Omega) - H(\Omega \mid S). \tag{D.1}$$

Substituting in the expression for $H(\Omega)$ and $H(\Omega \mid S)$ then gives

$$I(\Omega, S) = \sum_{\omega\in\Omega} p(\omega) \log \frac{1}{p(\omega)} - \sum_{s\in S} \sum_{\omega\in\Omega} p(\omega, s) \log \frac{1}{p(\omega \mid s)}. \tag{D.2}$$

To compute the channel capacity of our binary asymmetric channel we need to find the probability $q \equiv p(\omega = 1)$ that maximizes this expression. Use that

$$I(\Omega, S) = H(S) - H(S \mid \Omega) \tag{D.3}$$

so that

$$I(\Omega, S) = h\left(q(1-\varepsilon_0) + (1-q)\varepsilon_1\right) - qh(\varepsilon_0) - (1-q)h(\varepsilon_1). \tag{D.4}$$

where $h(x)$ is the entropy of a binary variable with probability x. Rearrange to get

$$I(\Omega,S)=h\left(q\left(1-\varepsilon_0-\varepsilon_1\right)+\varepsilon_1\right)-q\left(h\left(\varepsilon_0\right)-h\left(\varepsilon_1\right)\right)-h\left(\varepsilon_1\right) \tag{D.5}$$

and take the derivative with respect to q to get

$$I_d'(\Omega,S)=\left(1-\varepsilon_0-\varepsilon_1\right)\log_2\left(\frac{1}{q\left(1-\varepsilon_0-\varepsilon_1\right)+\varepsilon_1}-1\right)-\left(h\left(\varepsilon_0\right)-h\left(\varepsilon_1\right)\right) \tag{D.6}$$

Set $I_d'(\Omega,S)=0$ and rearrange

$$\frac{1}{q\left(1-\varepsilon_0-\varepsilon_1\right)+\varepsilon_1}-1=2^{\frac{(h(\varepsilon_0)-h(\varepsilon_1))}{(1-\varepsilon_0-\varepsilon_1)}} \tag{D.7}$$

solve for q

$$q=\frac{1}{\left(1-\varepsilon_0-\varepsilon_1\right)}\left(\frac{1}{2^{\frac{(h(\varepsilon_0)-h(\varepsilon_1))}{(1-\varepsilon_0-\varepsilon_1)}}+1}-\varepsilon_1\right). \tag{D.8}$$

Define

$$z\equiv 2^{\frac{(h(\varepsilon_0)-h(\varepsilon_1))}{(1-\varepsilon_0-\varepsilon_1)}} \tag{D.9}$$

so that we can simplify the expression for q to

$$q=\frac{1-\varepsilon_1\left(1+z\right)}{\left(1-\varepsilon_0-\varepsilon_1\right)\left(1+z\right)}. \tag{D.10}$$

Substituting the expression for q into the expression for mutual information (D.5) then gives the capacity Γ_C of the channel as

$$\Gamma_C=h\left(\frac{1-\varepsilon_1\left(1+z\right)}{1+z}+\varepsilon_1\right)-\frac{1-\varepsilon_1\left(1+z\right)}{\left(1-\varepsilon_0-\varepsilon_1\right)\left(1+z\right)}\left(h\left(\varepsilon_0\right)-h\left(\varepsilon_1\right)\right)-h\left(\varepsilon_1\right). \tag{D.11}$$

We can use the definition of $z\equiv 2^{\frac{(h(\varepsilon_0)-h(\varepsilon_1))}{(1-\varepsilon_0-\varepsilon_1)}}$, rewriting as $\log_2(z)=\frac{(h(\varepsilon_0)-h(\varepsilon_1))}{(1-\varepsilon_0-\varepsilon_1)}$ to write this as

$$\Gamma_C = h\left(\frac{1}{1+z}-\varepsilon_1+\varepsilon_1\right)-\frac{1-\varepsilon_1\left(1+z\right)}{\left(1+z\right)}\log_2(z)-h\left(\varepsilon_1\right) \tag{D.12}$$

$$\Gamma_C = h\left(\frac{1}{1+z}\right)-\frac{\log_2\left(z\right)}{1+z}+\varepsilon_1\log_2\left(z\right)-h\left(\varepsilon_1\right) \tag{D.13}$$

Now, use that

$$h\left(\frac{1}{1+z}\right) = -\frac{1}{1+z}\log_2\left(\frac{1}{1+z}\right)-\frac{z}{1+z}\log_2\left(\frac{z}{1+z}\right) \tag{D.14}$$

$$= \frac{1}{1+z}\log_2(1+z)-\frac{z}{1+z}\log_2(z)+\frac{z}{1+z}\log(1+z) \tag{D.15}$$

$$= \log_2\left(1+z\right)-\frac{z\log_2\left(z\right)}{z+1} \tag{D.16}$$

so that

$$\Gamma_C=\log_2\left(1+z\right)-\log_2\left(z\right)+\varepsilon_1\log_2\left(z\right)-h(\varepsilon_1) \tag{D.17}$$

which by again using the definition of z we can write as

$$\Gamma_C = \log_2(1+z) - \frac{1-\varepsilon_1}{(1-\varepsilon_0-\varepsilon_1)}\left(h(\varepsilon_0) - h(\varepsilon_1)\right) - h(\varepsilon_1) \tag{D.18}$$

which can then be simplified to give the desired expression

$$\Gamma_C = \log_2\left(1 + 2^{\frac{(h(\varepsilon_0)-h(\varepsilon_1))}{(1-\varepsilon_0-\varepsilon_1)}}\right) - \frac{1-\varepsilon_1}{(1-\varepsilon_0-\varepsilon_1)}h(\varepsilon_0) + \frac{\varepsilon_0}{(1-\varepsilon_0-\varepsilon_1)}h(\varepsilon_1). \tag{D.19}$$