Dynamic Information Preference and Communication with Diminishing Sensitivity Over News*

Jetlir Duraj† Kevin He‡

First version: July 1, 2019
This version: December 18, 2023

Abstract

A Bayesian agent experiences gain-loss utility each period over changes in belief about future consumption ("news utility"), with diminishing sensitivity over the magnitude of news. Diminishing sensitivity induces a preference over news skewness: gradual bad news, one-shot good news is worse than one-shot resolution, which is in turn worse than gradual good news, one-shot bad news. So, the agent's preference between gradual information and one-shot resolution can depend on his consumption ranking of different states. In a dynamic cheap-talk framework where a benevolent sender communicates the state over multiple periods, the babbling equilibrium is essentially unique when the receiver is not loss averse. Contrary to the commitment case, more loss-averse receivers may enjoy higher news utility in equilibrium. We characterize the family of gradual good news equilibria when facing such receivers and find the sender conveys progressively larger pieces of good news.

Keywords: diminishing sensitivity, news utility, dynamic information, cheap talk, preference over skewness of information

*This paper was previously titled "Dynamic Information Design with Diminishing Sensitivity Over News." We thank Krishna Dasaratha, David Dillenberger, Ben Enke, Drew Fudenberg, Simone Galperti, Jerry Green, Faruk Gul, David Hagmann, Marina Halac, Johannes Hörner, David Laibson, Shengwu Li, Jonathan Libgober, Elliot Lipnowski, Erik Madsen, Pietro Ortoleva, Matthew Rabin, Gautam Rao, Collin Raymond, Todd Sarver, Joel Sobel, Charlie Sprenger, Tomasz Strzalecki, the MIT information design reading group, our seminar participants, and anonymous referees for insightful comments. Kevin He thanks the California Institute of Technology for hospitality when some of the work on this paper was completed and the University Research Foundation Grant at the University of Pennsylvania for financial support. All remaining errors are our own.

†University of Pittsburgh. Email: jetlirduraj@gmail.com
‡University of Pennsylvania. Email: hesichao@gmail.com

1 Introduction

People are sometimes willing to pay a cost to change how they receive news over time, even when the information does not help them make better decisions. Consider the following scenario:

A student applies for a selective summer program. He knows that accepted applicants will be notified by email some time during the first week of February, while other applicants will not hear back until much later. In other words, each day of no news during the first week of February is bad news. To avoid experiencing multiple instances of disappointment during the week in case he does not hear back for several days, the student sets up an email filter to automatically redirect any email from the summer program into a holding tank, then releases all messages from the holding tank into his inbox at the end of the week.

In this scenario, the student may be willing to exert costly effort to modify his informational environment because he experiences *diminishingly sensitive* psychological reactions to good and bad news. He is elated by good news and disappointed by bad news in every period, and multiple congruent pieces of news carry a greater total emotional impact if they are experienced separately in different periods than if the aggregated lump-sum news arrives in a single period. This kind of psychological consideration also influences how people convey news to others. When CEOs announce earnings forecasts to shareholders and when organization leaders update their teams about recent developments, they are surely mindful of their information's emotional impact (in addition to its possible instrumental value). Finally, the psychological effects of news also play a prominent role in designing entertainment content, where the audience experiences positive and negative reactions over time to news and developments that have no bearing on their personal decision-making.

In this paper, we study the implications of diminishingly sensitive reactions to news for informational preference and dynamic communication. An individual's future consumption depends on an unknown state of the world. In each period, he observes some information about the state and experiences gain-loss utility over the *change* in his belief about said future consumption ("*news utility*"). How does this individual prefer to learn about the state over time? If there is another person who knows the state and who wants to maximize the first individual's expected welfare, how will this informed person communicate her information?

Of course, we are not the first to model news utility (see Kőszegi and Rabin (2009)) or to study psychological considerations in dynamic games (see the survey Battigalli and Dufwenberg (2022), for example). Our main innovation is the focus on the implications of diminishing sensitivity — a classical but surprisingly under-studied assumption. Diminishing sensitivity in reference dependence traces back to Kahneman and Tversky (1979)'s original formulation of prospect theory. Based on Weber's law and experimental findings about human perception, these authors envision a gain-loss utility based on deviations from a reference point, where larger deviations carry smaller marginal effects. But almost all subsequent work on reference-dependent preferences use two-part linear gain-loss utility functions, so their results are driven by loss aversion but not diminishing sensitivity.[1] Four decades since Kahneman and Tversky (1979)'s publication, O'Donoghue and Sprenger (2018)'s review of the ensuing literature summarizes the situation:

> "Most applications of reference-dependent preferences focus entirely on loss aversion, and ignore the possibility of diminishing sensitivity [...] The literature still needs to develop a better sense of when diminishing sensitivity is important."

We show that diminishing sensitivity leads to novel and testable predictions for preference over information. First, when the agent commits to an information structure ex-ante, diminishing sensitivity generates a preference over the direction of news *skewness*. Any information structure where good news arrives all at once but bad news arrives gradually in small pieces — such as waiting for the admission decision in scenario above — is strictly worse than resolving all uncertainty in one period ("one-shot resolution"). On the other hand, any information structure with the opposite skewness — good news arrives gradually but bad news all at once — is strictly better than one-shot resolution, provided loss aversion is weak enough. We relate this result to recent experiments about preference over the skewness of information in Section 3.1.

As Kőszegi and Rabin (2009) point out, the two-part linear news-utility model (without diminishing sensitivity) predicts that people prefer one-shot resolution over any other dynamic information structure. At the same time, some other theories

[1] Kőszegi and Rabin (2009)'s model of news utility allows for diminishing sensitivity and they argue it is a realistic feature. But their results either work with a special case without diminishing sensitivity, or are in a setting where news utility with and without diminishing sensitivity are behaviorally equivalent.

(e.g., Ely, Frankel, and Kamenica (2015)'s suspense and surprise utility) make the opposite prediction that one-shot resolution is the worst possible information structure. By contrast, the skewness preference induced by news utility with diminishing sensitivity implies the same person can make different choices between gradual information and one-shot resolution in different situations — in particular, it depends on his consumption ranking over the states.

Our second main result is that when an informed benevolent sender communicates the state to the receiver through cheap talk, the receiver's diminishing sensitivity leads to credibility problems for the sender. We show that if the receiver has diminishing sensitivity and low enough loss aversion, the lack-of-commitment problem is so severe that every equilibrium is payoff-equivalent to the babbling equilibrium. The reason is that the sender strictly prefers to lie and say the state is good even when it is bad. This temptation is driven by the receiver's diminishing sensitivity: even though the sender is far-sighted and knows false hope creates additional disappointment when the state is revealed, diminishing sensitivity limits the incremental disutility of this extra future disappointment. Diminishing sensitivity thus drives a wedge between the commitment solution and the equilibrium outcome, whereas the two coincide without it. We also show that high enough loss aversion can restore the equilibrium credibility of good-news messages by increasing the future disappointment cost of false hope in the bad state. As a consequence, receivers with higher loss aversion may enjoy higher equilibrium payoffs.

With enough loss aversion, there exist non-babbling equilibria featuring gradual good news. We characterize the entire family of such equilibria and study how quickly the receiver learns the state. For a class of news-utility functions that include a tractable quadratic specification, the sender always conveys progressively larger pieces of good news over time, so the receiver's equilibrium belief grows at an increasing rate in the good state. The idea is that in equilibrium, the sender must be made indifferent between giving false hope and telling the truth in the bad state, and diminishing sensitivity implies that sustaining said indifference requires a greater amount of false hope when the receiver's current belief is more optimistic. This conclusion also puts a uniform bound on the number of periods of informative communication across all time horizons and all equilibria in this family.

The rest of the paper is organized as follows. Section 2 defines the timing of events and introduces a model of news utility with diminishing sensitivity. Section

3 studies how diminishing sensitivity leads to a preference over information structures with different skewness. Section 4 considers an environment where an informed benevolent sender communicates the state to a receiver with news utility, focusing on the credibility problems in the resulting cheap-talk game. Section 5 discusses related literature and contrasts our results with the predictions of other models of preference over non-instrumental information. In particular, it looks at the model's prediction that an agent's choice between gradual information and one-shot resolution depends on his consumption ranking of the states. Section 6 concludes.

2 Model

2.1 Timing of Events

We consider a discrete-time model with periods $0, 1, 2, ..., T$, where $T \geq 2$. There is a binary state space $\Theta = \{A, B\}$. In the final period T, the agent experiences consumption utility $v(\theta)$ in state $\theta \in \Theta$. There is no consumption in other periods, and we assume that $v(A) \neq v(B)$. For our analysis, it is without loss to normalize $v(A) = 1$ and $v(B) = 0$.

The agent starts with a prior probability $0 < \pi_0 < 1$ of the state being A. In every period $t = 1, ..., T$, the agent observes some information and updates his belief about $\{\theta = A\}$ to the Bayesian posterior $0 \leq \pi_t \leq 1$. The information is non-instrumental in that no actions taken in these interim periods affect the state or the consumption utility in period T. In period T, he perfectly learns the true state θ at the moment of consumption, so we always have $\pi_T = 1$ if $\theta = A$ and $\pi_T = 0$ if $\theta = B$.

Given our normalization, π_t is also the agent's time-t expectation of the final-period consumption utility. We refer to information that increases this expectation *good news* and information that decreases it *bad news*.

2.2 News Utility

Although the agent only consumes in the final period, he experiences news utility over consumption in every period. He has a gain-loss utility function, $\mu : [-1, 1] \to \mathbb{R}$ that maps changes in expected final-period consumption utility into a felicity level. At the end of period $1 \leq t \leq T$, the agent experiences news utility $\mu(\pi_t - \pi_{t-1})$ — that is, he derives joy or pain based on the recent belief update from π_{t-1} to π_t. Utility flow is undiscounted and the agent has the same μ in all periods,[2] so his total payoff is $\sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1}) + v(\theta)$. The number of periods T is exogenous, with the length of each period capturing the amount of time that the agent needs to process the new information and to experience its psychological effect.

Throughout we assume μ is continuous, strictly increasing, twice differentiable except possibly at 0, and $\mu(0) = 0$. We make further assumptions on μ to reflect diminishing sensitivity and loss aversion.

Definition 1. Say μ satisfies *diminishing sensitivity* if $\mu^{''}(x) < 0$ and $\mu^{''}(-x) > 0$ for all $x > 0$. Say μ satisfies *(weak) loss aversion* if $-\mu(-x) \geq \mu(x)$ for all $x > 0$. There is *strict loss aversion* if $-\mu(-x) > \mu(x)$ for all $x > 0$.

For instance, the gain-loss function μ in Tversky and Kahneman (1992) where $\mu(x) = x^\alpha$ for $x \geq 0$, $\mu(x) = -\lambda|x|^\beta$ for $x < 0$ with $0 < \alpha, \beta < 1$, $\beta \leq \alpha$, and $\lambda > 1$ satisfies both diminishing sensitivity and strict loss aversion.

This model of diminishing sensitivity over the magnitude of news shares the same psychological motivation as Kahneman and Tversky (1979), who base their theory of human responses to monetary gains and losses on Weber's law and on psychology experiments about how people perceive changes in physical attributes like temperature or brightness. In the realm of news, we make the analogous assumption that the magnitude of news utility is strictly concave in the magnitude of belief update, both in the direction of good news and in the direction of bad news. This translates into the assumption that μ must be strictly concave in the positive domain (i.e., good news) and strictly convex in the negative domain (i.e., bad news). This pair of assumptions about the curvature of μ drives the results.[3]

This framework of deriving utility from changes in beliefs has been previously discussed in Kőszegi and Rabin (2009), but they mostly focus on another model that makes percentile-by-percentile comparisons between old and new beliefs and without diminishing sensitivity.[4] The model we use allows us to characterize the implications of diminishing sensitivity in the simplest setup with two states.

[2] Our preference satisfies Segal (1990)'s time neutrality axiom. We abstract away from preferences for early or late resolution of uncertainty.

[3] If we instead assume that $v(A) = 0$ and $v(B) = 1$, then the agent experiences higher news utility when he updates his posterior more in the direction of state B. Diminishing sensitivity would still require that the magnitude of the agent's news utility is strictly concave in the magnitude of belief update in either direction.

[4] In their model, suppose F and G are the distributions over future consumption utility given by the agent's old and new beliefs. If $F^{-1}(q)$ and $G^{-1}(q)$ correspond to the q percentiles of these

2.2.1 Quadratic News Utility

We discuss another tractable functional form of μ that is rich enough to exhibit both diminishing sensitivity and loss aversion. The quadratic news-utility function $\mu : [-1, 1] \to \mathbb{R}$ is given by

$$\mu(x) = \begin{cases} \alpha_p x - \beta_p x^2 & x \geq 0 \\ \alpha_n x + \beta_n x^2 & x < 0 \end{cases}$$

with $\alpha_p, \beta_p, \alpha_n, \beta_n > 0$. So we have

$$\mu^{'}(x) = \begin{cases} \alpha_p - 2\beta_p x & x > 0 \\ \alpha_n + 2\beta_n x & x < 0 \end{cases}, \qquad \mu^{''}(x) = \begin{cases} -2\beta_p & x > 0 \\ 2\beta_n & x < 0 \end{cases}.$$

The parameters α_p, α_n control the extent of loss aversion near 0, while β_p, β_n determine the amount of curvature — i.e., the second derivative of μ. The maintained general assumptions on μ imply the following parametric restrictions.

1. *Monotonicity*: $\alpha_p \geq 2\beta_p$ and $\alpha_n \geq 2\beta_n$. These inequalities hold if and only if μ is strictly increasing.

2. *Loss aversion*: $\alpha_n - \alpha_p \geq \max(0, \beta_n - \beta_p)$. This condition is equivalent to loss aversion from Definition 1 for this class of news-utility functions.

A family of quadratic news-utility functions that satisfy these two restrictions can be constructed by choosing any $\alpha > 2\beta > 0$ and $\lambda \geq 1$, then set $\alpha_p = \alpha$, $\alpha_n = \lambda\alpha$, $\beta_p = \beta$, $\beta_n = \lambda\beta$. Figure 1 plots some of these news-utility functions for different values of α, β, and λ.

3 Diminishing Sensitivity and Preference over News Skewness

In this section, we show that news utility with diminishing sensitivity makes novel predictions about preference over the skewness of information. When there is no loss

utility distributions for $0 \leq q \leq 1$, then the agent experiences news utility $\int_0^1 \mu(G^{-1}(q) - F^{-1}(q))dq$. Kőszegi and Rabin (2009) focus on the case where μ exhibits loss aversion but not diminishing sensitivity.

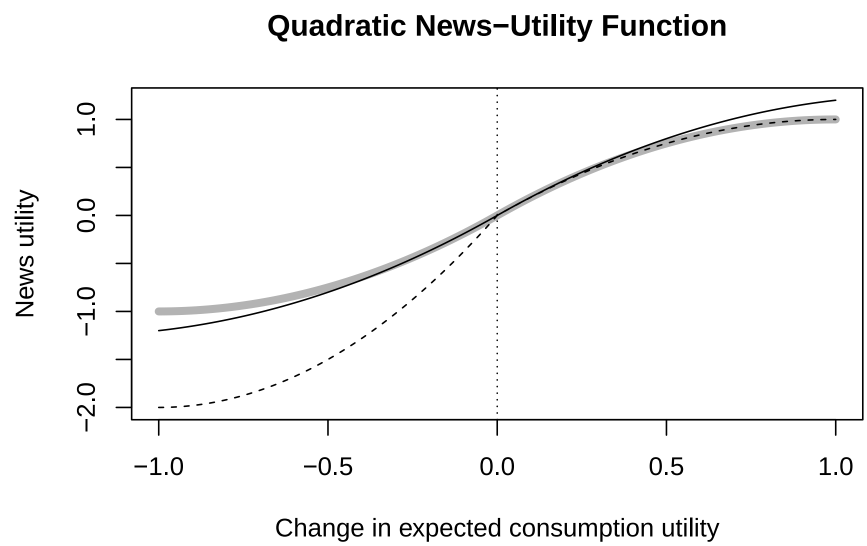

Figure 1: Examples of quadratic news-utility functions in the family $\alpha_p = \alpha$, $\alpha_n = \lambda\alpha$, $\beta_p = \beta$, $\beta_n = \lambda\beta$. Shaded curve: $\alpha = 2$, $\beta = 1$, $\lambda = 1$. Dashed curve: $\alpha = 2$, $\beta = 1$, $\lambda = 2$. Solid curve: $\alpha = 2$, $\beta = 0.8$, $\lambda = 1$.

aversion, one-shot resolution of uncertainty is neither the agent's most preferred way to get information nor the least preferred. Instead, the agent strictly prefers one-shot resolution over an information structure that delivers piecemeal bad news over time, and strictly prefers an information structure with the opposite skewness over one-shot resolution. By continuity, the same conclusions hold when loss aversion is present but sufficiently weak.

Definition 2. An information structure features *gradual good news, one-shot bad news* if

- $\mathbb{P}[\pi_t \geq \pi_{t-1}$ for all $1 \leq t \leq T \mid \theta = A] = 1$ and
- $\mathbb{P}[\pi_t < \pi_{t-1}$ for no more than one $1 \leq t \leq T \mid \theta = B] = 1$.

An information structure features *gradual bad news, one-shot good news* if

- $\mathbb{P}[\pi_t \leq \pi_{t-1}$ for all $1 \leq t \leq T \mid \theta = B] = 1$ and
- $\mathbb{P}[\pi_t > \pi_{t-1}$ for no more than one $1 \leq t \leq T \mid \theta = A] = 1$.

In the gradual good news, one-shot bad news information structures, the agent gets good news over time and gradually increases his expectation of future consumption. When the state is bad, the agent gets all the negative information at once — in the first period when his expectation π_t strictly decreases, he fully learns that the state is bad. Conversely, "gradual bad news, one-shot good news" refers to the opposite kind of information structure.

An information structure features *one-shot resolution* if

$$\mathbb{P}[\pi_t \neq \pi_{t-1} \text{ for at most one } 1 \leq t \leq T] = 1.$$

That is, almost surely the agent's belief only changes in one period (possibly the final period when true state is perfectly revealed). Note that one-shot resolution falls into both classes from Definition 2. We say that an information structure features *strictly gradual good news* if

$$\mathbb{P}[\pi_t > \pi_{t-1} \text{ and } \pi_{t'} > \pi_{t'-1} \text{ for two distinct } 1 \leq t, t^{'} \leq T \mid \theta = A] > 0.$$

That is, there is positive probability that the agent's expectation strictly increases at least twice in periods 1 through T. Similarly define *strictly gradual bad news.*

We now prove that whenever μ satisfies diminishing sensitivity and (weak) loss aversion, information structures featuring strictly gradual bad news, one-shot good news are *strictly* worse than one-shot resolution. The intuition is that an information structure in this class delivers small pieces of bad news but large clumps of good news, which is the exact opposite of what the agent wants when he experiences diminishing sensitivity to news. When $\theta = B$, the information structure gives several pieces of bad news, but the agent is better off getting all of the bad news in one period. When $\theta = A$, the information structure gives several pieces of bad news followed by conclusive good news. By diminishing sensitivity, this is worse than getting all of the bad news in one period and then the conclusive good news in the subsequent period. By loss aversion, this is in turn worse than directly learning the state in the first period.

Proposition 1. *Suppose μ satisfies diminishing sensitivity and weak loss aversion. Any information structure featuring strictly gradual bad news, one-shot good news provides strictly lower utility than one-shot resolution in expectation, and almost surely* *weakly lower utility ex-post.*

Proposition 1 identifies a class of information structures that are worse than one-shot resolution for news utility with diminishing sensitivity, distinguishing it from other models of information preference where one-shot resolution is the worst possible information structure. Utility models that make this other prediction include suspense and surprise (Ely, Frankel, and Kamenica, 2015) and news utility with a two-part linear, gain-loving (instead of loss-averse) value function (Chapman, Snowberg, Wang, and Camerer, 2022; Campos-Mercade, Goette, Graeber, Kellogg, and Sprenger, 2022).

Next, we show that if the agent has diminishing sensitivity but not loss aversion, then information structures with strictly gradual good news, one-shot bad news are *strictly* better than one-shot resolution.

Proposition 2. *Suppose μ satisfies diminishing sensitivity and it is symmetric around 0 with $-\mu(-x) = \mu(x)$ for all $x \geq 0$ (that is, it does not exhibit loss aversion). Any information structure featuring strictly gradual good news, one-shot bad news provides strictly higher utility than one-shot resolution in expectation, and almost surely weakly higher utility ex-post.*

In Kőszegi and Rabin (2009)'s model of news utility without diminishing sensitivity, one-shot resolution is optimal among all information structures.[5] By contrast, Proposition 2 can be combined with continuity to show that for news-utility functions with diminishing sensitivity and a small enough amount of loss aversion, there are information structures that are strictly better than one-shot resolution. To make this precise, consider the parametric class of λ-*scaled news-utility functions.* We fix some $\tilde{\mu}_{pos} : [0,1] \to \mathbb{R}_+$, strictly increasing and strictly concave with $\tilde{\mu}_{pos}(0) = 0$, and consider the family of news-utility functions given by $\mu_\lambda(x) = \tilde{\mu}_{pos}(x)$, $\mu_\lambda(-x) = -\lambda\tilde{\mu}_{pos}(x)$ for $x \geq 0$ as we vary the loss aversion parameter $\lambda \geq 1$.

Corollary 1. *Consider a class of λ-scaled news-utility functions $(\mu_\lambda)_{\lambda \geq 1}$ and any information structure featuring strictly gradual good news, one-shot bad news. There exists some $\bar{\lambda} > 1$ so that for any $1 \leq \lambda \leq \bar{\lambda}$, this information structure gives strictly higher utility than one-shot resolution in expectation.*

[5] Kőszegi and Rabin (2009) show this for their percentile-based model of news utility with binary states, while Dillenberger and Raymond (2020) prove the same also holds for arbitrarily many states.

In summary, provided loss aversion is low enough, diminishing sensitivity induces the following preference ranking: gradual good news, one-shot bad news is better than one-shot resolution, which is in turn better than gradual bad news, one-shot good news.[6]

3.1 Experiments on Information Preference

Our results relate to a number of experimental papers that test whether people prefer one-shot resolution by asking subjects to choose how they wish to learn about their prize for the experiment, with one-shot resolution as a feasible information structure. After accounting for preference over the timing of resolution,[7] Falk and Zimmermann (2023) and Bellemare, Krause, Kröger, and Zhang (2005) find that subjects prefer one-shot resolution, while Nielsen (2020); Masatlioglu, Orhun, and Raymond (2017); Zimmermann (2014); Budescu and Fischer (2001) find the opposite. News utility with diminishing sensitivity may explain these mixed results, as it predicts one-shot resolution is neither the best nor the worst information structure, so it may or may not be chosen depending on what other information structures are feasible in a particular experiment. On the other hand, these experimental results are harder to reconcile with theories that either predict people always choose one-shot resolution or predict people always avoid it.

Gul, Natenzon, Ozbay, and Pesendorfer (2020) find in an experiment that 59% of the subjects choose gradual information over early one-shot resolution when the gradual information features gradual good news, one-shot bad news. But only 40% of the subjects make the same choice when the gradual information features gradual bad news, one-shot good news instead. These findings are consistent with the mechanism discussed in this section.

Two experiments have examined people's preference over the skewness of news, with mixed results. Tables 10 and 11 in Nielsen (2020) report that subjects prefer negatively skewed news, as predicted by news utility with diminishing sensitivity. But, Masatlioglu, Orhun, and Raymond (2017) find that agents prefer positively

[6] Appendix B of the working paper version of this article (available at https://arxiv.org/abs/1908.00084v5) contains additional results about preference over information structures.

[7] Information structures that reveal the prize gradually will resolve uncertainty earlier than a one-shot resolution structure that reveals the prize at the end of the experiment, but later than a one-shot resolution structure that reveals the prize immediately.

skewed news. In showing that a classical assumption of reference dependence leads to a prediction about preference over news skewness, we hope to stimulate further empirical work on this topic.

4 Diminishing Sensitivity and the Credibility Problem

So far, we have assumed the agent commits to an information structure ex-ante. In many economic settings, it is instead an informed individual who communicates the state to the agent over time. Such communication often takes the form of unverifiable cheap-talk messages, especially if the speaker wishes to convey inconclusive news about the state.

We consider a cheap-talk game between a receiver who experiences news utility with diminishing sensitivity, and a benevolent sender who knows the state and wishes to maximize the receiver's welfare. At first glance, one may think that the sender can simply implement the receiver's favorite information structure in the equilibrium of the game, given that the two parties have aligned incentives. While this is true with two-part linear news utility, we show that the receiver's diminishing sensitivity leads to a credibility problem for the sender.

4.1 Cheap Talk with an Informed and Benevolent Sender

Let a finite set of cheap-talk messages M with $|M| \geq 2$ be fixed. The sender learns the true state of the world $\theta \in \{A, B\}$ in period $t = 0$. In every period $t = 1, 2, ..., T-1$, the sender conveys a message $m \in M$ to the receiver. The sender's communication strategy in period t is given by a mixture over messages $\sigma_t(\cdot \mid h^{t-1}, \theta) \in \Delta(M)$ that can depend on the history h^{t-1} of messages so far and the true state θ. The sender cannot commit to how she will communicate with the receiver in different states of the world.

The sender is benevolent and wants to maximize the receiver's welfare. At the end of period T, if the receiver has experienced the belief path $(\pi_t)_{t=0}^{T}$, then the sender's total payoff in the game is $\sum_{t=1}^{T} \mu\,(\pi_t - \pi_{t-1})$ (we may ignore the physical consumption term since neither party can affect it). The state of the world determines the final belief π_T and thus affects news utility in the final period, so the sender expects different payoffs from sending the same sequence of messages in different states.[8]

We analyze perfect-Bayesian equilibria of the cheap talk game, under some off-path belief refinements.

Definition 3. A *perfect-Bayesian equilibrium* consists of sender's strategy $\sigma^* = (\sigma_t^*)_{t=1}^{T-1}$ together with receiver's beliefs $p^* : \cup_{t=0}^{T-1} H^t \to [0,1]$, where:

- For every $1 \leq t \leq T-1$, $h^{t-1} \in H^{t-1}$ and $\theta \in \{A, B\}$, σ^* maximizes the receiver's total expected news utility in periods $t, ..., T-1, T$ conditional on having reached the public history h^{t-1} in state θ at the start of period t.
- p^* is derived by applying the Bayes' rule to σ^* whenever possible.

We make two belief-refinement restrictions:

- If $t \leq T-1$, h^t is a continuation history of $h^{\underline{t}}$, and $p^*(h^{\underline{t}}) \in \{0,1\}$, then $p^*(h^t) = p^*(h^{\underline{t}})$.
- The receiver's belief π_T in period T when state is θ puts probability 1 on θ, regardless of the preceding history $h^{T-1} \in H^{T-1}$.

We will abbreviate a perfect-Bayesian equilibrium satisfying our off-path belief refinements as an "equilibrium." Our definition requires that once the receiver updates his belief to 0 or 1, this belief stays constant through the end of period $T-1$. In other words, the support of his belief is non-expanding through the penultimate period.[9] In period T, the receiver updates his belief to reflect full confidence in the true state of the world, regardless of his (possibly dogmatically wrong) belief at the end of period $T-1$.

Babbling equilibria always exist for any news-utility function μ, message space M, time horizon T, and prior π_0. In a babbling equilibrium, the sender mixes over messages in a state-independent way, and the receiver keeps his prior belief π_0 after every history up until period T. A babbling equilibrium implements one-shot resolution for the receiver, as his belief stays constant until the final period and then fully resolves.

[8]In particular, this is not a cheap-talk game with state-independent sender payoffs, as in Lipnowski and Ravid (2020).

[9]This standard refinement was first used in Grossman and Perry (1986). It rules out pathological off-path belief updates if the sender deviates and sends a message perfectly indicative of one state following a history where the receiver is fully convinced of the other state.

4.2 The Credibility Problem and Babbling

Are there equilibria where the sender gets a higher expected payoff than the *babbling payoff* of $\pi_0\mu(1-\pi_0)+(1-\pi_0)\mu(-\pi_0)$? By Proposition 2, for a receiver who has diminishing sensitivity but not loss aversion, there exists a class of information structures that is strictly better than one-shot resolution. But the next result proves none of these information structures can be implemented in equilibrium.

Proposition 3. *Suppose μ is symmetric around 0 and $\mu''(x)<0$ for all $x>0$. For any M,T,π_0, the sender's payoff in every equilibrium is equal to the babbling payoff.*

To understand why, consider the final period of communication $t=T-1$. Suppose the state is bad and the sender must decide between revealing the truth to decrease the receiver's belief from π to 0, or sending a positive message that increases the receiver's belief by $z>0$. Such false hope in period $T-1$ gives positive news utility today at the cost of increasing disappointment in the final period. But diminishing sensitivity implies the marginal utility of positive news today is larger than the marginal disutility of the incremental future disappointment, $\mu(z)>\mu(-\pi)-\mu(-(\pi+z))$. This shows that in equilibrium, the sender cannot communicate good news in either state, otherwise she will be tempted to mimic the good-news messages when the state is bad, destroying the credibility of these messages. So the sender must babble in period $T-1$, so we could treat period $T-2$ as the last period of communication, and apply the same arguments by backward induction.

In summary, diminishing sensitivity leads to a credibility problem that prevents any informative communication, even though the players share the same payoff function. In a cheap-talk setting with instrumental information and anticipatory utility, Kőszegi (2006) shows that a benevolent sender also distorts equilibrium communication relative to the commitment benchmark. The breakdown in communication is more complete in our setting, for the players get the same payoffs as when communication is impossible.

The intuition we gave for the uniqueness of babbling up to payoffs assumes the receiver is not loss averse — that is, μ is symmetric around 0. Babbling remains unique with a small amount of loss aversion, but a high enough level of loss aversion can restore the sender's credibility and enable non-babbling equilibria. (In the next section, we will construct a family of such non-babbling equilibria.)

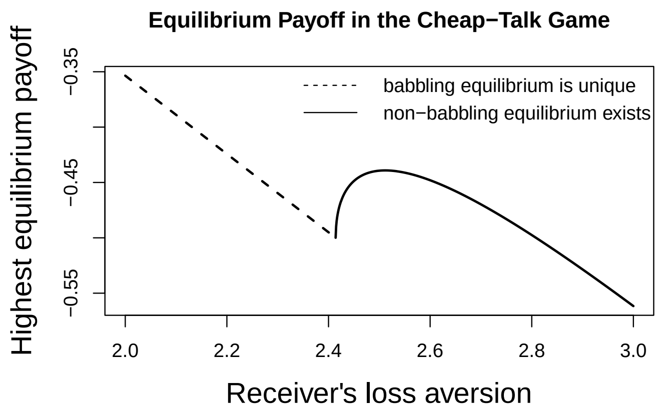

Figure 2: The babbling equilibrium is unique up to payoffs for low values of λ, but there exists an equilibrium with gradual good news for $\lambda \geq 1+\sqrt{2} \approx 2.414$. Due to the role of loss aversion in sustaining credible partial news, a receiver with higher loss aversion may experience higher expected news utility in equilibrium than a receiver with lower loss aversion.

Example 1. To illustrate, consider the following case: $\mu(x) = \sqrt{x}$ for $x \geq 0$, $\mu(x) = -\lambda\sqrt{-x}$ for $x < 0$, $T = 2$, and $\pi_0 = \frac{1}{2}$. The highest equilibrium payoff for different values of λ is depicted in Figure 2. (Detailed calculations supporting these results are available in Appendix A.5).

Receivers with higher λ may have higher equilibrium payoffs. This non-monotonicity is driven by the fact that, when $\lambda \leq 1+\sqrt{2}$, the babbling equilibrium is unique up to payoffs and increasing λ decreases expected news utility linearly. A new, non-babbling equilibrium emerges when λ exceeds $1+\sqrt{2}$. In this equilibrium, the sender induces the belief $\frac{1}{2} \cdot \left(\frac{\lambda^2+1}{\lambda^2-1}\right)^2$ in period $t=1$ if $\theta = A$, and induces either the belief $\frac{1}{2} \cdot \left(\frac{\lambda^2+1}{\lambda^2-1}\right)^2$ or the belief 0 in period $t=1$ if $\theta = B$. Increasing loss aversion then carries two countervailing effects: first, a *non-strategic effect* of hurting welfare when $\theta = B$, as it increases the disutility when the receiver eventually hears the bad news; second, an *equilibrium effect* of changing the relative amounts of good news in different periods conditional on $\theta = A$. As λ increases above $1+\sqrt{2}$, the receiver goes from getting all of the good news in the final period when $\theta = A$ to getting some partial good news in the first period when $\theta = A$. In other words, increasing λ helps by improving the equilibrium "consumption smoothing" of good news across two periods.

4.3 Deterministic Gradual Good News Equilibria

When the receiver's loss aversion is high enough, there can exist non-babbling equilibria in the cheap-talk game. We now analyze a family of such non-babbling equilibria, where the receiver's belief monotonically increases over time conditional on the good state. These equilibria show that the gradual good news, one-shot bad news information structures discussed in Section 3 can be sustained without commitment.

An equilibrium (M, σ^*, p^*) features *deterministic*[10]*gradual good news* (GGN equilibrium) if there exists a sequence of constants $p_0 \leq p_1 \leq ... \leq p_{T-1} \leq p_T$ with $p_0 = \pi_0$, $p_T = 1$, and the receiver always has belief p_t in period t when $\theta = A$. By Bayesian beliefs, in any GGN equilibrium the sender must induce a belief of either 0 or p_t in period t when $\theta = B$, as any message not inducing belief p_t is a conclusive signal of the bad state.

The class of GGN equilibria is non-empty, for it contains the babbling equilibrium where $\pi_0 = p_0 = p_1 = ... = p_{T-1} < p_T = 1$. The number of *intermediate beliefs* in a GGN equilibrium is the number of distinct beliefs in the open interval $(\pi_0, 1)$ along the sequence $p_0, p_1, ..., p_{T-1}$. The babbling equilibrium has zero intermediate beliefs.

The next proposition characterizes the set of all GGN equilibria with at least one intermediate belief.

Proposition 4. *Let $P^*(\pi) \subseteq (\pi, 1]$ be those beliefs $x > \pi$ satisfying $\mu(x - \pi) + \mu(-x) = \mu(-\pi)$. Suppose μ exhibits diminishing sensitivity and loss aversion. For $1 \leq J \leq T - 1$, there exists a gradual good news equilibrium with the J intermediate beliefs $q^{(1)} < ... < q^{(J)}$ if and only if $q^{(j)} \in P^*(q^{(j-1)})$ for every $j = 1, ..., J$, where $q^{(0)} := \pi_0$.*

To interpret, $P^*(\pi)$ contains the set of beliefs $x > \pi$ such that the sender is indifferent between inducing the two belief paths $\pi \to x \to 0$ and $\pi \to 0$. When μ is symmetric, this indifference condition is never satisfied, which is the source of the credibility problem for good-news messages. The same indifference condition pins down the relationship between successive intermediate beliefs in GGN equilibria. This condition ensures that in the bad state, the sender is willing to randomize between revealing the state and lying with an inconclusive piece of good news that moves the receiver to the next intermediate belief.

[10] This class of equilibria is slightly more restrictive than the gradual good news, one-shot bad news information structures from Definition 2, because the sender may not randomize between several increasing paths of beliefs in the good state.

We illustrate this result with the quadratic news utility.

Corollary 2. *1) With quadratic news utility,* $P^*(\pi) = \left\{\pi \cdot \frac{\beta_p+\beta_n}{\beta_p-\beta_n} - \frac{\alpha_n-\alpha_p}{\beta_p-\beta_n}\right\} \cap (\pi, 1)$.

2a) If $\beta_n > \beta_p$*, there cannot exist any gradual good news equilibrium with more than one intermediate belief.*

2b) If $\beta_n < \beta_p$*, there can exist gradual good news equilibria with more than one intermediate belief. For a given set of parameters of the quadratic news-utility function and prior* π_0*, there exists a uniform bound on the number of intermediate beliefs that can be sustained in equilibrium across all* T*.*

3) In any GGN equilibrium with quadratic news utility, intermediate beliefs in the good state grow at an increasing rate.

For the case of quadratic news utility, this result provides a closed-form characterization of the successive intermediate beliefs. It also shows every GGN equilibrium involves progressively larger pieces of good news in the good state, $q^{(j+1)} - q^{(j)} > q^{(j)} - q^{(j-1)}$. The convex time-path of equilibrium beliefs is due to diminishing sensitivity. If the sender is indifferent between providing d amount of false hope and truth-telling in the bad state when the receiver has prior belief π_L, then she strictly prefers providing the same amount of false hope over truth-telling at any more optimistic prior belief $\pi_H > \pi_L$. The false hope generates the same positive news utility in both cases, but an extra d units of disappointment matters less when added a baseline disappointment level of π_H rather than π_L, thanks to diminishing sensitivity.

Equilibrium beliefs in the good state grow at an increasing rate, but must be bounded above by 1. So, there exists some uniform bound $\bar{J}$ on the number of intermediate beliefs depending only on the prior belief π_0 and parameters of the news-utility function.

As an illustration, consider the quadratic news utility with $\alpha_p = 2$, $\alpha_n = 2.1$, $\beta_p = 1$, and $\beta_n = 0.2$. Starting at the prior belief of $\pi_0 = \frac{1}{3}$, Figure 3 shows the longest possible sequence of intermediate beliefs in any GGN equilibrium for arbitrarily large T. Since the P^* sets are either empty sets or singleton sets for the quadratic news utility, Figure 3 also contains all the possible beliefs in any state of any GGN equilibrium with these parameters.

Beyond the quadratic case, the intuition that diminishing sensitivity should cause the receiver to have a convex time-path of equilibrium beliefs holds more generally. The next result formalizes this relationship. It shows that when diminishing

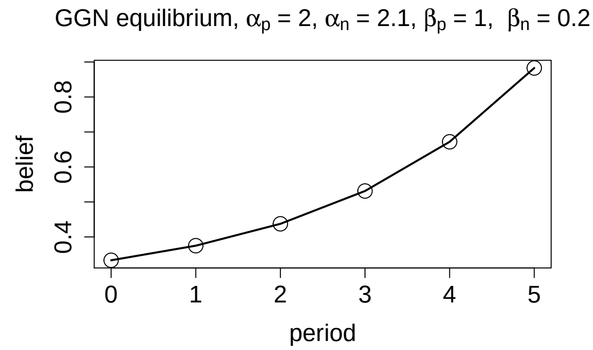

Figure 3: The longest possible sequence of GGN intermediate beliefs starting with the prior $\pi_0 = \frac{1}{3}$. For quadratic news utility, equilibrium GGN beliefs always increase at an increasing rate in the good state.

sensitivity is combined with a pair of sufficient regularity conditions, intermediate beliefs grow at an increasing rate in any GGN equilibrium. These conditions are satisfied, for example, by the square-roots news utility with loss aversion.

Proposition 5. *Suppose μ exhibits diminishing sensitivity, $|P^*(\pi)| \leq 1$ for all $0 < \pi < 1$, and $\mu'(0_-) \leq \mu'(0_+)$. Then, in any GGN equilibrium with intermediate beliefs $q^{(1)} < ... < q^{(J)}$, we get $q^{(j)} - q^{(j-1)} < q^{(j+1)} - q^{(j)}$ for all $1 \leq j \leq J-1$.*

The first regularity condition requires that the sender is indifferent between the belief paths $\pi \to x \to 0$ and $\pi \to 0$ for at most one $x > \pi$. It is a technical assumption that lets us prove our result, but we suspect the conclusion also holds under some relaxed conditions. The second regularity condition implies that in the bad state, the total news utility associated with an ϵ amount of false hope is higher than truth-telling for small $\epsilon > 0$. It is satisfied if $\mu'(0_+) = \infty$ or if μ is differentiable at 0.

5 Related Literature and Predictions of Other Belief-Based Utility Models

5.1 Predictions of Other Belief-Based Utility Models

In general, papers on belief-based utility have highlighted two sources of felicity: *levels* of belief about future consumption utility ("anticipatory utility," e.g., Kőszegi (2006); Eliaz and Spiegler (2006); Schweizer and Szech (2018)) and *changes* in belief about future consumption utility ("news utility" and "suspense and surprise" (Ely, Frankel, and Kamenica, 2015)). For the latter, some function of both the prior belief and the posterior belief serves as the carrier of utility. For the former, a given posterior belief brings the same anticipatory utility for all priors (Eliaz and Spiegler, 2006). The rich information preference under news utility with diminishing sensitivity contrast against more stark predictions of these other commonly used models.

5.1.1 News Utility without Diminishing Sensitivity

The literature on reference-dependent preferences and news utility has focused on two-part linear gain-loss utility functions, which violate diminishing sensitivity. If μ is two-part linear with loss aversion, then it follows from the martingale property of Bayesian beliefs that one-shot resolution is weakly optimal for the agent among all information structures. If there is strict loss aversion, then one-shot resolution does strictly better than any information structure that resolves uncertainty gradually.

5.1.2 Anticipatory Utility

Eliaz and Spiegler (2006) study a representation of anticipatory utility. The decision-maker has a utility u defined over posterior beliefs, and the ex-ante anticipatory utility of an information structure is the expectation of u evaluated at the Bayesian posterior beliefs. Kőszegi (2006) considers a cheap-talk setting where the receiver gets a message from the sender, updates his belief about the state, then takes an action. He experiences anticipatory utility proportional to the expectation of his future utility, based on his action choice and his belief about the state.

In our setup where the agent does not take any actions, the analogous model of the agent's anticipatory utility in period t is π_t, given our normalization $v(A) = 1$ and $v(B) = 0$. Our results would be unchanged if we let the agent experience both anticipatory utility and news utility. This is because by the martingale property, the agent's ex-ante expected anticipatory utility in a given period is the same across all information structures. So, the ranking of information structures entirely depends on the news utility they generate.

More generally, one could model anticipatory utility in period t as $W(\pi_t)$, where $W : \mathbb{R} \to \mathbb{R}$ is a strictly increasing anticipatory-utility function. If the agent only experiences anticipatory utility, not news utility, then there exists an optimal information structure that only releases information in $t = 1$ (see Appendix B.1). By contrast, this kind of information structure is in general not optimal when the agent has diminishing sensitivity and weak enough loss aversion.

Another difference is that under anticipatory utility, the agent prefers any setting where state A has any positive probability $\pi_0 > 0$ to a setting where $\pi_0 = 0$. By contrast, an agent with news utility may prefer the latter (see Appendix B.2). Even though the former always gives the agent weakly better outcomes and weakly higher expectation of future consumption utility, it does not always give better news. For agents who derive utility from *changes* in beliefs, the incentive to avoid disappointing news could make them choose to lower π_0 to 0 (Dreyfuss, Heffetz, and Rabin, 2022).

5.1.3 Suspense and Surprise

Ely, Frankel, and Kamenica (2015) propose and study an original utility function over belief paths where larger belief movements always bring greater felicity. By contrast, changes in beliefs may increase or decrease the receiver's utility in our setting. Information structures featuring gradual bad news, one-shot good news are worse than one-shot resolution in our problem, while one-shot resolution is the worst possible information structure in Ely, Frankel, and Kamenica (2015)'s problem. Ely, Frankel, and Kamenica (2015) also discuss state-dependent versions of suspense and surprise utilities, but this extension does not embed our model (see Appendix B.3).

5.2 Related Work in Decision Theory

Several paper in decision theory have studied models of preference over dynamic information structures. Dillenberger (2010) shows that preference for one-shot resolution of uncertainty is equivalent to a weakened version of independence, provided the preference satisfies recursivity. This result does not apply here because our model of news utility violates recursivity (see Appendix B.2). Dillenberger and Raymond (2020) axiomatize a general class of additive belief-based preferences in the domain of two-stage lotteries, relaxing recursivity and the independence axiom. In the case of $T = 2$, our news-utility model belongs to the class they characterize. Under this specialization, our work may be thought of as studying information preference and strategic communication using some of Dillenberger and Raymond (2020)'s additive belief-based preferences. Gul, Natenzon, and Pesendorfer (2021) axiomatize a class of preferences over non-instrumental information that they call risk consumption preferences. In contrast, we study the implications diminishing sensitivity in a model that is not a risk consumption preference (see Appendix B.4).

5.3 Consumption Preference and Information Preference

To further illustrate how diminishing sensitivity leads to different predictions than the models in Section 5.1, we study an extension of the baseline model from Section 2 by considering a heterogeneous population of agents who differ in their consumption preferences over the two states.

There is a binary state space $\Theta = \{A, B\}$, but some agents prefer A to B (so $v(A) = 1$, $v(B) = 0$ as in Section 2) while other agents prefer B to A (so $v(A) = 0$, $v(B) = 1$). All agents use the same gain-loss utility function μ to map changes in expected future consumption utility into a felicity level. Using π_t to denote Bayesian posterior belief about $\{\theta = A\}$ in period t, total news utility is $\sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1})$ for someone with $v(A) = 1, v(B) = 0$, and it is $\sum_{t=1}^{T} \mu(-(\pi_t - \pi_{t-1}))$ for someone with $v(A) = 0, v(B) = 1$.

Consider a setting where a sequence of signal realizations gradually determine the binary state. In each period $t = 1, 2, ..., T$, a binary signal X_t realizes, where $\mathbb{P}[X_t = 1] = q_t$ with $0 < q_t < 1$. Each X_t is independent of the other ones. If $X_t = 1$ for all t, then the state is A. Otherwise, when $X_t = 0$ for at least one t, the state is B.[11] At time 0, each agent chooses between observing the realizations of the signals $(X_t)_{t=1}^{T}$ in real time (*gradual information*), or only learning the state of the world at the end of period T (*one-shot resolution*).

For a concrete example, imagine a televised debate between two political candidates A and B where A loses as soon as she makes a "gaffe" during the debate. If A does not make any gaffes, then A wins. In this example, $\{X_t = 1\}$ corresponds to the event that candidate A does not make a gaffe during the t-th minute of the debate. States A and B correspond to candidates A and B winning the debate.

For someone who prefers candidate A, the debate provides gradual good news, one-shot bad news. For someone who prefers candidate B, the debate provides gradual

[11] Equivalently, we can think of state A having probability $\Pi_{t=1}^{T} q_t$ and state B having the complementary probability. Conditional on $\theta = A$, we always have $X_1 = ... = X_T = 1$. Conditional on $\theta = B$, for a sequence of signal realizations $(x_1, x_2, ..., x_T) \in \{0, 1\}^T$, we have $\mathbb{P}[(X_1, ..., X_T) = (x_1, ..., x_T) \mid \theta = B] = [\Pi_{t=1}^{T} q_t^{x_t} \cdot (1 - q_t)^{1-x_t}]/[1 - \Pi_{t=1}^{T} q_t]$ if at least one x_i is 0, otherwise $\mathbb{P}[(X_1, ..., X_T) = (x_1, ..., x_T) \mid \theta = B] = 0$.

bad news, one-shot good news. Proposition 1 and Corollary 1 imply that these two types of agents can make different choices about whether to watch the debate. So, heterogeneous consumption preferences can generate heterogeneous information preferences. By contrast, the related theories reviewed in Section 5.1 predict that the agent either always prefers one-shot resolution in all situations, or always prefers every other information structure to one-shot resolution in all situations.

Proposition 6. *The following models predict that the agent will not change his choice between gradual information and one-shot resolution when the sign of $v(A) - v(B)$ changes.*

1. *News utility with a two-part linear μ, where $\mu(x) = x$ for $x \geq 0$ and $\mu(x) = \lambda x$ for $x < 0$, with any $\lambda \geq 0$.*
2. *Anticipatory utility where the agent gets $W(\pi_t \cdot v(A) + (1 - \pi_t) \cdot v(B))$ in period t, with W an increasing, weakly concave function.*
3. *Ely, Frankel, and Kamenica (2015)'s "suspense and surprise" utility.*

5.4 Related Work on News Utility and Information Design

Since Kőszegi and Rabin (2009), several other authors have analyzed the implications of news utility in different settings (Pagel, 2016, 2017, 2018; Duraj, 2019). These papers focus on Bayesian agents with two-part linear gain-loss utilities and do not study the role of diminishing sensitivity to news.

Interpreting monetary gains and losses as news about future consumption, experiments that show risk-seeking behavior when choosing between loss lotteries and risk-averse behavior when choosing between gain lotteries provide evidence for diminishing sensitivity over consumption news (see e.g., Rabin and Weizsäcker (2009)). In the same vein, papers in the finance literature that use diminishing sensitivity over monetary gains and losses to explain the disposition effect (Shefrin and Statman, 1985; Kyle, Ou-Yang, and Xiong, 2006; Barberis and Xiong, 2012; Henderson, 2012) also provide indirect evidence for diminishing sensitivity over consumption news.

Bowman, Minehart, and Rabin (1999) study a consumption-based reference-dependent model with diminishing sensitivity. A critical difference is that their reference points are based on past habits, not rational expectations.

Ebert and Strack (2015, 2018) study dynamic gambling for agents with cumulative prospective theory preferences. The gambler's wealth process forms a martingale if the bets are fair, a property shared by the belief process in learning models. The results in these papers are driven by probability weighting, not diminishing sensitivity.

Lipnowski and Mathevet (2018) study a static model of information design with a psychological receiver whose welfare depends directly on posterior belief. They discuss an application to a mean-based news-utility model *without* diminishing sensitivity in their Appendix A, finding that either one-shot resolution or no information is optimal. We focus on the implications of diminishing sensitivity. Our work also differs in that we study a dynamic problem and examine equilibria without commitment.

Caplin and Leahy (2004) consider a psychological game where an informed sender interacts with a receiver who experiences both anticipatory utility and a preference over the timing of resolution of uncertainty. By contrast, we examine the implications of a different behavioral preference for news. Another difference is that they study a setting with verifiable evidence, but our sender uses cheap-talk messages.

6 Conclusion

In this work, we have studied how diminishingly sensitive gain-loss utilities applied to changes in beliefs affect the agent's informational preferences. If we think that diminishing sensitivity to the magnitude of news is psychologically realistic in this domain, then the stark predictions of the ubiquitous two-part linear models may be misleading. In the presence of diminishing sensitivity, richer informational preferences emerge.

An agent's consumption ranking over the states can determine his preference between an information structure that delivers news about the state gradually and another that results in one-shot resolution. In general, one-shot resolution is neither the best way to get information nor the worst way — skewness matters. One-shot resolution is strictly better than information structures with strictly gradual bad news, one-shot good news. But, it is strictly worse than information structures with strictly gradual good news, one-shot bad news, provided loss aversion is not too high.

For an informed sender who lacks commitment power, diminishing sensitivity leads to novel credibility problems that inhibit any meaningful communication when the receiver has no loss aversion. High enough loss aversion can restore the equilibrium credibility of good-news messages, and the receiver's equilibrium welfare may be non-monotonic in loss aversion. We construct a family of non-babbling equilibria with gradual good news when loss aversion is high enough, finding that the sender must communicate increasingly larger pieces of good news over time in the good state.

Appendix

A Proofs

This appendix contains the proofs of the results stated in the main text.

In the proofs, we will often use the following fact about news-utility functions with diminishing sensitivity. We omit its simple proof.

Fact 1. *Let $d_1, d_2 > 0$ and suppose $\mu(0) = 0$.*

- *(sub-additivity in gains) If $\mu''(x) < 0$ for all $x > 0$, then $\mu(d_1 + d_2) < \mu(d_1) + \mu(d_2)$.*
- *(super-additivity in losses) If $\mu''(x) > 0$ for all $x < 0$, then $\mu(-d_1 - d_2) > \mu(-d_1) + \mu(-d_2)$*

A.1 Proof of Proposition 1

Proof. When $\theta = B$, the agent gets $\mu(-\pi_0)$ with one-shot resolution, but $\sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1})$ with gradual bad news, one-shot good news. For each t, $\pi_t - \pi_{t-1} \leq 0$, and furthermore $\sum_{t=1}^{T} \pi_t - \pi_{t-1} = -\pi_0$ by telescoping and using the fact that $\theta = B$. Due to super-additivity in losses, we get that $\mu(-\pi_0) \geq \sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1})$ almost surely when the state is bad. Also, because there is *strictly* gradual bad news, $\mathbb{E}[\sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1}) \mid \theta = B] < \mu(-\pi_0)$.

When $\theta = A$, the agent gets $\mu(1 - \pi_0)$ with one-shot resolution. With gradual bad news, one-shot good news, let $\hat{T} \leq T$ be the first period where $\pi_{\hat{T}} > \pi_{\hat{T}-1}$. His news utility is $\left[\sum_{t=1}^{\hat{T}-1} \mu(\pi_t - \pi_{t-1})\right] + \mu(1 - \pi_{\hat{T}-1})$ where each $\pi_t - \pi_{t-1} \leq 0$ for $1 \leq t \leq \hat{T} - 1$. Again by super-additivity in losses, $\sum_{t=1}^{\hat{T}-1} \mu(\pi_t - \pi_{t-1}) \leq \mu(\pi_{\hat{T}-1} - \pi_0)$. By sub-additivity in gains, $\mu(1 - \pi_{\hat{T}-1}) \leq \mu(\pi_0 - \pi_{\hat{T}-1}) + \mu(1 - \pi_0) \leq -\mu(\pi_{\hat{T}-1} - \pi_0) + \mu(1 - \pi_0)$, where the last inequality is due to loss aversion. Putting these pieces together,

$$\left[\sum_{t=1}^{\hat{T}-1} \mu(\pi_t - \pi_{t-1})\right] + \mu(1 - \pi_{\hat{T}-1}) \leq \mu(\pi_{\hat{T}-1} - \pi_0) - \mu(\pi_{\hat{T}-1} - \pi_0) + \mu(1 - \pi_0) = \mu(1 - \pi_0).$$

Therefore, strictly gradual bad news, one-shot good news gives strictly lower utility than one-shot resolution in expectation, and almost surely weakly lower utility ex-post. □

A.2 Proof of Proposition 2

Proof. When $\theta = A$, the agent gets $\mu(1-\pi_0)$ with one-shot resolution, but $\sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1})$ with gradual good news, one-shot bad news. For each t, $\pi_t - \pi_{t-1} \geq 0$, and furthermore $\sum_{t=1}^{T} \pi_t - \pi_{t-1} = 1 - \pi_0$ by telescoping and using the fact that $\theta = A$. Due to sub-additivity in gains, we get that $\sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1}) \geq \mu(1 - \pi_0)$ when the state is good. Also, because there is *strictly* gradual good news, $\mathbb{E}[\sum_{t=1}^{T} \mu(\pi_t - \pi_{t-1}) \mid \theta = A] > \mu(1 - \pi_0)$.

When $\theta = B$, the agent gets $\mu(-\pi_0)$ with one-shot resolution. With gradual good news, one-shot bad news, let $\hat{T} \leq T$ be the first period where the $X_{\hat{T}} = 0$. His news utility is $\left[\sum_{t=1}^{\hat{T}-1} \mu(\pi_t - \pi_{t-1})\right] + \mu(-\pi_{\hat{T}-1})$ where each $\pi_t - \pi_{t-1} \geq 0$ for $1 \leq t \leq \hat{T} - 1$. Again by sub-additivity in gains, $\sum_{t=1}^{\hat{T}-1} \mu(\pi_t - \pi_{t-1}) \geq \mu(\pi_{\hat{T}-1} - \pi_0)$. By super-additivity in losses, $\mu(-\pi_{\hat{T}-1}) \geq \mu(-(\pi_{\hat{T}-1} - \pi_0)) + \mu(-\pi_0) = -\mu(\pi_{\hat{T}-1} - \pi_0) + \mu(-\pi_0)$, where we used the symmetry of μ around 0 in the last equality. Putting these pieces together,

$$\left[\sum_{t=1}^{\hat{T}-1} \mu(\pi_t - \pi_{t-1})\right] + \mu(-\pi_{\hat{T}-1}) \geq \mu(\pi_{\hat{T}-1} - \pi_0) - \mu(\pi_{\hat{T}-1} - \pi_0) + \mu(-\pi_0) = \mu(-\pi_0).$$

Therefore, strictly gradual good news, one-shot bad news provides strictly higher utility than one-shot resolution in expectation, and almost surely weakly higher utility ex-post. □

A.3 Proof of Corollary 1

Proof. This follows from Proposition 2 by continuity. □

A.4 Proof of Proposition 3

We begin by giving some additional definition and notation. For $p, \pi \in [0,1]$, let $N_B(x;\pi) := \mu(x-\pi) + \mu(-x)$ denote the total amount of news utility across two periods when the receiver updates his belief from π to $x > \pi$ today and updates it from x to 0 tomorrow. Similarly, $N_A(p;\pi) := \mu(p-\pi) + \mu(1-p)$.

We state some preliminary lemmas about N_A and N_B.

Lemma A.1. *If μ is symmetric around 0 and $\mu^{''}(x) < 0$ for all $x > 0$, then for any $0 < \pi < x < 1$ it holds $N_B(0;\pi) < N_B(x;\pi)$.*

Proof. Due to sub-additivity,

$$\mu(p) < \mu(p-\pi) + \mu(\pi). \tag{1}$$

Note that symmetry implies $\mu(-p) = -\mu(p)$ and that $\mu(-\pi) = -\mu(\pi)$. Rearranged, Equation (1) is precisely $N(0;\pi) < N(p;\pi)$. □

Say μ exhibits *greater sensitivity to losses* if $\mu^{'}(x) \leq \mu^{'}(-x)$ for all $x > 0$.

Lemma A.2. *Suppose μ exhibits diminishing sensitivity and greater sensitivity to losses. Then, $p \mapsto N_A(p;\pi)$ is strictly increasing on $[0,\pi]$ and symmetric on the interval $[\pi,1]$. For each $p_1 \in [\pi,1]$, there exists exactly one point $p_2 \in [\pi,1]$ so that $N_A(p_1;\pi) = N_A(p_2;\pi)$. For every $p_L < \pi$ and $p_H \geq \pi$, $N_A(p_L;\pi) < N_A(p_H;\pi)$. Also, $N_B(p;\pi)$ is symmetric on the interval $[0,\pi]$. For each $p_1 \in [0,\pi]$, there exists exactly one point $p_2 \in [0,\pi]$ so that $N_B(p_1;\pi) = N_B(p_2;\pi)$.*

Proof. We have $\frac{\partial N_A(p;\pi)}{\partial p} = \mu^{'}(p-\pi) - \mu^{'}(1-p)$. For $0 \leq p < \pi$ and under greater sensitivity to losses, $\mu^{'}(p-\pi) \geq \mu^{'}(\pi-p)$. Since $\mu^{''}(x) < 0$ for $x > 0$, $\mu^{'}(\pi-p) > \mu^{'}(1-p)$. This shows $\frac{\partial N_A(p;\pi)}{\partial p} > 0$ for $p \in [0,\pi)$.

The symmetry results follow from simple algebra and do not require any assumptions.

Note that $\frac{\partial^2 N_A(p;\pi)}{\partial p^2} = \mu^{''}(p-\pi) + \mu^{''}(1-p) < 0$ for any $p \in [\pi,1]$, due to diminishing sensitivity. Combined with the required symmetry, this means $\frac{\partial N_A(p;\pi)}{\partial p}$ crosses 0 at most once on $[\pi,1]$, so for each $p_1 \in [\pi,1]$, we can find at most one p_2 so that $N_A(p_1;\pi) = N_A(p_2;\pi)$. In particular, this implies at every intermediate $p_1 \in (\pi,1)$, we get $N_A(p_1;\pi) > N_A(\pi;\pi)$ since we already have $N_A(1;\pi) = N_A(\pi;\pi)$. This shows $N_A(\cdot;\pi)$ is strictly larger on $[\pi,1]$ than on $[0,\pi)$.

A similar argument, using $\mu^{''}(x) > 0$ for $x < 0$, establishes that for each $p_1 \in [0, \pi]$, we can find at most one p_2 so that $N_B(p_1; \pi) = N_B(p_2; \pi)$. □

Consider any period $T - 2$ history h_{T-2} in any equilibrium (M, σ^*, p^*) where $p^*(h_{T-2}) = \pi \in (0, 1)$. Let P_A and P_B represent the sets of posterior beliefs induced at the end of $T - 1$ with positive probability, in states A and B. The next lemma gives an exhaustive enumeration of all possible P_A, P_B.

Lemma A.3. *Suppose μ exhibits diminishing sensitivity and greater sensitivity to losses. The sets P_A, P_B belong to one of the following cases.*

1. $P_A = P_B = \{\pi\}$
2. $P_A = \{1\}$, $P_B = \{0\}$
3. $P_A = \{p_1\}$ for some $p_1 \in (\pi, 1)$ and $P_B = \{0, p_1\}$
4. $P_A = \{\pi, 1\}$ and $P_B = \{0, \pi\}$
5. $P_A = \{p_1, p_2\}$ for some $p_1 \in (\pi, \frac{1+\pi}{2})$, $p_2 = 1 - p_1 + \pi$, $P_B = \{0, p_1, p_2\}$.

Proof. Suppose $|P_A| = 1$.

If $P_A = \{\pi\}$, then any equilibrium message not inducing π must induce 0. By Bayes' rule, the sender cannot induce belief 0 with positive probability in the bad state, so $P_B = \{\pi\}$ as well.

If $P_A = \{1\}$, then any equilibrium message not inducing 1 must induce 0. Furthermore, the sender cannot send equilibrium messages inducing belief 1 with positive probability in the bad state, else the equilibrium belief associated with these messages should be strictly less than 1. Thus $P_B = \{0\}$.

If $P_A = \{p_1\}$ for some $0 \le p_1 < \pi$, then any equilibrium message not inducing p_1 must induce 0. This is a contradiction since the posterior beliefs do not average out to π.

This leaves the case of $P_A = \{p_1\}$ for some $\pi < p_1 < 1$. Any equilibrium message not inducing p_1 must induce 0. Furthermore, the sender must induce the belief p_1 in the bad state with positive probability, else we would have $p_1 = 1$. At the same time, the sender must also induce belief 0 with positive probability in the bad state, else we violate Bayes' rule. So $P_B = \{0, p_1\}$.

Now suppose $|P_A| = 2$.

In the good state, the sender must be indifferent between two beliefs p_1, p_2 both induced with positive probability. By Lemma A.2, $N_A(p;\pi)$ is strictly increasing on $[0,\pi]$ and strictly higher on $[\pi,1]$ than on $[0,\pi)$, while for each $p_1 \in [\pi,1]$, there exists exactly one point $p_2 \in [\pi,1]$ so that $N_A(p_1;\pi) = N_A(p_2;\pi)$. This means we must have $p_1 \in [\pi, \frac{1+\pi}{2}]$, $p_2 = 1 - p_1 + \pi$.

If $P_A = \{\pi, 1\}$, any equilibrium message not inducing π or 1 must induce 0. Also, $1 \notin P_B$, because any message sent with positive probability in the bad state cannot induce belief 1. We cannot have $P_B = \{0\}$, because then the message inducing belief π actually induces 1. We cannot have $P_B = \{\pi\}$ for then we violate Bayes' rule. This leaves only $P_B = \{0, \pi\}$.

If $P_A = \{p_1, p_2\}$ for some $p_1 \in (\pi, \frac{1+\pi}{2})$, then any equilibrium message not inducing p_1 or p_2 must induce 0. Also, $p_1, p_2 \in P_B$, else messages inducing these beliefs give conclusive evidence of the good state. By Bayes' rule, we must have $P_B = \{0, p_1, p_2\}$.

It is impossible that $|P_A| \geq 3$, since, by Lemma A.2, $N_A(p;\pi)$ is strictly increasing on $[0,\pi]$ and strictly higher on $[\pi,1]$ than on $[0,\pi)$, while for each $p_1 \in [\pi,1]$, there exists exactly one point $p_2 \in [\pi,1]$ so that $N_A(p_1;\pi) = N_A(p_2;\pi)$. So the sender cannot be indifferent between 3 or more different posterior beliefs of the receiver in the good state. □

We now give the proof of Proposition 3.

Proof. The hypothesis that μ is symmetric also implies that it exhibits greater sensitivity to losses. So, Lemmas A.1 and A.3 apply. Consider any period $T-2$ history h^{T-2} with $p^*(h^{T-2}) \in (0,1)$. By Lemma A.1, $N_B(p; p^*(h^{T-2})) > N_B(0; p^*(h^{T-2}))$ for all $p \in (p^*(h^{T-2}), 1]$. Therefore, cases 3 and 5 are ruled out from the conclusion of Lemma A.3. This shows that after having reached history h^{T-2}, the receiver will get total news utility of $\mu(1 - p^*(h^{T-2}))$ in the good state and $\mu(-p^*(h^{T-2}))$ in the bad state. This conclusion applies to all period $T-2$ histories (including those with equilibrium beliefs 0 or 1). So, the sender gets the same utility as if the state is perfectly revealed in period $T-1$ rather than T, and the equilibrium up to period $T-1$ form an equilibrium of the cheap talk game with horizon $T-1$. By backwards induction, we see that along the equilibrium path, whenever the receiver's belief updates, it is updated to the dogmatic belief in θ. □

A.5 Detailed Calculations for Example 1

Example 1's μ exhibits diminishing sensitivity and greater sensitivity to losses, so Lemma A.3 applies. We use the classification from Lemma A.3 with $T = 2$ and $\pi = 1/2$. An equilibrium belonging to case 1 or case 2 gives the same payoff as the babbling equilibrium, since all uncertainty is resolved in one period. For an equilibrium belonging to case 4, in each state, the sender fully reveals the state with positive probability. The indifference condition implies the sender must get the same payoff as she would from always fully revealing the state in period $t = 1$. So, such an equilibrium would again have the same payoff as the babbling equilibrium.

Since only cases 3 and 5 remain, there is an equilibrium with a payoff different than that of the babbling equilibrium only if there exists some $x \in (1/2, 1)$ so that $\sqrt{x - 0.5} - \lambda\sqrt{x} = -\lambda\sqrt{0.5}$ — that is, when $\theta = B$, the sender is indifferent between inducing the belief x and revealing the state in period $t = 1$. By straightforward algebra, the two solutions for x are $\frac{1}{2} \cdot \left(\frac{\lambda^2+1}{\lambda^2-1}\right)^2$ and $\frac{1}{2}$. The latter is not in the open interval $(1/2, 1)$, and the former is in this interval if and only if $\lambda > 1 + \sqrt{2}$.

This analysis shows the sender can be indifferent between 0 and up to one belief in the interval $(1/2, 1)$ when the $\theta = B$, so case 5 from Lemma A.3 is also ruled out. There exists an equilibrium with strictly higher payoff than the babbling equilibrium if and only if $\lambda > 1 + \sqrt{2}$. In this other equilibrium, case 3 must hold: that is, the sender induces the belief $\frac{1}{2} \cdot \left(\frac{\lambda^2+1}{\lambda^2-1}\right)^2$ in period $t = 1$ if $\theta = A$, and induces either the belief $\frac{1}{2} \cdot \left(\frac{\lambda^2+1}{\lambda^2-1}\right)^2$ or the belief 0 in period $t = 1$ if $\theta = B$.

A.6 Proof of Proposition 4

Proof. Let J intermediate beliefs satisfying the hypotheses be given. We construct a gradual good news equilibrium where $p_t = q^{(t)}$ for $1 \le t \le J$, and $p_t = q^{(J)}$ for $J + 1 \le t \le T - 1$.

Without loss of generality, let $M = \{a, b\}$ and consider the following strategy profile. In period $t \le J$ where the public history so far h^{t-1} does not contain any b, let $\sigma(h^{t-1}; A)(a) = 1$, $\sigma(h^{t-1}; B)(a) = x$ where $x \in (0, 1)$ satisfies $\frac{p_{t-1}}{p_{t-1}+(1-p_{t-1})x} = p_t$. But if public history contains at least one b, then $\sigma(h^{t-1}; A)(b) = 1$ and $\sigma(h^{t-1}; B)(b) = 1$. Finally, if the period is $t > J$, then $\sigma(h^{t-1}; A)(b) = 1$ and $\sigma(h^{t-1}; B)(b) = 1$. In terms of beliefs, suppose h^t has $t \le J$ and every message so far has been a. Such histories are on-path and get assigned the Bayesian posterior belief. If h^t has $t \le J$ and contains at least one b, then it gets assigned belief 0. Finally, if h^t has $t > J$, then h^t gets assigned the same belief as the subhistory constructed from its first J elements. It is easy to verify that these beliefs are derived from Bayes' rule whenever possible.

We verify that the sender has no incentive to deviate. Consider period $t \leq J$ with history h^{t-1} that does not contain any b. The receiver's current belief is p_{t-1} by construction.

In state B, we first calculate the sender's equilibrium payoff after sending a. The receiver will get some I periods of good news before the bad state is revealed, either by the sender or by nature in period T. That is, the equilibrium news utility with I periods of good news is given by $\sum_{i=1}^{I} \mu(p_{t-1+i} - p_{t-2+i}) + \mu(-p_{t-1+I})$. Since $p_{t-1+I} \in P^*(p_{t-2+I})$, we have $N_B(p_{t-1+I}; p_{t-2+I}) = N_B(0; p_{t-2+I})$, that is to say $\mu(p_{t-1+I} - p_{t-2+I}) + \mu(-p_{t-1+I}) = \mu(-p_{t-2+I})$. We may therefore rewrite the receiver's total news utility as $\sum_{i=1}^{I-1} \mu(p_{t-1+i} - p_{t-2+i}) + \mu(-p_{t-2+I})$. But by repeating this argument, we conclude that the receiver's total news utility is just $\mu(-p_{t-1})$. Since this result holds regardless of I's realization, the sender's expected total utility from sending a today is $\mu(-p_{t-1})$, which is the same as the news utility from sending b today. Thus, sender is indifferent between a and b and has no profitable deviation.

In state A, the sender gets at least $\mu(1 - p_{t-1})$ from following the equilibrium strategy. This is because the receiver's total news utility in the good state along the equilibrium path is given by $\sum_{i=1}^{J-(t-1)} \mu(p_{t-1+i} - p_{t-2+i}) + \mu(1 - p_{t-1+I})$. By sub-additivity in gains, this sum is strictly larger than $\mu(1 - p_{t-1})$. If the sender deviates to sending b today, then the receiver updates belief to 0 today and belief remains there until the exogenous revelation, when belief updates to 1. So this deviation gives the total news utility $\mu(-p_{t-1}) + \mu(1)$. We have

$$\begin{aligned}\mu(1) &< \mu(1 - p_{t-1}) + \mu(p_{t-1}) \\ &\leq \mu(1 - p_{t-1}) - \mu(-p_{t-1}),\end{aligned}$$

where the first inequality comes from sub-additivity in gains, and the second from weak loss aversion. This shows $\mu(-p_{t-1}) + \mu(1) < \mu(1 - p_{t-1})$, so the deviation is strictly worse than sending the equilibrium message.

Finally, at a history containing at least one b or a history with length J or longer, the receiver's belief is the same at all continuation histories. So the sender has no deviation incentives since no deviations affect future beliefs.

For the other direction, suppose by way of contradiction there exists a gradual good news equilibrium with the J intermediate beliefs $q^{(1)} < ... < q^{(J)}$. For a given $1 \leq j \leq J$, find the smallest t such that $p_t = q^{(k-1)}$ and $p_{t+1} = q^{(k)}$. At every on-path history $h^t \in H^t$ with $p^*(h^t) = p_t$, we must have $\sigma^*(h^t; B)$ inducing both 0 and $q^{(j)}$ with strictly positive probability. Since we are in equilibrium, we must have $\mu(-q^{(j-1)})$ being equal to $\mu(q^{(j)} - q^{(j-1)})$ plus the continuation payoff. If $j = J$, then this continuation payoff is $\mu(-q^{(j)})$ as the only other period of belief movement is in period T when the receiver learns the state is bad. If $j < J$, then find the smallest $\bar{t}$ so that $p_{\bar{t}+1} = q^{(j+1)}$. At any on-path $h^{\bar{t}} \in H^{\bar{t}}$ which is a continuation of h^t, we have $p^*(h^{\bar{t}}) = q^{(j)}$ and the receiver has not experienced any news utility in periods $t+2, ..., \bar{t}$. Also, $\sigma^*(h^{\bar{t}}; B)$ assigns positive probability to inducing posterior belief 0, so the continuation payoff in question must be $\mu(-q^{(j)})$. So we have shown that $\mu(-q^{(j-1)}) = \mu(q^{(j)} - q^{(j-1)}) + \mu(-q^{(j)})$, that is $N_B(q^{(j)}; q^{(j-1)}) = N_B(0; q^{(j-1)})$. $\square$

A.7 Proof of Corollary 2

Proof. We apply Proposition 4 to the case of quadratic news utility. Recall the relevant indifference equation in the good state.

$$\mu(-q_t) = \mu(q_{t+1} - q_t) + \mu(-q_{t+1}). \tag{2}$$

Plugging in the quadratic specification and algebraic transformations lead to

$$0 = (\alpha_p - \alpha_n)(q_{t+1} - q_t) - \beta_p(q_{t+1} - q_t) + \beta_n(q_{t+1} - q_t)(q_{t+1} + q_t)$$

Define $r = q_{t+1} - q_t$. Then this relation can be written as

$$(\beta_p - \beta_n)r^2 + (\alpha_n - \alpha_p - 2\beta_n q_t)r = 0,$$

i.e. r is a zero of a second-order polynomial. For P^* to be non-empty we need this root r to be in $(0, 1 - q_t)$. In particular the critical point $\bar{r}$ of the second-order polynomial should satisfy $\bar{r} \in (0, \frac{1-q_t}{2})$. Given that $\bar{r} = \frac{2\beta_n q_t - (\alpha_n - \alpha_p)}{2(\beta_p - \beta_n)}$ for the case that $\beta_p \neq \beta_n$, we get the equivalent condition on the primitives $0 < \frac{2\beta_n q_t - (\alpha_n - \alpha_p)}{2(\beta_p - \beta_n)} < \frac{1-q_t}{2}$. The root

r itself is given by $r = \frac{2\beta_n q_t - (\alpha_n - \alpha_p)}{\beta_p - \beta_n}$, which leads to the recursion

$$q_{t+1} = q_t \frac{\beta_p + \beta_n}{\beta_p - \beta_n} - \frac{\alpha_n - \alpha_p}{\beta_p - \beta_n}. \tag{3}$$

This leads to the formula for $P^*(\pi)$ in the statement of the corollary.

Case 1: When $\beta_p < \beta_n$ the coefficient in front of q_t is negative so that the recursion in Equation (3) leads to

$$q_{t+1} - q_t = q_t \frac{2\beta_n}{\beta_p - \beta_n} - \frac{\alpha_n - \alpha_p}{\beta_p - \beta_n} < 0.$$

This also shows that for the case that $\beta_p < \beta_n$, a GGN equilibrium with 1 or more intermediate beliefs only exists when the prior is low enough: namely $\pi_0 < \frac{\alpha_n - \alpha_p}{2\beta_n}$.

Case 2: When $\beta_p > \beta_n$ the slope in Equation (3) is above 1. So, for all priors π_0 large enough, we get an increasing sequence q_t which satisfies Equation (2). It is also easy to see from Equation (3) that

$$(q_{t+2} - q_{t+1}) - (q_{t+1} - q_t) = \left(\frac{\beta_p + \beta_n}{\beta_p - \beta_n} - 1 \right) > 0,$$

proving the final statement of the corollary.

The existence of an equilibrium with more than one intermediate belief is shown by the example in Figure 3. □

A.8 Proof of Proposition 5

Proof. Let any $0 < \pi < 1$ be given, and consider $N_B(p;\pi) - N_B(0;\pi)$ as a function of $p \geq \pi$. When $p = \pi$, we have $N_B(p;\pi) - N_B(0;\pi) = 0$. For any $\epsilon > 0$ such that $\pi + \epsilon \leq 1$, we get $N_B(\pi + \epsilon;\pi) - N_B(0;\pi) = \int_0^\epsilon \mu'(x)dx - \int_{-\pi-\epsilon}^{-\pi} \mu'(x)dx$. Since $\mu'(0_-) \leq \mu(0_+)$ and μ' is strictly increasing in $[-1, 0)$, $\mu'(x) < \mu'(0_+)$ for all $x \leq -\pi$. Also, since μ' is continuous in $(0, 1]$, for small enough $\epsilon > 0$ we also get $\mu'(x) < \mu'(y)$ for all $x \leq -\pi$, $y \in (0, \epsilon]$. Thus, $\int_0^\epsilon \mu'(x)dx - \int_{-\pi-\epsilon}^{-\pi} \mu'(x)dx > 0$ for $\epsilon > 0$ close enough to 0.

This analysis shows $N_B(p;\pi) - N_B(0;\pi)$ is strictly positive for some range of p slightly above π. Given that $|P^*(\pi)| \leq 1$, if we find some $p' > \pi$ with $N_B(p';\pi) - N_B(0;\pi) > 0$, then any solution to $N_B(p;\pi) - N_B(0;\pi) = 0$ in $(\pi, 1)$ must lie to the right of $p^{'}$.

If $q^{(j)}, q^{(j+1)}$ are intermediate beliefs in a GGN equilibrium, then by Proposition 4, $q^{(j)} \in P^*(q^{(j-1)})$ and $q^{(j+1)} \in P^*(q^{(j)})$. Let $p^{'} = q^{(j)} + (q^{(j)} - q^{(j-1)})$. Then,

$$\begin{aligned} N_B(p^{'}; q^{(j)}) - N_B(0; q^{(j)}) &= \mu(p^{'} - q^{(j)}) + \mu(-p^{'}) - \mu(-q^{(j)}) \\ &= \mu(q^{(j)} - q^{(j-1)}) + \mu(-q^{(j)} - (q^{(j)} - q^{(j-1)})) - \mu(-q^{(j)}) \\ &> \mu(q^{(j)} - q^{(j-1)}) + \mu(-q^{(j-1)} - (q^{(j)} - q^{(j-1)})) - \mu(-q^{(j-1)}), \end{aligned}$$

where the last inequality comes from diminishing sensitivity. But, the final expression is $N_B(q^{(j)}; q^{(j-1)}) - N_B(0; q^{(j-1)})$, which is 0 since $q^{(j)} \in P^*(q^{(j-1)})$. This shows we must have $q^{(j+1)} - q^{(j)} > q^{(j)} - q^{(j-1)}$. □

A.9 Proof of Proposition 6

Proof. (1) Suppose μ is two-part linear with $\mu(x) = x$ for $x \geq 0$, $\mu(x) = \lambda x$ for $x < 0$, where $\lambda \geq 0$. Suppose $v(A) = 1$, $v(B) = 0$. In each period, $\mathbb{E}[\mu(\pi_t - \pi_{t-1})] = \mathbb{E}[(\pi_t - \pi_{t-1})^+ - \lambda(\pi_t - \pi_{t-1})^-]$. By the martingale property, $\mathbb{E}[(\pi_t - \pi_{t-1})^+] = \mathbb{E}[(\pi_t - \pi_{t-1})^-]$, so $\mathbb{E}[\mu(\pi_t - \pi_{t-1})] = \frac{1}{2}(1 - \lambda)\mathbb{E}[|\pi_t - \pi_{t-1}|]$. This shows total expected news utility is $\mathbb{E}[\sum_{t=1}^T \mu(\pi_t - \pi_{t-1})] = \frac{1}{2}(1 - \lambda)\mathbb{E}[\sum_{t=1}^T |\pi_t - \pi_{t-1}|]$. Note that $\mathbb{E}[\sum_{t=1}^T |\pi_t - \pi_{t-1}|]$ is strictly larger for gradual information than for one-shot resolution. If $\lambda > 1$, the agent strictly prefers one-shot resolution. If $0 \leq \lambda < 1$, the agent strictly prefers gradual information. If $\lambda = 1$, the agent is indifferent.

Now suppose $v(A) = 0$, $v(B) = 1$. By the same arguments, total expected news utility is $\mathbb{E}[\sum_{t=1}^T \mu(-(\pi_t - \pi_{t-1}))] = \frac{1}{2}(1 - \lambda)\mathbb{E}[\sum_{t=1}^T |\pi_t - \pi_{t-1}|]$. Note that $\mathbb{E}[\sum_{t=1}^T |\pi_t - \pi_{t-1}|]$ is strictly larger for gradual information than for one-shot resolution. So again, if $\lambda > 1$, the agent strictly prefers one-shot resolution. If $0 \leq \lambda < 1$, the agent strictly prefers gradual information. If $\lambda = 1$, the agent is indifferent.

(2) If W is linear, then the agent is indifferent between gradual information and one-shot resolution regardless of the sign of $v(A) - v(B)$. If W is strictly concave, let $\rho_t = \pi_t \cdot v(A) + (1 - \pi_t) \cdot v(B)$ be the agent's expectation of future consumption utility in period t. For $1 \leq t \leq T - 1$, $\mathbb{E}[W(\rho_t)] < W(\rho_0)$ by combining the martingale property and Jensen's inequality. So the agent strictly prefer to keep his prior beliefs until the last period and will therefore choose one-shot resolution, regardless of the sign of $v(A) - v(B)$.

(3) Ely, Frankel, and Kamenica (2015) mention a "state-dependent" specification of their suspense and surprise utility functions. With two states, A and B, their specification uses weights $\alpha_A, \alpha_B > 0$ to differentially re-scale belief-based utilities for movements in the two different directions. Specifically, their re-scaled suspense utility is

$$\sum_{t=0}^{T-1} u\left(\mathbb{E}_t\left[\alpha_A \cdot (\pi_{t+1} - \pi_t)^2 + \alpha_B \cdot ((1 - \pi_{t+1}) - (1 - \pi_t))^2\right]\right)$$

and their re-scaled surprise utility is

$$\mathbb{E}\left[\sum_{t=1}^{T} u\left(\alpha_A \cdot (\pi_{t+1} - \pi_t)^2 + \alpha_B \cdot ((1 - \pi_{t+1}) - (1 - \pi_t))^2\right)\right].$$

We may consider agents with opposite preferences over states A and B as agents with different pairs of scaling weights (α_A, α_B). Specifically, say there are $\alpha^{\text{High}} > \alpha^{\text{Low}} > 0$. For an agent preferring A, $\alpha_A = \alpha^{\text{High}}, \alpha_B = \alpha^{\text{Low}}$. For an agent preferring B, $\alpha_A = \alpha^{\text{Low}}, \alpha_B = \alpha^{\text{High}}$. But note that we always have $\pi_{t+1} - \pi_t = -[(1 - \pi_{t+1}) - (1 - \pi_t)]$, so along every realized path of beliefs, $(\pi_{t+1} - \pi_t)^2 = ((1 - \pi_{t+1}) - (1 - \pi_t))^2$. This means these two agents with the opposite scaling weights actually have identical objectives and therefore will have the same preference over gradual information or one-shot resolution. □

B Further Results

B.1 Optimal Information Structure for Anticipatory Utility

We show that if the agent has anticipatory utility and gets $W(\pi_t)$ when he ends period t with posterior belief π_t, then with commitment power, there exists an optimal information structure that only discloses information in period $t = 1$.

Consider any information structure. Find the period t^* with the highest ex-ante anticipatory utility under this information structure, i.e., $t^* \in \underset{1 \le t \le T-1}{\arg\max} \mathbb{E}\left[W(\pi_t)\right]$. Consider another information structure that generates the (feasible) distribution of beliefs π_{t^*} in period 1, then reveals no additional information in periods $2, ..., T-1$. This new information structure gives weakly higher expected anticipatory utility than the given information structure in every period. Therefore there exists an optimal information structure that only discloses information in $t = 1$.

B.2 Preference for Dominated Consumption Lotteries

So far, we have taken the prior distribution over states π_0 as exogenously given. Fixing an information structure, a news-utility agent may strictly prefer a dominated distribution over states. This distinguishes our news-utility preference from other preferences, such as recursive preferences and Gul, Natenzon, and Pesendorfer (2021)'s risk consumption preference.

We now give an example. Suppose $T = 2$ and there are two states, $\Theta = \{A, B\}$. Normalize consumption utility to be $v(A) = 1$, $v(B) = 0$. Let the news utility function be $\mu(z) = \sqrt{z}$ for $z \geq 0$, $\mu(z) = -\lambda\sqrt{-z}$ for $z < 0$, where $\lambda \geq 1$. At time $t = 0$, the agent holds a prior belief $\pi_0 \in [0, 1]$. At time $t = 1$, the agent learns the state perfectly, so π_1 is degenerate with probability 1. Consumption takes place at time $t = 2$. For any λ, the agent strictly prefers state A for sure ($\pi_0 = 1$) over state B for sure ($\pi_0 = 0$), as both environments provide zero news utility. But, the agent may strictly prefer state B for sure over an interior probability of state A, $\pi_0 = p \in (0, 1)$. In fact, this happens when $p + p\sqrt{1-p} - \lambda(1-p)\sqrt{p} < 0$, which says $\lambda > \frac{\sqrt{p}(1+\sqrt{1-p})}{1-p}$. A sufficiently loss-averse agent may strictly prefer no chance of winning a consumption lottery than a low chance of winning.

B.3 State-Dependent Suspense and Surprise

Suppose there are two states, $\Theta = \{A, B\}$, and suppose that the agent has either the suspense objective $\sum_{t=0}^{T-1} u\left(\mathbb{E}_t(\sum_\theta \alpha_\theta \cdot (\pi_{t+1}(\theta) - \pi_t(\theta))^2\right)$ or the surprise objective $\sum_{t=1}^{T} u\left(\sum_\theta \alpha_\theta \cdot (\pi_t(\theta) - \pi_{t-1}(\theta))^2\right)$, where $\alpha_A, \alpha_B > 0$ are state-dependent scaling weights. We must have $\pi_{t+1}(A) - \pi_t(A) = -(\pi_{t+1}(B) - \pi_t(B))$, so pathwise $(\pi_{t+1}(A) - \pi_t(A))^2 = (\pi_{t+1}(B) - \pi_t(B))^2$. This shows that the new objectives obtained by applying two possibly different scaling weights $\alpha_A \neq \alpha_B$ to states A and B are identical to the ones that would be obtained by applying the *same* scaling weight $\alpha = \frac{\alpha_A + \alpha_B}{2}$ to both states. Due to this symmetry in preference, the optimal information structure for entertaining an agent with state-dependent suspense or surprise utility treats the two states symmetrically, in contrast to a central prediction of diminishing sensitivity in our model.

B.4 Risk Consumption Preferences

Gul, Natenzon, and Pesendorfer (2021) study a model of preference over random evolving lotteries and propose a class of risk consumption preferences. Translated into our setting, an agent with risk consumption preference values an information structure (M, σ) according to utility function

$$\mathbb{E}\left[\int v(u_2(\pi_t))d\eta\right].$$

Here $u_2 : \Delta(\Theta) \to \mathbb{R}$ is affine and v is strictly increasing. The term $v(u_2(\pi_t))$ is viewed as a function from the time periods $\{0, 1, ..., T-1\}$ into the reals and $d\eta$ denotes the Choquet integral with respect to a capacity η on $\{0, 1, ..., T-1\}$.

To show that our model of mean-based news utility is not nested under the class of risk consumption preferences, we show that risk consumption preferences cannot exhibit the preference patterns from Appendix B.2: that is, strictly preferring winning a lottery for sure to not winning it for sure, but also strictly preferring not winning for sure to winning with some interior probability $p \in (0, 1)$ in the $T = 2$ setup.

By an abuse of notation, the belief assigning probability q to state A will simply be denoted q. The first part of the preference gives $v(u_2(1)) > v(u_2(0))$, since Choquet integral of a constant function returns the same constant. When the prior winning probability is $p \in (0, 1)$, the Choquet integrand is either $f_A : \{0, 1\} \to \mathbb{R}$ with $f_A(0) = v(u_2(p))$ and $f_A(1) = v(u_2(1))$, or $f_B : \{0, 1\} \to \mathbb{R}$ with $f_B(0) = v(u_2(p))$ and $f_B(0) = v(u_2(0))$. The two integrands correspond to belief paths where the agent wins or loses the lottery. Since v is strictly increasing, u_2 is affine, and $v(u_2(1)) > v(u_2(0))$, we have $v(u_2(p)) > v(u_2(0))$. Thus both f_G and f_B dominate the constant function $v(u_2(0))$ in every period. By monotonicity of the Choquet integral, the agent must prefer p probability of winning the lottery to no chance of winning it.

References

Barberis, N. and W. Xiong (2012): "Realization utility," *Journal of Financial Economics*, 104, 251–271.

Battigalli, P. and M. Dufwenberg (2022): "Belief-dependent motivations and psychological game theory," *Journal of Economic Literature*, 60, 833–82.

Bellemare, C., M. Krause, S. Kröger, and C. Zhang (2005): "Myopic loss aversion: Information feedback vs. investment flexibility," *Economics Letters*, 87, 319–324.

Bowman, D., D. Minehart, and M. Rabin (1999): "Loss aversion in a consumption–savings model," *Journal of Economic Behavior and Organization*, 38, 155–178.

Budescu, D. V. and I. Fischer (2001): "The same but different: an empirical investigation of the reducibility principle," *Journal of Behavioral Decision Making*, 14, 187–206.

Campos-Mercade, P., L. Goette, T. Graeber, A. Kellogg, and C. Sprenger (2022): "Heterogeneity of gain-loss attitudes and expectations-based reference points," *Working Paper*.

Caplin, A. and J. Leahy (2004): "The supply of information by a concerned expert," *The Economic Journal*, 114, 487–505.

Chapman, J., E. Snowberg, S. W. Wang, and C. Camerer (2022): "Looming large or seeming small? Attitudes towards losses in a representative sample," *Working Paper*.

Dillenberger, D. (2010): "Preferences for one-shot resolution of uncertainty and Allais-type behavior," *Econometrica*, 78, 1973–2004.

Dillenberger, D. and C. Raymond (2020): "Additive-belief-based preferences," *Working Paper*.

Dreyfuss, B., O. Heffetz, and M. Rabin (2022): "Expectations-based loss aversion may help explain seemingly dominated choices in strategy-proof mechanisms," *American Economic Journal: Microeconomics*, 14, 515–555.

Duraj, J. (2019): "Mechanism design with news utility," *Working Paper*.

Ebert, S. and P. Strack (2015): "Until the bitter end: On prospect theory in a dynamic context," *American Economic Review*, 105, 1618–1633.

——— (2018): "Never, ever getting started: On prospect theory without commitment," *Working Paper*.

Eliaz, K. and R. Spiegler (2006): "Can anticipatory feelings explain anomalous choices of information sources?" *Games and Economic Behavior*, 56, 87–104.

Ely, J., A. Frankel, and E. Kamenica (2015): "Suspense and surprise," *Journal of Political Economy*, 123, 215–260.

Falk, A. and F. Zimmermann (2023): "Attention and dread: Experimental evidence on preferences for information," *Management Science, forthcoming.*

Grossman, S. J. and M. Perry (1986): "Sequential bargaining under asymmetric information," *Journal of Economic Theory*, 39, 120–154.

Gul, F., P. Natenzon, E. Y. Ozbay, and W. Pesendorfer (2020): "The thrill of gradual learning," *Working Paper.*

Gul, F., P. Natenzon, and W. Pesendorfer (2021): "Random evolving lotteries and intrinsic preference for information," *Econometrica*, 89, 2225–2259.

Henderson, V. (2012): "Prospect theory, liquidation, and the disposition effect," *Management Science*, 58, 445–460.

Kahneman, D. and A. Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica*, 47, 263–292.

Kőszegi, B. (2006): "Emotional agency," *Quarterly Journal of Economics*, 121, 121–155.

Kőszegi, B. and M. Rabin (2009): "Reference-dependent consumption plans," *American Economic Review*, 99, 909–36.

Kyle, A. S., H. Ou-Yang, and W. Xiong (2006): "Prospect theory and liquidation decisions," *Journal of Economic Theory*, 129, 273–288.

Lipnowski, E. and L. Mathevet (2018): "Disclosure to a psychological audience," *American Economic Journal: Microeconomics*, 10, 67–93.

Lipnowski, E. and D. Ravid (2020): "Cheap talk with transparent motives," *Econometrica*, 88, 1631–1660.

Masatlioglu, Y., A. Y. Orhun, and C. Raymond (2017): "Intrinsic information preferences and skewness," *Working Paper.*

Nielsen, K. (2020): "Preferences for the resolution of uncertainty and the timing of information," *Journal of Economic Theory*, 189, 105090.

O'Donoghue, T. and C. Sprenger (2018): "Reference-dependent preferences," *Handbook of Behavioral Economics-Foundations and Applications 1*, 1.

Pagel, M. (2016): "Expectations-based reference-dependent preferences and asset pricing," *Journal of the European Economic Association*, 14, 468–514.

——— (2017): "Expectations-based reference-dependent life-cycle consumption," *Review of Economic Studies*, 84, 885–934.

——— (2018): "A news-utility theory for inattention and delegation in portfolio choice," *Econometrica*, 86, 491–522.

Rabin, M. and G. Weizsäcker (2009): "Narrow bracketing and dominated choices," *American Economic Review*, 99, 1508–43.

Schweizer, N. and N. Szech (2018): "Optimal revelation of life-changing information," *Management Science*, 64, 5250–5262.

Segal, U. (1990): "Two-stage lotteries without the reduction axiom," *Econometrica*, 349–377.

Shefrin, H. and M. Statman (1985): "The disposition to sell winners too early and ride losers too long: Theory and evidence," *Journal of Finance*, 40, 777–790.

Tversky, A. and D. Kahneman (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 5, 297–323.

Zimmermann, F. (2014): "Clumped or piecewise? Evidence on preferences for information," *Management Science*, 61, 740–753.