Dynamic Information Design with Diminishing Sensitivity Over News

Jetlir Duraj Kevin He

February 27, 2020

The Environment

Motivation: when giving news, people mindful of info's psychological impact

- Examples: CEOs announcing earnings forecasts to investors, leaders updating team members on recent developments, ...
- Main question: How does audience's psychological reaction to good and bad news shape the dynamic communication of info over time?
- Will focus on transmission of non-instrumental info, an under-studied topic
    - Abstract away from instrumental channel of info transmission, which is already well-studied and understood

The Environment

Informed sender communicates **non-instrumental info** over time

- Sender knows **state** that affects receiver's future consumption
- Different states change receiver's welfare, but not receiver's optimal actions — info irrelevant for decision-making
- Sender is benevolent and maximizes receiver's welfare
  - Cannot hide info forever: receiver learns state at moment of consumption in period T
  - Sender chooses how to communicate in periods $1, 2, ..., T-1$

Receiver Preference and Optimal Info Structure

Classical preference: indifferent between all info structures

Anticipatory utility: utility from level of belief

- e.g. Kőszegi (2006), Eliaz and Spiegler (2006)
- Optimal: sender only communicates in period 1

Suspense and surprise: utility from magnitude of belief change

- Ely, Frankel, Kamenica (2015)
- Optimal: a symmetric info structure, because model cannot talk about "good" or "bad" news

News utility with linear gains and losses

- Utility from changes in belief, Kőszegi and Rabin (2009)
  - In each period, good news elates while bad news disappoints
- A form of ref-dep preference (Kahneman and Tversky (1979))
  - Gain-loss utility based on deviation from reference point
  - Reference point = old rational expectation about consumption
- Optimal: one-shot resolution (provided **linear** gains/losses)

Key Assumption: Diminishing Sensitivity

Summary of this project:

- Incorporate **diminishing sensitivity** into news utility
- Novel, testable predictions about dynamic information transmission that seem less "stark" than existing models

What is DS and why should we care?

- Widely accepted, "textbook" assumption of ref-dep preferences since Kahneman and Tversky (1979)
- Larger deviations carry smaller marginal effects

But, almost all papers on ref-dep in past 40 years **assume away** DS for simplicity, instead use a two-part linear gain-loss utility

O'Donoghue and Sprenger (2018)'s survey of ref-dep models:

> *"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**."*

How Diminishing Sensitivity over News Matters

This project takes a first step in this direction, showing that DS is important in the domain of info design

Implications of news utility with DS that distinguish it from existing models discussed before:

1. DS leads to **more nuanced info preference.** Optimal info structure is generally not one-shot, but treats good and bad news asymmetrically.
2. DS leads to an **endogenous diversity** in informational preference: receivers with different consumption-based preferences over two states choose differently between gradual and one-shot news.
3. DS leads to **novel credibility problems** when the sender lacks commitment power.

▶ More Related Literature

Outline

1. Model: Sender's Problem and News Utility
2. Results with Commitment: DS and Info Preference
    - The optimal info structure
    - How consumption preference affects info preference
3. Results without Commitment: DS and Credibility Problem
    - When is the babbling equilibrium unique?
    - Equilibrium payoff and loss aversion
    - Rate of learning in equilibrium

1. Model (with Commitment)

Model: Timing of Events

Discrete time $t \in \{0, 1, ..., T\}$

Two players, sender and receiver, finite state space Θ

$\theta \in \Theta$ determines receiver's consumption c_θ in final period $t = T$

- No consumption in other periods, different c in different states
- Common prior $\pi_0 \in \Delta(\Theta), \pi_0(\theta) > 0$ for all θ
- θ is sender's private info, observed at start of $t = 1$
- Receiver exogenously learns θ at $t = T$

At $t = 0$, sender publicly commits to an **info structure** (M, σ)

- M – finite message space
- $h^t \in (M)^t$ – public history of messages at end of period t
- $\sigma = (\sigma_t)_{t=1}^{T-1}$, $\sigma_t(\cdot \mid \theta, h^{t-1}) \in \Delta(M)$ – how sender mixes over messages in period t, depending on state and public history

Model: Beliefs and Payoffs

Receiver knows (M, σ), forms Bayesian posterior belief $\pi_t \in \Delta(\Theta)$ after history h^t at end of t

Two sources of payoffs for receiver:

- Consumption utility
  - Get $v(c_\theta)$ in period T, v strictly increasing
  - Normalize without loss $0 \leq v(c_\theta) \leq 1$ for all θ
- News utility (NU)
  - NU function $N : \Delta(\Theta) \times \Delta(\Theta) \to \mathbb{R}$, continuous, $N(\pi \mid \pi) = 0$
  - Get $N(\pi_t \mid \pi_{t-1})$ at end of period $t \in \{1, ..., T\}$
  - Utility from recent change in belief

Sender wishes to maximize receiver's expected total welfare

- Non-instrumental info: no action affects consumption
- So, choose (M, σ) to maximize $\mathbb{E}_{(M,\sigma)} \left[\sum_{t=1}^{T} N(\pi_t \mid \pi_{t-1}) \right]$

Model: News Utility Function $N(\cdot \mid \cdot)$

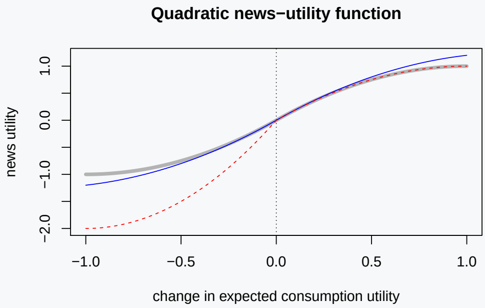
Focus on mean-based N for talk:

$$N(\pi_t \mid \pi_{t-1}) = \mu\left(\left[\sum_{\theta\in\Theta} \pi_t(\theta) v(c_\theta)\right] - \left[\sum_{\theta\in\Theta} \pi_{t-1}(\theta) v(c_\theta)\right]\right)$$

Throughout, assume gain-loss function $\mu : [-1, 1] \to \mathbb{R}$ satisfies

- $\mu(0) = 0$, μ strictly increasing
- **DS**: $\mu''(x) < 0$ for $x > 0$, $\mu''(x) > 0$ for $x < 0$
- **Weak loss aversion (LA)**: $-\mu(-x) \geq \mu(x)$ for all $x > 0$
  - Sometimes require **strict LA**: $-\mu(-x) > \mu(x)$ for all $x > 0$

Example: Quadratic μ



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

	α_p	α_n	β_p	β_n
Grey	2	2	1	1
Blue	2	2	0.8	0.8
Red	2	4	1	2

- $\alpha_n - \alpha_p$: loss aversion near 0; β_p, β_n: curvature, i.e. DS
- Add parameter restrictions to ensure monotonic, weak LA

Example: Power-Function μ

$$\mu(x) = \begin{cases} x^{\alpha} & x \geq 0 \\ -\lambda|x|^{\beta} & x < 0 \end{cases}$$

with $0 < \alpha, \beta < 1$ and $\lambda \geq 1$.

- α, β: DS to good and bad news; λ: LA
- Square-roots case when $\alpha = \beta = 0.5$
- The only class of gain-loss functions in Tversky and Kahneman (1992)

Cross-State Constraints

Suppose $\Theta = \{$**G**ood, **B**ad$\} = \{G, B\}$, $c_G > c_B$

Naive intuition about DS: **concentrate** all bad news in $t = 1$, but give **equally-sized** pieces of good news in $t = 1, 2, 3, \ldots$

These belief paths are infeasible!

- Bayesian receiver who knows this strategy and does not get bad news in $t = 1$ conclusively infers $\theta = G$
- Does not judge subsequent messages as further good news
- If $m \in M$ conveys positive but partial news in $t = 1$ when $\theta = G$, then also sent with positive prob. in $t = 1$ when $\theta = B$
- **Cross-state** constraints on information imply
  - Info structures that deliver news gradually must give false hope
  - Optimal info structure gives different amounts of good news in different periods, imperfect "consumption smoothing" of news

2(a). The Optimal Information Structure

An Example

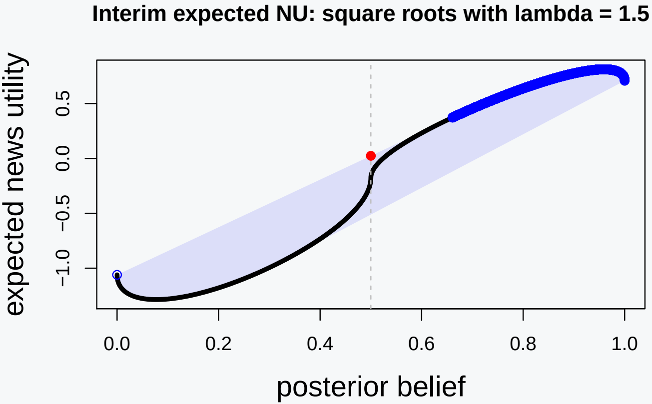
As an example, consider

- $T = 5$, $\Theta = \{G, B\}$, $\pi_0 = 0.5$, $v(c_G) = 1$, $v(c_B) = 0$
- $\mu(x) = \sqrt{x}$ for $x \geq 0$, $\mu(x) = -1.5\sqrt{-x}$ for $x < 0$

If $\pi_{T-2} = 0.5$, **interim expected NU** if update posterior to p:

$$U_{T-1}(p \mid 0.5) := \underbrace{\mu(p - 0.5)}_{\text{NU in } T-1} + \underbrace{p\mu(1-p) + (1-p)\mu(-p)}_{\text{expected NU in } T \text{ (by MG property)}}$$

Solving the Auxiliary Problem



- U **non-monotonic** due to μ DS
  - If μ two-part linear, then U monotonic
- Denote best expected NU following $\pi_{T-2} = 0.5$ by $U^*_{T-1}(0.5)$

Optimal Info Structure via Backward Induction

Once **value function** $U^*_{t+1}(\pi_t)$ defined for all $\pi_t \in \Delta(\Theta)$, let

$$U_t(\pi_t \mid \bar{\pi}_{t-1}) := \underbrace{\mu(\pi_t - \bar{\pi}_{t-1})}_{\text{NU in } t} + \underbrace{U^*_{t+1}(\pi_t)}_{\text{expected NU } \geq t+1}$$

$$U^*_t(\bar{\pi}_{t-1}) := \left[\mathrm{cav}\, U_t(\cdot \mid \bar{\pi}_{t-1})\right]_{\pi_t = \bar{\pi}_{t-1}}$$

for each $\bar{\pi}_{t-1} \in \Delta(\Theta)$ — the value function one period back

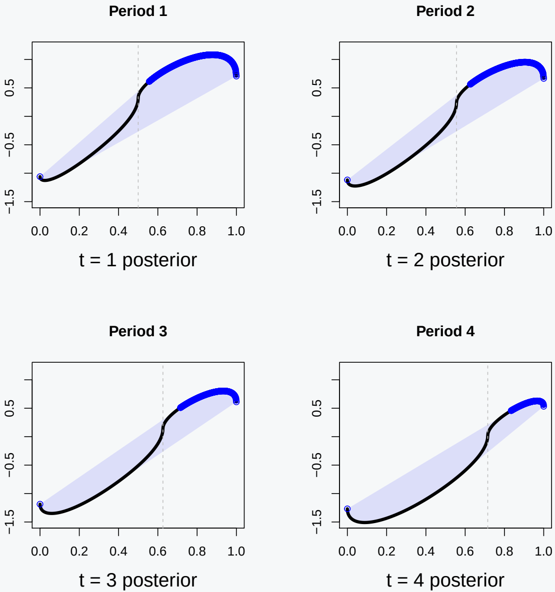
Proposition

The optimal news utility is $U^*_1(\pi_0)$.

Corollary

There exists an optimal info structure with $|M| = |\Theta|$ where a receiver entering period t with belief $\bar{\pi}_{t-1}$ always updates to a new belief π_t that supports concavification of $U_t(\cdot \mid \bar{\pi}_{t-1})$ at $\bar{\pi}_{t-1}$.

Optimal Info Structure in the Example



Optimal Info Structure in the Example

- If $\theta = B$, send b in a **random** period $\tilde{t}$ (or not at all)
  - In periods before $\tilde{t}$, send g
- If $\theta = G$, send g every period
- b = conclusive bad news, g = partial good news
- Optimal info structure features **gradual good news, one-shot bad news**

When $\theta = G$, belief grows at **increasing rates** across periods:

period	0	1	2	3	4	5
belief in $\theta = G$	0.500	0.556	0.626	0.715	0.834	1.000
change in belief		0.056	0.070	0.089	0.119	0.166

Uneven growth because of cross-state constraints

When Is One-Shot Resolution Suboptimal?

$v_0 := \mathbb{E}_{\theta\sim\pi_0}(v(c_\theta))$ — ex-ante expected future consumption utility

Proposition

Some info structure with "gradual good news, one-shot bad news" strictly dominates one-shot resolution of uncertainty if

$$\lim_{\epsilon\to 0^+} \frac{N(1_H \mid (1-\epsilon)1_H \oplus \epsilon 1_L)}{\epsilon} + N(1_H \mid \pi_0) - N(1_L \mid \pi_0)$$
$$> \lim_{\epsilon\to 0^+} \frac{N(1_H \mid \pi_0) - N((1-\epsilon)1_H \oplus \epsilon 1_L \mid \pi_0)}{\epsilon} - N(1_L \mid 1_H).$$

$\theta_H, \theta_L \in \Theta$ — *states with highest and lowest* c_θ
$1_H, 1_L \in \Delta(\Theta)$ — *degenerate beliefs in states* θ_H *and* θ_L

With mean-based news utility, condition says:

$$\underbrace{\mu(1-v_0) - \mu(-v_0)}_{\text{always positive}} + \underbrace{\mu'(0^+) - \mu'(1-v_0)}_{\text{positive by DS}} + \underbrace{\mu(-1)}_{\text{DS vs. LA}} > 0$$

Need "enough DS" (in positive or negative region) relative to LA

When Is One-Shot Resolution Suboptimal?

Example: quadratic news utility with $\alpha_n - \alpha_p \leq \beta_n + \beta_p$.

- LHS is size of kink at 0, $\mu^{'}(0^{-}) - \mu^{'}(0^{+})$
- RHS relates to $\mu^{''}(x)$ in negative and positive regions, i.e. DS
- So, condition holds if "sufficient DS relative to LA"

Example: any power-function news utility

- "infinite" DS near 0, as $\mu^{''}(0^{+}) = -\infty$ and $\mu^{''}(0^{-}) = \infty$

When Is One-Shot Resolution Suboptimal?

Kőszegi and Rabin (2009) study a percentile-based N:

$$N(\pi_t \mid \pi_{t-1}) = \int_0^1 \mu \left(v(F_{\pi_t}(p)) - v(F_{\pi_{t-1}}(p)) \right) dp$$

$F_{\pi_t}(p)$ and $F_{\pi_{t-1}}(p)$ are the p-th percentile consumption levels in beliefs π_t and π_{t-1}. Without DS, one-shot resolution is optimal.

Proposition

For Kőszegi and Rabin (2009)'s N, if $\mu''(x) < 0$ for all $x > 0$ and $|\Theta| \geq 3$, then one-shot resolution is strictly suboptimal.

Shows two-part linearity is crucial for Kőszegi and Rabin (2009)'s well-known one-shot optimality result – does not hold with any non-zero amount of DS in positive region.

Gradual Good News and Gradual Bad News

Specialize now to mean-based news utility and $\Theta = \{G, B\}$.

Definition

(M, σ) features **gradual good news, one-shot bad news** if:

- $\mathbb{P}_{(M,\sigma)}[\pi_t \geq \pi_{t-1}$ for all $1 \leq t \leq T \mid G] = 1$
- $\mathbb{P}_{(M,\sigma)}[\pi_t < \pi_{t-1}$ for at most one $1 \leq t \leq T \mid B] = 1$

There is **strictly gradual good news** if

- $\mathbb{P}_{(M,\sigma)}[\pi_t > \pi_{t-1}, \pi_{t'} > \pi_{t'-1}$ for two distinct $t \neq t' \mid G] > 0$

Similarly define **gradual bad news**.

Proposition

Any info structure featuring strictly gradual bad news, one-shot good news is strictly worse than one-shot resolution.

Gradual Good News and Gradual Bad News

Conversely, when does the **optimum** feature the "opposite skewness," i.e. gradual good news, one-shot bad news?

Here is a sufficient condition for $T = 2$.

Proposition

If the chord connecting $(0, U_{T-1}(0 \mid \pi_0))$ *and* $(\pi_0, U_{T-1}(\pi_0 \mid \pi_0))$ *lies strictly above* $U_{T-1}(p \mid \pi_0)$ *for all* $p \in (0, \pi_0)$, *then info structures with* $\mathbb{P}_{(M,\sigma)}[\pi_1 < \pi_0$ *and* $\pi_1 \neq 0] > 0$ *are strictly suboptimal.*

Can analytically prove the condition holds for quadratic μ

- Also holds (numerically) for μ power function, exponential, tanh, ...

For receiver with quadratic news utility with $\alpha_n - \alpha_p \leq \beta_n + \beta_p$, combining this result and sub-optimality of one-shot $\Rightarrow$ optimal info structure has strictly gradual good news, one-shot bad news

2(b)*. Consumption Preference and Info Preference

*(commitment case in a different environment)

Environment: Gradual vs One-Shot Info

New environment: sequence of exogenous signal realizations determine the state gradually

- $\Theta = \{A, B\}$, "**A**lternative" and "**B**aseline"
- In each $t = 1, 2, ..., T$, random variable $X_t \in \{0, 1\}$ realizes, $\mathbb{P}[X_t = 1] = q_t \in (0, 1)$, X_t's independent
- If $X_t = 1$ for **all** t, state is A. Else, state is B

At $t = 0$, agent commits to either observing realizations $(X_t)_{t=1}^T$ in real time, or only learning the final state in period T

- **Example**: debate between political candidates A and B, A loses as soon as she makes a "gaffe," but wins if no gaffe
  - Agent chooses between watching the debate or reading its outcome tomorrow
  - $c_A \gtrless c_B$ depends on agent's preference over candidates
- Assume there is $\lambda \geq 1$ s.t. $\mu(-x) = -\lambda\mu(x)$ for all $x > 0$

Endogenous Diversity of Information Preferences

Proposition

For any $\lambda \geq 1$, an agent who prefers state B will choose one-shot resolution of uncertainty. There exists some $\bar{\lambda} > 1$ so that for any $1 \leq \lambda \leq \bar{\lambda}$, an agent who prefers state A will choose gradual resolution of uncertainty.

- $c_A > c_B$: watch events unfold in real time to celebrate small victories
- $c_B > c_A$: only learn final state to avoid piecemeal bad news

Agents with opposite **consumption** preferences over A, B exhibit opposite **info** preferences. Testable prediction of the model.

Proposition

*These models do **not** predict different info choices for agents with opposite consumption rankings: (1) News utility with two-part linear μ; (2) Anticipatory utility; (3) Suspense and surprise.*

Application 1: Media Competition

- Suppose realization of some state A depends on a series of smaller events
- Some news outlets may cover these small events in detail
- Others only report final outcome
- Viewers sort between news sources based on $c_A \gtrless c_B$

Application 2: Game Shows

- 5 rounds, 1 contestant who will win either \$100,000 or nothing
- Empathetic audience gets news utility $\mu(\pi_t - \pi_{t-1})$ at the end of round t, $\pi_t =$ Bayesian prob. of contestant winning \$100k
- Format 1: "sudden death" — 5 easy rounds each with 87% winning probability, win \$100k if win **all** 5 rounds
- Format 2: "repêchage" — 5 hard rounds each with 13% winning probability, win \$100k by winning **any** round
- Proposition ⇒ audience whose LA is not too high will watch Format 1, but no one will watch Format 2
- Vast majority of game shows resemble Format 1
  - Even though two formats have same prob. of winning and same suspense and surprise utilities

3. DS and the Credibility Problem

Cheap Talk Model when Sender Lacks Commitment

- Let finite message space M be fixed with $|M| \geq 2$
- Sender **cannot commit** to strategy $\sigma_t(\cdot \mid \theta, h^{t-1})$
- Solve for **perfect-Bayesian eqm** under a belief refinement:
  - Once receiver updates belief to 0 or 1, belief stays constant through end of period $T-1$
  - In period T, belief $=$ full confidence in true state
- The babbling equilibrium always exists
- There is no other equilibrium payoff if μ exhibits DS and symmetry around origin

Proposition

When μ is symmetric around the origin and $\mu''(x) < 0$ for all $x > 0$, the babbling equilibrium is unique up to payoffs for any T.

Diminishing Sensitivity and the Credibility Problem

Consider period $T-1$

- When $\theta = B$, sender strictly prefers to lie and send a positive message intended for $\theta = G$
  - A trade-off for lying: positive news utility in period $T-1$ versus greater disappointment in $t = T$ due to false hope
  - **Due to DS**, trade-off favors lying
  - marginal utility of positive news today $>$ marginal disutility of heightened future disappointment
- So in eqm, sender cannot convey good news in **any** state
  - Else, tempted to mimic good-news message when $\theta = B$, destroying the news' credibility

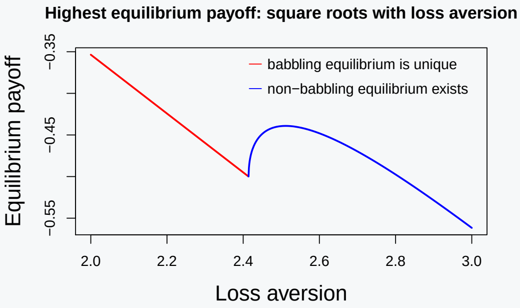
Sender must babble in period $T-1$

Now apply backward induction

This argument requires μ symmetric (i.e. $\lambda = 1$). Non-babbling equilibria can exist when LA high.

Loss Aversion and Receiver Welfare

Suppose $\mu(x) = \sqrt{x}$ for $x \geq 0$, $\mu(x) = -\lambda\sqrt{-x}$ for $x < 0$, $T = 2$



- Red: higher λ linearly decreases welfare (non-strategic effect)
- Blue: non-babbling eqm exists, λ additionally changes the eqm allocation of good news to different periods (eqm effect)
- For intermediate λ, eqm effect improves welfare by improving "consumption smoothing" of good news

GGN Equilibria

Definition

An equilibrium features deterministic gradual good news (**GGN equilibrium**) if there are constants $p_0 \leq p_1 \leq \ldots \leq p_T$ with $p_0 = \pi_0, p_T = 1$ s.t. $\pi_t = p_t$ in every t when $\theta = G$. The number of distinct beliefs in $(\pi_0, 1)$ is the number of **intermediate beliefs**.

Can characterize set of GGN equilibria using a relationship between successive intermediate beliefs.

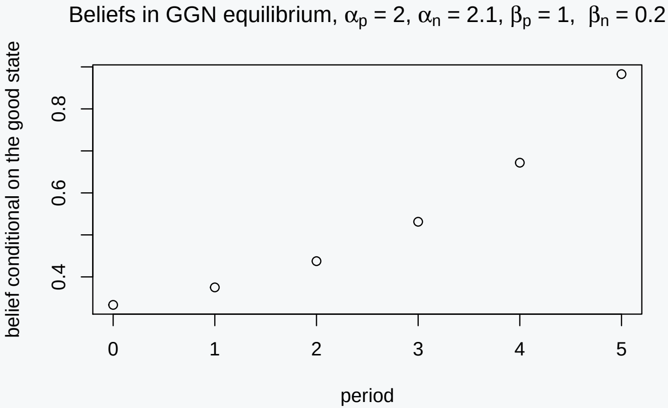
Proposition

There exists a GGN equilibrium with intermediate beliefs $q^{(1)} < \ldots < q^{(J)}$ if and only if $q^{(j)} \in P^*(q^{(j-1)})$ for all j (we put $q^{(0)} := \pi_0$), where

$$P^*(\pi) := \{\pi < x < 1 : \mu(x - \pi) + \mu(-x) = \mu(-\pi)\}.$$

GGN Equilibria for Quadratic μ

Example: For quadratic μ, $P^*(\pi) = \left\{\pi \cdot \frac{\beta_p+\beta_n}{\beta_p-\beta_n} - \frac{\alpha_n-\alpha_p}{\beta_p-\beta_n}\right\} \cap (\pi, 1)$.
GGN with multiple intermediate beliefs only exist for $\beta_n < \beta_p$.



Speed of Learning in GGN Equilibrium

Corollary

For μ quadratic or square roots, intermediate beliefs satisfy $q^{(j)} - q^{(j-1)} < q^{(j+1)} - q^{(j)}$ in every GGN equilibrium.

- Sender releases progressively larger pieces of good news
  - Intuition: when $\theta = B$, if sender indifferent between d amount of false hope and truth-telling at belief π_L, then strictly prefers d false hope over truth-telling at any $\pi_H > \pi_L$ due to DS
- Uniform bound on number of periods of informative communication, no matter how long the horizon T is
  - Bound is 5 for the quadratic example shown before
- More generally, "increasing rate of learning" holds whenever:
  - Sender indiff betw $\pi \to x \to 0$ and $\pi \to 0$ for at most one x
  - An ϵ amount of false hope better than truth-telling

Summary: Implications of Diminishing Sensitivity

Replying to O'Donoghue and Sprenger (2018), DS important b/c...

DS leads to **more nuanced info preference**

- The optimal info structure...
  - Generally treats good and bad news asymmetrically
  - Releases good news gradually, bad news in one-shot (under some conditions we identify)
  - Opposite "skewness" (i.e. gradual bad news) always worse than one-shot resolution
- Agents with opposite consumption preferences over states can exhibit opposite preferences between gradual and one-shot info

DS leads to **novel credibility problems** without commitment

- In bad state, marginal utility of false hope today $>$ marginal disutility of heightened future disappointment
- If low loss aversion, babbling is the unique equilibrium
- Receiver's equilibrium welfare non-monotonic in loss aversion
- In GGN equilibria, for quadratic and square-roots μ, receiver gets progressively larger pieces of good news over time

Thank you!

More Related Literature

- News utility without DS: Kőszegi and Rabin (2009), Pagel (2016, 2017, 2018), Duraj (2018 WP)
- Choosing beliefs without constraints (rather than info structure): Brunnermeier and Parker (2005), Macera (2014)
- Axioms of belief-based utility for $T = 2$: Dillenberger and Raymond (2018 WP)
- Bowman, Minehart, and Rabin (1999): ref-dep consumption model **with DS**
  - But, ref points are past habits, not rational expectations
  - All consumption loss in period 1, spread out consumption gains
  - Analogous strategy infeasible in sender's info design problem
- Lipnowski and Mathevet (2018) discuss an application to news utility without DS
  - We study information preference implications of DS, look at dynamics, relax commitment
- Li and Norman (2018 WP), Wu (2018 WP)
  - Group of senders move sequentially to persuade one receiver
  - Receiver acts once after all senders, only final belief matters
  - Our problem: final belief of receiver exogenous, but **stochastic process of interim beliefs** is the object of design

◀ Return