News Media as Suppliers of Narratives (and Information)*

Kfir Eliaz and Ran Spiegler†

January 19, 2024

Abstract

We present a model of news media that shape consumer beliefs by providing information (signals about an exogenous state) and narratives (models of what determines outcomes). To amplify consumers' engagement, media maximize consumers' anticipatory utility. Focusing on a class of separable consumer preferences, we show that a monopolistic media platform facing homogenous consumers provides a false "empowering" narrative coupled with an optimistically biased signal. Consumer heterogeneity gives rise to a novel menu-design problem due to a "data externality" among consumers. The optimal menu features multiple narratives and creates polarized beliefs. These effects also arise in a competitive media market model.

*Spiegler acknowledges financial support from Leverhulme Trust grant RPG-2023-120, "Modeling the Role of Narratives in Economics".

†Eliaz: School of Economics, Tel-Aviv University and David Eccles School of Business, University of Utah. E-mail: kfire@tauex.tau.ac.il. Spiegler: School of Economics, Tel-Aviv University and Economics Dept., University College London. E-mail: rani@tauex.tau.ac.il.

1 Introduction

Standard models of news media regard them as suppliers of information, providing noisy signals of an underlying state of Nature. A complementary view, which is absent from standard models, is that news media are a vehicle for spreading *narratives*, as reflected in the following quotes by prominent journalists:

> *"It's all storytelling, you know. That's what journalism is all about."* (Tom Brokaw)
>
> *"We're supposed to be tellers of tales as well as purveyors of facts. When we don't live up to that responsibility, we don't get read."* (William Blundell)

"Stories" or "narratives" are of course loaded terms with rich meanings in the context of news reporting. We conceive of narratives as models or frames that condition media consumers' thinking about the significance of reported information. For example, while many exogenous variables can be reported, the media often selects only some of them as relevant for outcomes of interest and therefore worthy of reporting. Another example is the shaping of popular perceptions about the role of personal agency and external factors in life outcomes. In particular, when reporting about discrimination, the media can peddle a narrative that focuses on the role of personal effort in achieving success. Alternatively, it can offer a narrative that attributes economic outcomes solely to discrimination; or a complex narrative that incorporates both factors.[1]

A related example involves narratives about financial investments. A popular narrative aimed at retail investors is that the value of their portfolio depends on how they manage it (how active are they? how do they allocate the portfolio between types of risk?). Such a narrative is misleading because it neglects the possibility that market fundamentals (interest rates, business sectors' risk or growth potential) are already reflected in security prices and thus affect investment returns. Thus, when the media report about market

[1] This example echoes Glenn Loury's (2020) distinction between "development" and "bias" narratives.

fundamentals (e.g., that certain sectors are experiencing growth), the simplistic narrative may lead media consumers to draw wrong conclusions from this information (e.g., that they should invest in a growth sector).

This paper presents a model of news media (in a broad sense that includes content platforms) that is based on a fusion of the two views: The media provides information about exogenous states as well as a narrative, which is a model of the determination of outcomes as a function of states and actions. Media consumers use the narrative to interpret statistical regularities, forming beliefs about the mapping from states and actions to outcomes. A false narrative is a misspecified model, which can therefore induce distorted beliefs.

The fusion of the information-based and narrative-based views enables us to offer a new model of media bias. There is a common intuition that the source of this phenomenon is some aspect of consumer demand (Gentzkow and Shapiro (2010) back this intuition with empirical evidence). Yet, the standard model of consumer behavior assumes that demand for information is purely instrumental. Expected-utility maximizers weakly prefer more informative signals. Therefore, unless there are frictions on the supply side that prevent media from providing complete and objective information, the market will provide it. Even if consumers have heterogeneous preferences, they all want more informative news.

This observation has inspired models of media bias in which demand for news is not purely instrumental (see Prat and Strömberg (2013) and Gentzkow et al. (2015) for surveys). We propose that media consumers have intrinsic preferences over *posterior* beliefs, which they arrive at via Bayesian updating of prior beliefs shaped by the narrative they adopt. To our knowledge, both the role of narratives in belief formation and the integration of hedonic and Bayesian aspects of consumer beliefs are new to the literature on media bias. We discuss the relation to the literature in further detail in Section 6.

One can distinguish between two types of preferences over posterior beliefs. *Retrospective* preferences rank beliefs about past outcomes — e.g., wanting to believe that you are the just side in a dispute (see Chopra et al. (2023)). *Prospective* preferences rank beliefs about future outcomes, taking into account private or public actions in response to beliefs. We focus on the latter type: Consumers approach news media with the desire to maximize their *anticipatory utility* — i.e., the expected indirect utility from their posterior beliefs.

According to this view, demand for news is driven by consumers' pursuit of optimistic posterior beliefs. By "optimism", we do not necessarily mean that the media paint a rosy picture of reality, but rather that it gives hope which may depend on one's actions. Real-life manifestations of this idea include business news channels conveying the impression that retail investors can beat the market; patriotic coverage of international conflicts or sporting events that amplify the prospect of victory; and reports of police brutality or climate change which send a message that policy reforms ("defunding the police", switching to alternative energy sources) can improve social welfare.

However, under the pure information-based model of news media, prospective preferences over (Bayesian posterior) beliefs cannot give rise to media bias. The reason is that expected anticipatory utility is the upper envelope of linear functions of posterior beliefs (hence convex in these beliefs). As a result, when the media caters to consumer demand by offering a signal function that maximizes consumers' anticipatory utility, it will weakly prefer full information provision. Thus, even when we allow for prospective non-instrumental demand for information, the standard view of the media as information providers cannot generate media bias.

This is where our view of media as joint providers of narratives and information enters. We show that this more comprehensive approach provides a non-trivial model of media bias, such that distortion of the truth consists of biased/inaccurate reports *together with* false narratives. Moreover, there is a synergy between these two instruments: They complement each other in producing the distorted, optimistic beliefs that consumers seek.

Overview of model and results

In the basic version of our model, a representative consumer takes an action after observing a signal about a state of Nature. There is an objective stochastic mapping from states and actions to outcomes. The consumer is endowed with a vNM utility function which is separable in the action. A monopolistic media outlet commits ex-ante to a "media strategy", which consists of: (i) a Blackwell experiment (namely, a stochastic mapping from states to signals), and (ii) a narrative, which selects a subset of the outcome's true causes.

There are four feasible narratives. The true narrative acknowledges both states and actions as causes. The "empowering" narrative postulates that actions are the sole cause of outcomes (in terms of the discrimination example above, this narrative says that only personal effort determines economic outcomes, thus suppressing the role of discrimination). The "fatalistic" narrative postulates that only the state matters for the outcome (e.g., it says that only the external force of discrimination determines outcomes, thus suppressing the role of personal agency). Finally, the "denial" narrative removes both the action and the state as causes, thus implicitly attributing outcomes to unspecified other factors.

Given an empirical long-run distribution over states, actions and outcomes, a narrative produces a subjective (and possibly distorted) belief by "fitting" the narrative to this distribution. For example, the empowering narrative interprets the empirical correlation between actions and outcomes as a causal quantity — i.e., it attributes the variation in outcomes entirely to variation in actions. Once the consumer adopts a narrative, his strategy (a stochastic mapping from signals to actions) prescribes actions that maximize expected utility with respect to the narrative-induced belief. In equilibrium, this strategy is consistent with the empirical long-run distribution. The need for an equilibrium definition of consumer response to a given narrative is typical of models of decision making under misspecified models (e.g., Esponda and Pouzo (2016), Spiegler (2016), Eliaz and Spiegler (2020)). The reason is that changes in long-run behavior can lead to changes in the consumer's perceived mapping from actions to outcomes.

The media's problem is to find a strategy and an equilibrium (induced by the strategy) that maximize the consumer's ex-ante expected anticipatory utility. The rationale for maximizing anticipatory utility is that a consumer's engagement with the media increases with the amount of optimism he can derive from its consumption. The better the media performs in generating optimistic beliefs, the higher the demand for it. A crucial feature of the problem is that the media takes into account equilibrium effects when designing its strategy. This is similar in spirit to information-design problems (e.g., Bergemann and Morris (2019)). However, in standard models equilibrium effects arise in multi-agent settings with payoff externalities. In contrast, in the present model equilibrium effects arise because of misspecified beliefs induced by false narratives.

This account of news media raises a number of questions: Will the media provide accurate, unbiased information? If not, what is the structure of media inaccuracy/bias, and which narratives will it peddle? Our analysis of the baseline model in Section 3 addresses these questions. We begin with a full characterization of the optimal media strategy in a specification of our model — inspired by the above discrimination and financial-investment scenarios — which serves as a running example in our paper. The optimal strategy consists of the empowering narrative and a signal with an optimistic bias (i.e., always correctly reporting good news and sometimes misrepresenting bad news). The magnitude of the bias is tailored to consumer preferences.

We then show that this combination is a robust feature: For any action-separable utility function, if a media strategy outperforms the true-narrative, perfect-information benchmark, then it must involve the empowering narrative. Also, it must provide information that induces different behavior from the benchmark (as long as the benchmark leads to state-contingent actions). Thus, *there is a synergy between false narratives and biased information, which is essential for the media's mission to maximize consumers' anticipatory utility.* This result also demonstrates the value of our model in making specific predictions about the structure of media narratives and media bias.

Section 4 introduces preference heterogeneity among consumers in the context of our running example. This naturally calls for a model in which consumers can choose between multiple narrative-information combinations. We now envisage our monopolistic media provider as a gatekeeper or a platform that restricts the entry of these combinations. Effectively, this means that the media chooses a menu of media strategies, aiming to maximize aggregate anticipatory utility.[2] From the menu, each consumer chooses the narrative-information pair that maximizes his own anticipatory utility, evaluated according to the equilibrium joint distribution over states, actions and outcomes that arises from the aggregate behavior of *all* consumer types.

At first glance, it may appear that incentive-compatibility should be moot in this model, because media and consumers have a common objective. However, this is not the case because of the "*data externality*" that exists among consumer types. Although they are separate individuals with idiosyncratic preferences, they all rely on the same aggregate data to form beliefs given the narratives they adopt. Consequently, changes in the behavior of one segment of the consumer population can change how another segment evaluates narrative-information pairs. Dealing with this externality in the context of a

[2] For tractability, we restrict media strategies to report good news in the good state, which is an endogenous feature of the optimal strategy in the baseline model.

menu design problem is a methodological novelty of our paper, and one of our motivations for introducing heterogeneity in the first place.

The data externality turns out to have a significant effect on the optimal menu, compared with the representative-consumer case. The contrast is particularly stark when consumer types are uniformly distributed. Instead of biased, partially informative signals that are finely tailored to consumer types, now none of the consumers receive *any* information. Nevertheless, the population is split into two "camps" with starkly different beliefs, driven by different narratives they consume. One segment opts for the empowering narrative and takes one constant action, while the other segment opts for the denial narrative and always takes the opposite action.[3] Thus, our model shows how a heterogeneous population of consumers trying to make sense of the same aggregate data can end up holding highly polarized beliefs based on no information, simply because of the narrative-peddling aspect of media strategy.

We then explore the role of market structure by examining a "perfect competition" version of the heterogeneous-consumers model. Each media provider is "small" in the sense that it takes the joint distribution over states, actions and outcomes as given, without taking into account how its media strategy affects this distribution via its effect on consumers' beliefs and actions. A "competitive equilibrium" is a profile of media strategies, one for each consumer type, such that: (i) the strategy associated with a type maximizes his anticipatory utility given the aggregate distribution; and (ii) this distribution arises from each type best-replying to the belief induced by the media strategy he adopts. We show that in the essentially unique equilibrium, only the true and fatalistic narratives prevail, where the former narrative is coupled with complete information. When consumer types are uniformly distributed, the segment of consumers who act on informative signals is larger than in the monopolistic case. Thus, while perfect competition is worse than monopoly in terms of consumers' anticipatory utility (because media providers fail to incorporate the data externality), it provides more accurate information, even though it does not eradicate wrong beliefs due to false narratives.

[3] The consumers who select the denial narrative could receive information, but that would have no effect on their behavior or anticipatory utility. For the consumers who adopt the empowering narrative, the lack of any information is strictly optimal.

In Section 5 we return to the baseline model and extend it in two directions. Our first extension mixes the consumer population with rational consumers, who know the true model and whose demand for information is conventionally instrumental. This extension introduces a screening problem along an unconventional dimension: consumers' willingness to adopt a false hopeful narrative. We show that it is optimal for the media to provide a *singleton* menu. When there are few rational consumers, this menu consists of the empowering narrative and a biased signal (albeit less biased than in the baseline model). When there are few non-rational consumers, the menu consists of the fatalistic narrative and a fully informative signal.

Next, we consider different separable specifications of the consumer's utility function. When it is separable in the *state*, the only false narrative that can outperform the true narrative is the *fatalistic* narrative. We illustrate this finding with an example in which actions have objective "unintended consequences" that the false narrative neglects. When the utility function is separable in the outcome, the true narrative, coupled with complete information, is optimal.

The extensions cement the main insight of our paper: When media demand is driven by motivated reasoning, peddling false narratives is a key feature of media bias.

2 A Model

We begin by introducing the primitives of our model. There are four relevant variables: a state of Nature t, an action a taken by a representative consumer, a signal s that the consumer observes before taking the action (he can only condition his action on s), and an outcome y. All variables take finitely many values. The consumer's vNM utility takes the form:

$$u(t, a, y) = v(t, y) - c(a)$$

The objective data-generating process is a probability distribution p defined over the four variables, which can be factorized as follows:

$$p(t, s, a, y) = p(t)p(s \mid t)p(a \mid s)p(y \mid t, a) \tag{1}$$

The first and last terms on the R.H.S are exogenously given; they describe the prior distribution of the state of Nature, and the outcome distribution conditional on the state and the consumer's action, respectively. Note that the signal has no direct effect on the outcome. The term $p(s \mid t)$ describes the signal distribution conditional on the state. This distribution is determined ex-ante by a monopolistic media outlet. Finally, the term $p(a \mid s)$ represents the consumer's strategy (i.e., his action distribution conditional on the signal). The strategy's endogenous determination will be described below.

The causal structure underlying this data-generating process can be described by the following directed acyclic graph (DAG), denoted N^*:

$$\begin{array}{ccc} t & \rightarrow & s \\ \downarrow & & \downarrow \\ y & \leftarrow & a \end{array}$$

In this graphical representation, borrowed from the Statistics/AI literature on probabilistic graphical models (Pearl (2009)), a node represents a variable, and an arrow represents a direct causal relation (e.g., s is the only direct cause of a).

Let us now describe the interaction between the media and the representative consumer. The media moves first, committing ex-ante to a pair (I, N), where: I is a signal function, which is a Blackwell experiment assigning a distribution over signals to each state (this is the conditional probability distribution $(p(s \mid t))_{t,s}$ described above); and N is a narrative, which is a causal model represented by a DAG over the set of nodes representing the four variables t, s, a, y.

We restrict N to be one of the following possible DAGs, in addition to N^*:

$$\begin{array}{ccc} t & \rightarrow & s \\ & & \downarrow \\ y & \leftarrow & a \end{array} \qquad \begin{array}{ccc} t & \rightarrow & s \\ \downarrow & & \downarrow \\ y & & a \end{array} \qquad \begin{array}{ccc} t & \rightarrow & s \\ & & \downarrow \\ y & & a \end{array}$$

N^a (empowering) $\quad$ N^t (fatalistic) $\quad$ $N^{\emptyset}$ (denial)

All narratives represent causal models that coincide with the true model N^* in how they depict the causal relations among t, s, and a. That is, the narratives correctly account for the causal chain that leads from the state to the action via the signal. The only difference between the three narratives and N^* is in their account of the direct causes of the outcome y. The "empowering" narrative N^a fully attributes the outcome to actions. The "fatalistic" narrative N^t fully attributes the outcome to the exogenous state. The "denial" narrative $N^{\emptyset}$ suppresses the causal effect of both a and t, thus implicitly attributing the outcome to other, unspecified exogenous factors. These three narratives are false in the sense that they remove at least one of the causal links into y.

Given an objective joint distribution p, each of the false narratives induces a subjective belief over the four variables, given as follows:

$$p_{N^x}(t, s, a, y) = p(t)p(s \mid t)p(a \mid s)p(y \mid x) \tag{2}$$

where x is a stand-in for either a, t or $\emptyset$ (that is, $p(y \mid x) = p_{N^x}(y \mid t, a)$). Under this notation, $p(y \mid \emptyset)$ is simply $p(y)$. Note that p_{N^*} coincides with p, given by (1).

The interpretation is that p represents a long-run empirical distribution (reflecting the decisions of other consumers who faced the same problem); and the narrative N makes sense of this distribution by imposing a particular explanation for what causes variation in outcomes. The distribution p_N is a systematic distortion of the objective distribution p through the prism of the causal model N.

Thus, the media affects the consumer's beliefs via two channels: (i) the signal function given by I, which determines the reporting of current events; and (ii) the narrative N, which provides a perspective — based on an interpretation of historical regularities — for drawing implications from the signal.

As long as the terms in (2) do not involve conditioning on zero-probability events, p_{N^x} is a well-defined probability distribution, which induces the following conditional belief:

$$p_{N^x}(t, y \mid s, a) = p(t \mid s)p(y \mid x) \tag{3}$$

This conditional probability is not invariant to the long-run consumer average behavior given by $(p(a \mid s))_{a,s}$. To see why, elaborate $p(y \mid x)$ in (3) for each of the three false narratives:

$$p(y \mid a) = \sum_{t'} p(t' \mid a)p(y \mid t', a) = \sum_{s',t'} p(s' \mid a)p(t' \mid s')p(y \mid t', a) \tag{4}$$

$$p(y \mid t) = \sum_{a'} p(a' \mid t) p(y \mid t, a') = \sum_{s', a'} p(s' \mid t) p(a' \mid s') p(y \mid t, a') \tag{5}$$

$$p(y) = \sum_{t'} p(t') \sum_{a'} p(a' \mid t') p(y \mid t', a') = \sum_{t'} p(t') \sum_{s', a'} p(s' \mid t') p(a' \mid s') p(y \mid t', a') \tag{6}$$

It is evident that the terms $p(s' \mid a)$ and $p(a' \mid s')$ involve the consumer's strategy. In other words, long-run consumer behavior affects narrative-based perception of the mapping from actions to consequences (given a signal), which in turn affects the consumer's subjectively optimal decisions. Thus, if we view the long-run distribution p as a *steady state*, we need an equilibrium notion of the consumer's subjective optimization.

Definition 1 (Equilibrium) *Given (I, N), a strategy $(p(a \mid s))_{a,s}$ is an ε-equilibrium if, whenever $p(a \mid s) > \varepsilon$, a maximizes*

$$V_{I,N}(s, a) = \sum_{t,y} p_N(t, y \mid s, a) u(t, a, y) \tag{7}$$

A strategy is an equilibrium if it is a limit of a sequence of ε-equilibria, where $\varepsilon \to 0$.

This is the definition of personal equilibrium in Spiegler (2016), and it coincides with Berk-Nash equilibrium (Esponda and Pouzo (2016)) when the consumer's subjective model is given by N. The role of trembles in this definition is merely to avoid conditioning on null events. Trembles play no meaningful role in our analysis.

We assume that the media chooses (I, N) ex-ante to maximize

$$U(I, N) = \sum_{t} p(t) \sum_{s} p(s \mid t) \max_{a} V_{I,N}(s, a) \tag{8}$$

subject to the constraint that $(p(a \mid s))_{a,s}$ is an equilibrium. The media's objective function is the consumer's *expected anticipatory utility.* The idea behind this objective function is that anticipatory utility generates the consumer's demand for news media. The higher his anticipatory utility, the greater his media engagement. Our task is to characterize the media's optimal strategy.

In solving its problem, the media takes into account the consumer's equilibrium response to the media strategy. This naturally raises the question of whether the media knowingly anticipates equilibrium effects. One interpretation is that the media is not aware of them a priori. Instead, it reacts to past data about consumer engagement, possibly using algorithmic learning. The equilibrium effects that shape consumers' media engagement will be reflected in the learning process.

The necessity of false narratives for media bias
Suppose that the media is restricted to providing the true narrative N^*. This reduces the model to standard information provision by a sender who can commit ex-ante to a Blackwell experiment. The sender faces a Bayesian receiver whose indirect utility from a posterior belief μ over t is

$$\max_a \sum_t \mu(t) \sum_y p(y \mid t, a) u(t, a, y)$$

Since this indirect utility is a maximum over functions that are linear in μ, it is convex in μ. Therefore, it is (weakly) optimal for the sender to commit to a fully informative signal — i.e., $p(s = t \mid t) = 1$ for every t. It follows that in our model, given the media's objective of maximizing the consumer's ex-ante anticipatory utility, the media has no strict incentive to provide partial or biased information unless it also peddles a false narrative. Throughout the paper, we refer to the maximal anticipatory utility attained by the true narrative and complete information as the *rational-expectations benchmark*.

Comment on the interpretation of a and y
According to one interpretation of our model, a represents a *private* action that an individual media consumer takes, and y is a personal outcome of his choice. For example, a can represent the agent's career decision or a dietary choice, in which case y represents his earnings or health outcome, respectively. The data that the consumer relies on to form beliefs (via the factorization according to N) is *aggregate*, reflecting the historical choices and outcomes of other consumers.

An alternative interpretation is that a represents a *public* choice (such as economic or foreign policy), and y represents a public outcome (economic growth, national security). According to this interpretation, the media consumer is a representative *voter*, and the probability $p(a \mid s)$ is the frequency with which society selects a political leadership that implements a. This is a reduced-form representation of a democratic process, such that society's choice matches what the representative voter deems optimal.

Our model departs from the canonical information-design framework (see Bergemann and Morris (2019)), since it allows the designer to influence the subjective model that the receiver holds. Nevertheless, the assumption that the consumer always correctly perceives $p(t, s, a)$ ensures that the standard revelation principle in the information-design literature can be adapted to the present setting.

Remark 1 (A revelation principle) *Without loss of optimality, we can restrict attention to signal functions that assign a distribution over recommended actions to each state, and to equilibria in which $a = s$ with probability one for each s.*

The proof of this remark follows the footsteps of Theorem 1 in Bergemann and Morris (2016) — adapted to the single-player setting — and is therefore omitted. The proof involves manipulating the signal function given by $(p(s \mid t))_{t,s}$ and the consumer's strategy given by $(p(a \mid s))_{a,s}$. In general, when the consumer forms beliefs according to a misspecified model N, such changes may affect $p_N(y \mid t, a)$, which could violate the revelation principle. The reason the principle does hold in our setting is that the manipulation of $(p(s \mid t))_{t,s}$ and $(p(a \mid s))_{a,s}$ in the proof leaves $(p(t, a))_{t,a}$ unchanged. As evident from (4)-(6), this means that $p_N(t, a)$ and $p_N(y \mid t, a)$ also remain unchanged, regardless of how t and a are jointly distributed with s. This enables the standard proof to go through. The revelation principle will simplify our analysis in the sequel.

3 Analysis

In this section we analyze the media's optimal strategy. We begin with a specification that serves as a running example in the paper. We then show that the qualitative features of the optimal media strategy in our example are robust.

3.1 An Example: Investment Narratives

In this example, all variables take values in $\{0, 1\}$. The exogenous components of the data-generating process are:

$$p(t = 1) = \frac{1}{2} \qquad\qquad p(y = 1 \mid t, a) = \frac{1}{2}a\,(2 - t)$$

The consumer's payoff function is

$$u(a, t, y) = ty - ca$$

where $c \in (0, \frac{1}{2})$. (We will later handle the case of $c > \frac{1}{2}$.) The action a represents a private decision whether to initiate a costly economic activity. The outcome y indicates whether the activity is successful. The state t represents the return from a successful outcome. High returns are associated with lower chances of a successful outcome.

For a specific example, the consumer is a retail financial investor choosing whether to engage in active or passive portfolio management. The realization $y = 1$ represents an increase in the portfolio's value. The state t represents an effective discount factor that reflects economic fundamentals such as the interest rate for alternative investments, the expected timing of a successful outcome, or macroeconomic uncertainty. The negative correlation between t and y has a natural interpretation in this context. Market fundamentals tend to be reflected in current security prices. For instance, a low-risk environment is reflected in high current prices, lowering the prospects of an increase in the portfolio's value.

An alternative story is that the consumer is a college student who decides how seriously to take his studies. A successful outcome means graduating (rather than dropping out). The realization $t = 1$ represents a college wage premium. The negative correlation between t and y is due to the fact that as returns to high education rise, colleges respond by becoming more demanding, such that graduating becomes less likely.

Under both stories, the media provides information about the fundamentals represented by t, as well as a narrative about what drives the outcome y. Our task is to characterize the optimal media strategy, considering each of the feasible narratives.

Rational-expectations benchmark

Suppose the media offers the true narrative N^*. As we saw, it is optimal to couple this narrative with a fully informative signal. When $t = 0$, the consumer knows that $ty = 0$, and therefore plays $a = 0$. When $t = 1$, he knows that $p(y = 1 \mid t = 1, a = 1) = \frac{1}{2}$. Since $c < \frac{1}{2}$, the consumer plays $a = 1$. It follows that the rational-expectations benchmark in this example is $\frac{1}{4} - \frac{1}{2}c$.

Narratives that omit the link $a \to y$

Under the narratives N^t and $N^{\emptyset}$, the consumer believes that his action has no effect on y, and therefore prefers to take the costless action $a = 0$. In any equilibrium, $a = 0$ with certainty for every t. However, since $y = 0$ whenever $a = 0$, it follows that $p(y = 1 \mid t) = p(y = 1) = 0$ for every t. Therefore, the consumer's anticipatory utility is necessarily zero, which is below the rational-expectations benchmark. It follows that the media will necessarily offer a narrative that retains the link $a \to y$.

The empowering narrative

Under the narrative N^a,

$$p_{N^a}(ty = 1 \mid a, s) = p(t = 1 \mid s)p(y = 1 \mid a)$$

Since $p(y = 1 \mid a = 0) = 0$, the consumer's subjective payoff from $a = 0$ is zero for every signal he receives. Let us calculate the consumer's subjective expected payoff from $a = 1$. Denote $q_t = p(s = 1 \mid t)$. By the revelation principle, we can restrict attention to binary signals and an equilibrium in which the consumer always plays $a = s$. Then,

$$p(y = 1 \mid a = 1) = \frac{1}{2} + \frac{1}{2}p(t = 0 \mid a = 1) = \frac{1}{2} + \frac{1}{2} \cdot \frac{q_0}{q_1 + q_0}$$

Therefore,

$$p_{N^a}(ty = 1 \mid s = 1, a = 1) = \frac{q_1}{q_1 + q_0} \cdot \left(\frac{1}{2} + \frac{1}{2} \cdot \frac{q_0}{q_1 + q_0}\right) \tag{9}$$

and

$$p_{N^a}(ty = 1 \mid s = 0, a = 1) = \frac{1 - q_1}{2 - q_1 - q_0} \cdot \left(\frac{1}{2} + \frac{1}{2} \cdot \frac{q_0}{q_1 + q_0}\right) \tag{10}$$

In order for the consumer's strategy to be an equilibrium, we need (9) and

(10) to be weakly above and below c, respectively. If these constraints hold, the consumer's anticipatory utility is $p(s=1)\cdot[p_{N^a}(ty=1 \mid s=1, a=1)-c]$, given by

$$\frac{q_1+q_0}{2}\cdot\left[\frac{q_1}{q_1+q_0}\cdot\left(\frac{1}{2}+\frac{1}{2}\cdot\frac{q_0}{q_1+q_0}\right)-c\right] \tag{11}$$

Observe that when the media offers a fully informative signal ($q_1=1$, $q_0=0$), this expression coincides with the payoff from N^*. Thus, if the false narrative N^a is to outperform the true narrative, it must be coupled with incomplete information. We now proceed to calculate the optimal $I=(q_0,q_1)$ that accompanies N^a.

Claim 1 *When $c<\frac{1}{2}$, it is optimal to set $q_1=1$.*

Thus, if the optimal signal function has a bias, it must be an optimistic one, as the media always reports good news ($s=1$) when the state is good ($t=1$). The proof of this claim (like all proofs in this paper) is in the Appendix. The claim implies the following simplified expression for the consumer's anticipatory utility:

$$\frac{1}{2}\left[\frac{1}{2}+\frac{1}{2}\cdot\frac{q_0}{1+q_0}-c(1+q_0)\right] \tag{12}$$

Note that $q_1=1$ also implies that (10) is below c, such that playing $a=0$ when $s=0$ is optimal for the consumer. It is now straightforward to derive the optimal value of q_0:

$$q_0=\min\left\{1,\sqrt{\frac{1}{2c}}-1\right\} \tag{13}$$

Plugging (13) in (12), the consumer's ex-ante anticipatory payoff is

$$\begin{array}{ccc}\frac{1}{2}-\sqrt{\frac{c}{2}} & if & c\in\left[\frac{1}{8},\frac{1}{2}\right) \\ \frac{3}{8}-c & if & c\in\left(0,\frac{1}{8}\right)\end{array}$$

which exceeds the rational-expectations benchmark.

Thus, when $c<\frac{1}{2}$, the optimal media strategy involves the narrative N^a coupled with positively biased information: always sending a good signal in the good state, and sending it with positive probability in the bad state.

In terms of the interpretation we offered for this example, the false narrative N^a attributes a successful investment outcome entirely to the consumer's actions, without taking into account the role of the fundamentals. The accompanying signal function has an optimistic bias in the direction of claiming that returns are high even when they are not. Thus, on one hand the media exaggerates the attractiveness of the investment environment, while on the other hand it suppresses — via the empowering narrative — the negative effect that good fundamentals have on the chances of good investment outcomes.

So far, we assumed that $c < \frac{1}{2}$. It is easy to see that when $c \geq \frac{1}{2}$, none of the feasible narratives can generate positive utility. As we already saw, the narratives N^t and $N^\emptyset$ generate zero utility for every c. It is also immediate from the expressions for the anticipatory utility induced by N^* and N^a that when $c \geq \frac{1}{2}$, these narratives cannot generate positive payoffs.

3.2 A Characterization Result

The investment-narrative example has two noteworthy features. First, the empowering narrative emerges as optimal. Second, it is coupled with biased information that impacts consumer behavior. We now show that both features hold generally under the model of Section 2.

Proposition 1 *If the media can outperform the rational-expectations benchmark, then N^a is part of an optimal strategy.*

Thus, the empowering narrative N^a is an essential feature of a media strategy that outperforms the rational-expectations benchmark. The logic behind the result is as follows. Because u is action-separable, a false narrative can have an effect on ex-ante anticipatory utility only when it distorts the joint distribution of (t, y). By definition, the fatalistic narrative N^t cannot do that. In principle, the denial narrative $N^\emptyset$ can attain such a distortion. However, this effect can be replicated by N^a coupled with no information.

The next result addresses the consumer behavior that the optimal media strategy induces. We say that the payoff function and the exogenous data-generating process form a *regular environment* if, under the true narrative and complete information, the consumer has a unique best-reply which is a one-to-one function of the state. That is, in regular environments different states prescribe different unique actions under rational expectations.

Proposition 2 *Suppose the environment is regular. If the optimal media strategy outperforms the rational-expectations benchmark, then its induced conditional distribution* $(p(a \mid t))_{t,a}$ *is different from that benchmark.*

Thus, when the media deviates from the rational-expectations benchmark, it necessarily induces changes in consumer behavior. Since regularity assumes a unique optimal action in each state (under rational expectations), this means that the outcome induced by the media's strategy departs from what a paternalistic social planner (aiming to maximize consumers' material payoffs) would prescribe.

Note that our result does not claim that the media necessarily departs from fully informative signals. We cannot rule out the possibility that the media sends a fully informative signal in every state and that the consumer's subjective best-reply involves mixing, which will be sustained in equilibrium thanks to the false narrative N^a.

Regularity plays a key role in the result. To see why, consider the payoff specification of Section 3.1, and let the data-generating process satisfy $p(t=1)=\frac{1}{2}$ and $p(y=1 \mid t,a)=1-t$ for every t,a. Under rational expectations, the consumer's optimal action is $a=0$ for every t, and the rational-expectations payoff is 0 (because $a=0$ and $ty=0$ with probability one). Using similar arguments as in Section 3.1, it can be shown that it is optimal for the media to provide $N^{\emptyset}$ (or, equivalently, N^a) and no information. The consumer responds by playing $a=0$. His anticipatory payoff is $\frac{1}{4}$, beating the rational-expectations benchmark, although the behavior is the same. Thus, without regularity, it is possible for the media strategy to outperform the benchmark without any effect on consumer behavior.

4 Heterogeneous Consumers

In this section we revisit the investment example of Section 3.1, and extend the model by introducing consumer heterogeneity. Specifically, we assume that the cost parameter c is distributed over $[0,1]$ in the consumer population according to some continuous and strictly increasing *cdf* F. Given that the optimal media strategy in Section 3.1 varied with c, it is natural to consider media markets with a supply of multiple pairs (I, N).

We consider two market structures. In Section 4.1, we consider a monopolistic media platform acting as a gatekeeper who restricts the entry of media providers (each represented by a distinct pair (I, N)). The monopolist's objective is to maximize consumers' aggregate anticipatory utility — reflecting the continued assumption that this corresponds to maximizing their platform engagement. In Section 4.2, we remove the gatekeeper and analyze a "perfectly competitive" media market, in which each provider targets a particular consumer type and tries to maximize his anticipatory utility.

This is not a mechanical extension from homogenous to heterogeneous consumer populations. The reason is that while each consumer type maximizes his own anticipatory utility, this utility — shaped by the narrative he adopts — is evaluated according to the joint distribution over actions and outcomes, which reflects the *aggregate* behavior of all consumers. In other words, when consumers adopt a false narrative, they exert a "data externality" on one another, because the choices of one segment of consumer types can affect the belief formed by another segment. This externality requires methodological innovations when dealing with the monopolistic and competitive market structures. The difference between the two market structures is that the monopolist is an "externality maker" (who internalizes the data externality) while competitive media providers are "externality takers". This leads to qualitatively different characterizations of media strategies that emerge in heterogeneous markets.

4.1 Monopoly

In this version of the model, the monopolist commits ex-ante to a menu M of pairs (I, N). We assume that signals are binary, $s \in \{0, 1\}$, and restrict the set of feasible signal functions such that $\Pr(s = 1 \mid t = 1) = 1$. Unlike the homogenous case, the latter restriction entails a loss of generality in the present setting; we impose it for tractability, as it lowers the dimensionality of media strategies. However, the restriction also means that we cannot apply the revelation principle. Accordingly, we will not take it for granted that consumers' actions mimic the signal they receive. Thus, each signal function

I is identified with q, which is the probability of submitting $s = 1$ when $t = 0$. The probability of $t = 1$ conditional on s is thus

$$\pi_{s,q} = \Pr(t = 1 \mid s) = \frac{s}{1+q} \tag{14}$$

For expositional simplicity, we will rule out the possibility that consumers play mixed strategies. Because c is continuously distributed, this restriction entails no loss of generality. Let $a_c(s \mid q, N)$ denote type c's subjectively optimal action in response to the signal s, given that he selected the media strategy (q, N).

Under any feasible (q, N), when the consumer observes the signal $s = 0$, he infers that $t = 0$ and therefore $ty = 0$ with probability one. Hence, his optimal action is $a = 0$ and he earns a payoff of 0. It follows that the evaluation of (q, N) by a consumer of type c, denoted $U_c(q, N)$, is

$$\begin{aligned}
& \left(\tfrac{1}{2} + \tfrac{1}{2}q\right) \cdot [p_N(t = 1|s = 1)p_N(y = 1|t = 1, a_c(s \mid q, N)) - ca_c(s \mid q, N)] \\
= \; & \tfrac{1+q}{2} \cdot [p(t = 1|s = 1)p_N(y = 1|t = 1, a_c(s \mid q, N)) - ca_c(s \mid q, N)] \\
= \; & \tfrac{1+q}{2} \cdot \left[\tfrac{\frac{1}{2}\cdot 1}{\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot q} \cdot p_N(y = 1 \mid t = 1, a_c(s \mid q, N)) - ca_c(s \mid q, N)\right] \\
= \; & \tfrac{1}{2}p_N(y = 1 \mid t = 1, a_c(s \mid q, N)) - \tfrac{c(1+q)}{2}a_c(s \mid q, N)
\end{aligned}$$

where $\frac{1}{2} + \frac{1}{2}q$ is the probability that type c receives the signal $s = 1$; and $p_N(t = 1|s = 1) = p(t = 1|s = 1)$ because all narratives correctly account for the joint distribution of (t, s). Importantly, the conditional distribution $p_N(y \mid t, a)$ is generated by applying the factorization formula for N to the *aggregate* joint distribution p, which arises from the choice behavior of all consumer types.

The monopolist's problem is to choose a menu $M = \{(q_c, N_c)\}_{c\in[0,1]}$ to maximize

$$\int_0^1 U_c(q_c, N_c)dF(c)$$

where

$$U_c(q_c, N_c) = \max_{(q,N)\in M} U_c(q, N),$$

subject to the following constraints. First, the aggregate conditional action distribution p satisfies

$$p(a=1\mid t=1)=\int_0^1 a_c(1)dF(c) \qquad\qquad p(a=1\mid t=0)=\int_0^1 q_c a_c(1)dF(c)$$

where $a_c(s) = a_c(s \mid q_c, N_c)$. Second, $a_c(1)$ needs to be subjectively optimal for type c in response to $s = 1$ (we have already established that $a_c(0) = 0$ for every c):

$$\begin{aligned} & \pi_{1,q_c}\cdot p_{N_c}(y=1\mid t=1,a=a_c(1))-ca_c(1) \\ \geq \;& \pi_{1,q_c}\cdot p_{N_c}(y=1\mid t=1,a=1-a_c(1))-c(1-a_c(1)) \end{aligned}$$

This design problem is a novel type of "second degree" discrimination, which arises because the monopolist cannot prevent consumers from freely choosing their favorite media strategy from the menu. Specifically, each consumer chooses the (I, N) in M that leads to a higher anticipatory utility given the distribution p.

At first glance, this seems to be a trivial problem, since there is no conflict of interest between the two parties: Both the consumer and the monopolist are guided by maximizing consumer anticipatory utility. However, consumers' choices exert a non-standard externality on one another: They all base their decision on the aggregate distribution p, which reflects the individual choices of all consumers. This novel interdependence among consumers is what makes the menu-design problem non-trivial.

Unlike the basic model, here consumers potentially face a choice from a variety of alternative media strategies. Each consumer selects the strategy that maximizes his own anticipatory utility (evaluated according to the aggregate distribution p). This choice involves choosing between different models, and it is obviously not a "rational" or "scientific" method for comparing models. Nevertheless, given that media consumers in our model are motivated by the pursuit of optimistic beliefs, such hedonic choice between narratives appears descriptively appropriate (Eliaz and Spiegler (2020) applied this criterion to how voters choose between political narratives).

Proposition 3 *The media maximizes its objective function with a menu that has the following features:*

(i) The menu includes exactly one pair (q^a, N^a)*; furthermore,* $q^a > 0$*, and there is* $c^* \in (0,1)$ *such that all consumer types in* $[0, c^*]$ *choose* (q^a, N^a) *and play* $a = s$*.*

(ii) The menu includes exactly one narrative $N \in \{N^t, N^\emptyset\}$*, coupled with an arbitrary* q*; there is* $c^{**} \in [c^*, 1)$ *such that all consumer types in* $[c^{**}, 1]$ *choose* (q, N) *and play* $a = 0$ *with probability one.*

(iii) If $c^{**} > c^*$*, then the menu also includes exactly one pair* (q^*, N^*)*; furthermore,* $q^* < q^a$*, and all consumer types in* (c^*, c^{**}) *choose* (q^*, N^*) *and play* $a = s$*.*

There are a few noteworthy differences from the homogenous case. First, under homogeneity, a single narrative (N^a) serves all consumers. The differentiation between consumer populations (characterized by distinct c) is done through the signal function. In contrast, differentiation between types in the heterogeneous case is carried out by offering a menu of narratives. Each of the narratives that keep the link $a \to y$ is coupled with a specific signal function. The reason is that given our restricted domain of signal functions, different media strategies that share the same narrative are Blackwell-ordered — and therefore unambiguously ranked in terms of the anticipatory utility they confer. As a result, no consumer will select dominated media strategies.

More specifically, the menu includes the narrative N^a, which is coupled with biased information toward $a = 1$; the narrative N^* (which need not be in the menu) has a smaller, potentially zero bias in that direction; while the other narratives generate the action $a = 0$. Thus, we have a proliferation of narratives, which lead to polarized beliefs and polarized behavior.

Second, in the homogenous case, market coverage is partial: Consumer types $c > \frac{1}{2}$ receive zero payoffs; they are effectively unserved. In contrast, in the heterogeneous case they earn positive anticipatory payoffs, thanks to the narratives N^t or $N^\emptyset$. This is made possible by the externality between types: Types with high c "free ride" on low-c types, who play $a = 1$ with positive probability.

The following result completes the characterization of the optimal menu when c is uniformly distributed.

Proposition 4 *When $c \sim U[0,1]$, the optimal menu consists of two media strategies: $(q = 1, N^a)$ and $(q = 1, N^\emptyset)$. Consumers with $c < \frac{3}{11}$ choose the former pair and always play $a = 1$; whereas consumers with $c > \frac{3}{11}$ choose the latter pair and always play $a = 0$.*

Thus, under uniformly distributed types, the media never provides *any* information to any consumer. Consumer behavior is highly polarized: Consumers with high c always play $a = 0$ whereas consumers with low c always play $a = 1$. What generates this polarization is the different narratives that the two consumer segments adopt: Low-c consumers opt for the empowering narrative while high-c consumers opt for the denial narrative.

4.2 Perfect Competition

Let us now define a notion of perfect competition in our media market. In this case, we need not restrict the set of feasible signal functions, except the purely expositional restriction to binary signals that take the values 0 or 1.

Definition 2 (Competitive equilibirum) *A a profile of media strategies $(I_c, N_c)_{c\in[0,1]}$ and a profile of consumer strategies $(a_c(s))_{c\in[0,1]}$ constitute a competitive equilibrium if:*

(i) For every c, $(I_c, N_c) \in \arg\max_{(I,N)} U_c(I,N)$ given the joint distribution p induced by $(a_c(s))_{c\in[0,1]}$.

(ii) For every c and s, $a_c(s)$ maximizes $\pi_{s,I_c} \cdot p_{N_c}(y = 1|t = 1, a) - ca$, where $\pi_{s,I}$ is given by (14).[4]

Unlike the monopoly case, here each media strategy targets a consumer type and maximizes his anticipatory utility, taking the entire distribution p as given. Media suppliers do not internalize the data externality between types, because they take the distribution p as given.

Proposition 5 *There is an essentially unique competitive equilibrium. Specifically, there is $\bar{c} \in (0,1)$ given by $2\bar{c} + F(\bar{c}) = 1$, such that: (i) for every $c < \bar{c}$,*

[4] The subscript I in $\pi_{s,I}$ plays the same role as the subscript q in (14).

I_c is the fully informative signal function and $N_c = N^*$; and (ii) for every $c > \bar{c}$, $N_c = N^t$.

By essential uniqueness, we mean that there could be other media strategies that implement the same beliefs and actions. For example, when a consumer chooses $N^{\emptyset}$, the exact signal function is irrelevant for his beliefs and actions. Also, we could replace N^* with N^a in the characterization, and consumers' beliefs would be identical.

It follows that for consumer types with low c, perfect competition leads to an unambiguous improvement in the informativeness of news media compared with monopoly. To see why, note that under monopoly, a positive measure of consumer type with c close to 0 receive biased information (which is coupled with the narrative N^a), whereas under competition they have rational expectations and full information. When $c \sim U[0,1]$, we have $\bar{c} = \frac{1}{3}$, which is above the cutoff $c^* = \frac{3}{11}$ of the monopoly case — i.e., competition improves informativeness for all consumer types.

5 Variations

This section returns to the homogenous-consumer case and explores variations on the basic model.

5.1 Introducing Rational Consumers

So far, we assumed that consumers are homogenous in the sense that their demand for news is not instrumental — rather, they use news media to cultivate desirable beliefs. This gives a role to false narratives as a vehicle for sustaining such beliefs. Consumers are not dogmatic: They are willing to accept any narrative, and their sole criterion for selecting a narrative is the anticipatory utility it induces.

Let us return to the basic setting in which $u(t,a,y) = ty - ca$, where all consumers have the same $c \in (0, \frac{1}{2})$, and introduce a different form of heterogeneity. A fraction λ of the consumer population have traditionally instrumental demand for information — that is, they aim to maximize their objective expected material payoff, rather than their anticipatory payoff. Furthermore, these consumers have rational expectations: They know the true model N^* and are therefore immune to narrative peddling by the media. We refer to these consumers as rational, and to the remaining consumers (who behave as in previous sections) as non-rational.

The way rational consumers evaluate a media strategy (I, N) is thus quite simple, because it is equivalent to the way non-rational consumers evaluate the strategy (I, N^*). The reason is that under the true model N^*, the distinction between anticipatory utility and objective expected material utility disappears: $U(I, N^*)$ is the ex-ante expected material payoff when the consumer subjectively best-replies to the beliefs induced by I.

We handle the new heterogeneity as we handled the heterogeneity in c in Section 4. In particular, we adopt the simplifying assumption that I must involve two signals, 0 and 1, such that in state $t = 1$ the signal is $s = 1$ with probability one. In this setting, this entails no loss of generality but facilitates exposition by emphasizing the methodological connection to the model of Section 4.1. Thus, each I is identified with the probability q of $s = 1$ in $t = 0$. The media's problem is to choose a menu of media strategies, $\{(q_r, N_r), (q_{nr}, N_{nr})\}$ to maximize

$$\lambda \cdot U_r(q_r, N_r) + (1 - \lambda) \cdot U_{nr}(q_{nr}, N_{nr})$$

subject to the constraint that (q_i, N_i) maximizes U_i for every type $i \in \{r, nr\}$, given the aggregate distribution induced by the strategy each type plays; and subject to the constraint that each type i's strategy is a best-reply given the beliefs induced by the type, the pair (q_i, N_i) he chooses, and the objective distribution.

Proposition 6 *There is an optimal menu that consists of a single pair (q, N).*

Thus, the media's problem of screening consumers according to their rationality is degenerate: The media can offer the same strategy to all consumers. The result holds more generally when $u(t, a, y) = v(t, y) - ca$, when $a, t \in \{0, 1\}$, v is an arbitrary function, and $c \in (0, \frac{1}{2})$.

For brevity, we do not provide a detailed derivation of the optimal media strategy in this environment for all values of λ. We make do with illustrating it for extreme values of λ.

Proposition 7 *When λ is sufficiently close to* 0*, the optimal media strategy is* $(\min\left\{1, \sqrt{\frac{1-\lambda}{2c}} - 1\right\}, N^a)$*. When λ is sufficiently close to* 1*, the optimal media strategy is* $(0, N^t)$*.*

Thus, introducing a small group of rational consumers into a population of non-rational consumers increases the informativeness of the signal that the media provides, while still keeping the empowering narrative. In contrast, introducing a small group of non-rational consumers into a population of rational consumers causes the media to offer the fatalistic narrative while continuing to give full information.

5.2 Other Separable Utility Specifications

In this sub-section we examine alternative specifications of $u(t, a, y)$. All definitions are adapted straightforwardly.

5.2.1 An Example: "Whac-a-Mole"

Impose the following structure on the data-generating process:

$$p(t = 1) = \frac{1}{2} \qquad\qquad p(y = 1 \mid t, a) = \beta(1 - a) + (1 - \beta)t$$

where $\beta \in (\frac{1}{3}, 1)$. The consumer's payoff function is $u(a, t, y) = \mathbf{1}[a = y]$.

We adopt the following interpretation for this specification. The action a represents a public decision how to allocate a scarce resource between two sectors or locations. For example, the dilemma is whether to allocate policing effort to one area of criminal activity or another. The state t indicates which sector is more dangerous. The outcome y indicates which sector ends up being active. Public policy is successful if it allocates the policing effort to the relevant sector. However, criminal activity exhibits a "whac-a-mole" property: When the government cracks down on one area of activity, criminals partly divert their activity to the other area. This explains the negative correlation between a and y. In this context, the media reports on the dangers posed by various sectors, and conveys a narrative about what ultimately determines the active sector. Consumer choice represents support for a certain public policy (e.g., voting for a political party that runs on this policy).

As before, let us begin our quest for optimal media strategies with the case in which the narrative is N^*. As usual, we can assume that the media provides full information. When $t = 1$, the consumer's payoff from $a = 1$ is $1 - \beta$, and the payoff from $a = 0$ is 0. Therefore, the consumer plays $a = 1$ when $t = 1$, and his payoff is $1 - \beta$. The case of $t = 0$ is handled symmetrically: the consumer plays $a = 0$, and earns a payoff of $1 - \beta$. It follows that the consumer's ex-ante anticipatory utility is $1 - \beta$. Thus, when the media conveys the true narrative and fully informs the consumer about t, the consumer correctly identifies the dangerous sector and plays $a = t$. At the same time, the consumer correctly takes the whac-a-mole effect into account.

We will later see that the narratives N^a and $N^\emptyset$ (which omit the link $t \to y$) are weakly inferior to N^*. Therefore, let us focus on the narrative N^t. We apply the revelation principle and take it for granted that $a = s$ in equilibrium. By definition,

$$p_{N^t}(y = 1 \mid s, a) = \sum_t p(t \mid s) p(y = 1 \mid t)$$

where

$$p(t = 1 \mid s = 1) = \frac{q_1}{q_0 + q_1} \qquad p(t = 1 \mid s = 0) = \frac{1 - q_1}{2 - q_0 - q_1}$$

$$\begin{aligned} p(y &= 1 \mid t = 1) = \sum_s p(s \mid t = 1) p(y = 1 \mid a = s, t = 1) \\ &= q_1 \cdot (1 - \beta) + (1 - q_1) \cdot 1 = 1 - \beta q_1 \end{aligned}$$

and

$$\begin{aligned} p(y &= 1 \mid t = 0) = \sum_s p(s \mid t = 0) p(y = 1 \mid a = s, t = 0) \\ &= q_0 \cdot 0 + (1 - q_0) \cdot \beta = \beta(1 - q_0) \end{aligned}$$

It follows that the consumer's payoff from playing $a = 1$ when $s = 1$ is

$$U_{N^t}(s = 1) = \frac{q_1}{q_0 + q_1} \cdot (1 - \beta q_1) + \frac{q_0}{q_0 + q_1} \cdot \beta(1 - q_0)$$

Likewise, the consumer's payoff from playing $a = 0$ when $s = 0$ is

$$U_{N^t}(s=0) = 1 - \left[\frac{1-q_1}{2-q_0-q_1} \cdot (1-\beta q_1) + \frac{1-q_0}{2-q_0-q_1} \cdot \beta(1-q_0)\right]$$

In order for this strategy to be an equilibrium, we need both expressions to be weakly above $\frac{1}{2}$. We will confirm this below. The strategy $a = s$ induces the following ex-ante anticipatory utility:

$$\frac{q_0+q_1}{2} \cdot U_{N^t}(s=1) + \left(1 - \frac{q_0+q_1}{2}\right) \cdot U_{N^t}(s=0)$$

This expression reduces to

$$1 + \frac{1}{2} \cdot [(2q_1-1)(1-\beta q_1) - q_1] + \frac{1}{2} \cdot [\beta(2q_0-1)(1-q_0) - q_0]$$

If the media employs a fully informative signal (i.e., $q_1 = 1$, $q_0 = 0$), this expression is equal to $1-\beta$, which is the maximal payoff from the true narrative N^*. It follows that as in the example of Section 3.1, the false narrative N^t can only be optimal when accompanied by imperfectly informative signals. The optimal signal function is

$$q_1 = \frac{1}{4} + \frac{1}{4\beta} \qquad\qquad q_0 = \frac{3}{4} - \frac{1}{4\beta}$$

Note that the optimal signal treats the two states symmetrically (since $q_0 + q_1 = 1$).

Plugging these values of q_0 and q_1, we can confirm that $U_{N^t}(s) > \frac{1}{2}$ for every s. The consumer's ex-ante anticipatory payoff is

$$\frac{(1+\beta)^2}{8\beta}$$

which is greater than $1 - \beta$.

The false narrative N^t that emerges from this exercise neglects the effect of a on y, and thus effectively pretends that the whac-a-mole effect does not exist. This enables the consumer to be more optimistic about the success of policies, but only when the narrative is accompanied by imprecise information.

5.2.2 A Characterization Result

The following result shows that the optimality of the narrative N^t in the whac-a-mole example is not a coincidence.

Proposition 8 *Suppose that* $u(t, a, y) = v(a, y) + w(t)$. *If the media can outperform the rational-expectations benchmark, then* N^t *is part of an optimal strategy.*

Thus, when u is separable in t, the fatalistic narrative is optimal. It is the analogue of the result that N^a is optimal when u is separable in a (Proposition 1). It can also be shown that in regular environments, this narrative can outperform the rational-expectations benchmark only if it leads to different behavior than the benchmark. The proof is the same as that of Proposition 2.

Finally, consider utility functions that are separable in y. This turns out to be a degenerate case, in the sense that it does not give rise to false narratives.

Proposition 9 *Suppose that* $u(t, a, y) = v(t, a)+w(y)$. *The benchmark media strategy is optimal.*

This completes the characterization of utility functions that are separable in at least one of the variables.

6 Discussion of Related Literature

This paper belongs to a research program on the role of causal narratives in economic and political interactions. Eliaz and Spiegler (2020) presented a modeling framework that formalizes causal narratives as directed acyclic graphs (building on Spiegler (2016)), where agents' adoption of narratives is based on the anticipatory utility they generate. Eliaz and Spiegler (2020) and Eliaz et al. (2022) applied this framework to political competition. The present paper brings the modeling approach to the market for news, focusing on the role of media as suppliers of narratives. Methodologically, its main contributions are: (i) modeling the media's joint provision of narratives and information; (ii) the novel screening problem that arises under consumer heterogeneity (in preferences or in rationality), due to the "data externality" between consumer types; and (iii) a new conception of a competitive media market.[5]

In terms of economic substance, our paper is part of the literature on media bias. This phenomenon has been extensively studied from various points of view. Prat and Stromberg (2013) and Gentzkow et al. (2015) provide comprehensive reviews of this literature. Our paper contributes to a theoretical strand in this literature that tries to explain media bias as a demand-based phenomenon arising from non-instrumental aspects of consumers' attitude to information. The basic idea in this literature is that consumer derive intrinsic utility from beliefs or from the news they consume, independently of their effect on decisions.

For example, Mullainathan and Shleifer (2005) model states of Nature and news as points along an interval. When a consumer confronts news, he incurs a cost that increases in the distance between the news and the mean of his prior belief. Media's strategic choices are thus reduced to a Hotelling-style model, where the consumer's psychological cost is analogous to a transportation cost in the standard Hotelling model.

Gentzkow et al. (2015) present a model in which consumers' utility has two additively separable components. The first component is a standard material expected-utility term that employs the consumer's posterior beliefs, which are obtained conventionally via Bayesian updating. This component treats beliefs in the usual instrumental manner. The second component is a function of the consumer's prior belief and the distribution of signals, such that if the prior leans in the direction of one state, then the function increases in the frequency of the signal whose label coincides with that state's label. This captures the idea that people like consuming news that support their prior beliefs. Note that this non-standard component does not reflect any belief updating. In particular, if the media always sends a signal that coincides with the state the consumer deems more likely (such that effectively the signal is entirely uninformative), the non-instrumental term reaches its maximal possible level given

[5] Recent empirical and experimental approaches to causal economic narratives include Ash et al. (2021), Andre et al. (2022), Charles and Kendall (2022), Macaulay and Song (2023) and Ambuehl and Thysen (2023).

the consumer's prior belief. Thus, both Mullainathan and Shleifer (2005) and Gentzkow et al. (2015) assume that the hedonic effect of news is orthogonal to Bayesian belief updating.

Against this background, our model introduces two innovations. To our knowledge, it is the first model of news media as suppliers of narratives in addition to information. It also appears to be the first model in which the hedonic aspect of media consumers' beliefs is fully integrated with Bayesian updating of these beliefs. Consumers' intrinsic utility from beliefs is a function of Bayesian posteriors induced by the information the media provides and the narrative it peddles. Eliaz and Spiegler (2006) is a precedent for this aspect of our model. In that paper, we presented of demand for information — represented by prior-dependent preferences over Blackwell experiments — which is driven by maximization of expected utility from (correctly specified) Bayesian posterior beliefs. Since that model allowed for non-convex utility from beliefs, it could accommodate demand for information that is non-increasing in Blackwell informativeness. Lipnowski and Mathevet (2018) examined optimal information provision for agents with such preferences. Thaler (2023) is an experimental study of the supply of information to agents who exhibit motivated reasoning (defined as non-Bayesian updating that is affected by the valence of beliefs).

The idea that misspecified models can be used to manipulate agents' beliefs has been studied in other contexts. Eliaz et al. (2021a) analyzed a cheap-talk model in which the sender provides not only information but also statistical data (or, equivalently, a model) that enables the receiver to interpret the information. Eliaz et al. (2021b) characterized the maximal distortion of perceived correlation between two variables that a causal model can generate in Gaussian environments. Schwartzstein and Sunderam (2021) and Aina (2023) studied persuasion problems in which the sender proposes models, formalized as likelihood functions, and the receiver chooses among them according to how well they fit historical data.

Finally, our paper is related to a small literature on strategic communication with agents whose inference from signals departs from the standard Bayesian, rational-expectations model (e.g., Hagenbach and Koessler (2020), Levy et al. (2022), de Clippel and Zhang (2022)).

References

[1] Aina, C. (2023). Tailored stories. Mimeo.

[2] Ambuehl S., & Thysen, H. (2023). Competing causal interpretations: An experimental study. Mimeo.

[3] Andre, P., Roth, C., Haaland, I., & Wohlfart, J. (2021). Inflation narratives. Mimeo.

[4] Ash, E., Gauthier, G., & Widmer, P. (2021). Text semantics capture political and economic narratives. Center for Law & Economics Working Paper Series, 2108-01720.

[5] Bergemann, D., & Morris, S. (2016). Bayes correlated equilibrium and the comparison of information structures in games. Theoretical Economics 11.2, 487-522.

[6] Bergemann, D., & Morris, S. (2019). Information design: A unified perspective. Journal of Economic Literature, 57(1), 44-95.

[7] Kendall, C. W., & Charles, C. (2022). Causal narratives. National Bureau of Economic Research, WP no. 30346.

[8] Chopra, F., Haaland, I., & Roth, C. (2023). The demand For news: Accuracy concerns versus belief confirmation motives. NHH Dept. of Economics Discussion Paper.

[9] de Clippel, G., & Zhang, X. (2022). Non-bayesian persuasion. Journal of Political Economy, 130(10), 2594-2642.

[10] Eliaz, K., Galperti, S., and Spiegler, R. (2022). False narratives and political mobilization. arXiv preprint arXiv:2206.12621.

[11] Eliaz, K., & Spiegler, R. (2006). Can anticipatory feelings explain anomalous choices of information sources?" Games and Economic Behavior 56.1, 87-104.

[12] Eliaz, K., & Spiegler, R. (2020). A model of competing narratives. American Economic Review, 110(12), 3786-3816.

[13] Eliaz, K., Spiegler, R., & Thysen, H. C. (2021a). Strategic interpretations. Journal of Economic Theory, 192, 105192.

[14] Eliaz, K., R. Spiegler, and Y. Weiss (2021b). Cheating with models. American Economic Review: Insights 3.4, 417-434.

[15] Esponda, I., and Pouzo, D. (2016). Berk–Nash equilibrium: A framework for modeling agents with misspecified models. Econometrica, 84(3), 1093-1130.

[16] Gentzkow, M., and Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. Econometrica, 78(1), 35-71.

[17] Gentzkow, M., Shapiro, J. M., and Stone, D. F. (2015). Media bias in the marketplace: Theory. In Handbook of media economics (Vol. 1, pp. 623-645). North-Holland.

[18] Levy, G., Morreno de Barreda, I., and Razin, R. (2022). Persuasion with correlation neglect: A full manipulation result. American Economic Review: Insights, 4(1), 123-138.

[19] Levy, G., Razin, R., and Young, A. (2022). Misspecified politics and the recurrence of populism. American Economic Review, 112(3), 928-962.

[20] Lipnowski, E., and Mathevet, L. (2018). Disclosure to a psychological audience. American Economic Journal: Microeconomics, 10(4), 67-93.

[21] Loury, G. C. (2020). The bias narrative vs. the development narrative. City Journal, December 2020, https://www.city-journal.org/article/the-bias-narrative-v-the-development-narrative.

[22] Macaulay, A., and Song, W. (2023). News media, inflation, and sentiment. In AEA Papers and Proceedings (Vol. 113, pp. 172-176). American Economic Association.

[23] Mullainathan, S., and Shleifer, A. (2005). The market for news. American Economic Review, 95(4), 1031-1053.

[24] Pearl, J. (2009). Causality. Cambridge university press.

[25] Prat, A., and Strömberg, D. (2013). The political economy of mass media. Advances in economics and econometrics, 2, 135.

[26] Schwartzstein, J., and Sunderam, A. (2021). Using models to persuade. American Economic Review, 111(1), 276-323.

[27] Spiegler, R. (2016). Bayesian networks and boundedly rational expectations. Quarterly Journal of Economics, 131, 1243-1290.

[28] Spiegler, R. (2020). Behavioral implications of causal misperceptions. Annual Review of Economics, 12, 81-106.

[29] Thaler, M. (2023). The supply of motivated beliefs. arXiv preprint arXiv:2111.06062.

Appendix: Proofs

Claim 1

Let us begin by showing that we can set $q_1 = 1$. First, note that we can rewrite (11) as

$$\frac{1}{2}\left[\frac{q_1}{2} + \frac{1}{2}\cdot\frac{q_1 q_0}{q_1+q_0} - c(q_1+q_0)\right]$$

The second and third terms inside the brackets are invariant to permuting q_0 and q_1, whereas the first term is increasing in q_1 and invariant to q_0. Therefore, it is optimal to set $q_1 \geq q_0$.

Second, note that we can rewrite (11) as

$$\frac{q_1+q_0}{2}\cdot\left[\frac{1}{1+\frac{q_0}{q_1}}\cdot\left(\frac{1}{2}+\frac{1}{2}\cdot\frac{\frac{q_0}{q_1}}{1+\frac{q_0}{q_1}}\right) - c\right]$$

Thus, the expression inside the square brackets only depends on the ratio q_0/q_1, while the term outside them increases in both q_0 and q_1. It follows that $q_1 = 1 \geq q_0$ in optimum. ■

Proposition 1

Consider the narrative N^t. In this case, the consumer believes that a has no causal effect on y. Therefore, for every s, he will only mix over actions that minimize c. Denote $\min_a c(a) = c^*$. Then, in equilibrium,

$$\begin{aligned}
U_{I,N^t}(s,a) &= \sum_s p(s) \sum_a p(a \mid s) \sum_t p(t \mid s) \sum_y p(y \mid t) v(t,y) - c^* \\
&= \sum_s p(s) \sum_t p(t \mid s) \sum_y \left(\sum_{a'} p(a' \mid t) p(y \mid t, a') \right) v(t,y) - c^* \\
&= \sum_t p(t) \sum_s p(s \mid t) \sum_y \left(\sum_{a'} p(a' \mid t) p(y \mid t, a') \right) v(t,y) - c^* \\
&= \sum_t p(t) \sum_{a'} p(a' \mid t) \sum_y p(y \mid t, a') v(t,y) - c^* \\
&\leq \sum_t p(t) \max_a \left[\sum_y p(y \mid t, a) v(t,y) - c(a) \right]
\end{aligned}$$

The final expression is the maximal anticipatory utility under the true narrative N^* coupled with full information. We can see that I is irrelevant for the consumer's anticipatory utility from action a. It follows that his ex-ante anticipatory utility can be written as

$$\begin{aligned}
& \textstyle\sum_a p(a) \sum_y p(y \mid a) v(a,y) \\
= & \textstyle\sum_a p(a) \sum_y \left(\sum_t p(t \mid a) p(y \mid t, a) \right) v(a,y) \\
= & \textstyle\sum_t p(t) \sum_a p(a \mid t) \sum_y p(y \mid t, a) v(a,y) \\
\leq & \textstyle\sum_t p(t) \max_a \sum_y p(y \mid t, a) v(a,y)
\end{aligned}$$

The final expression is the rational-expectations benchmark. Therefore, N^t cannot be part of a media strategy that outperforms it.

Now consider the narrative $N^{\emptyset}$. In this case,

$$\begin{aligned}
U_{I,N^{\emptyset}}(s,a) &= \sum_t p(t \mid s) \sum_a p(a \mid s) \left[\sum_y p(y) v(t,y) - c(a) \right] \\
&= \sum_t p(t \mid s) \left[\sum_y p(y) v(t,y) - \sum_a p(a \mid s) c(a) \right]
\end{aligned}$$

As in the previous case, the consumer believes that a has no effect on y. Therefore, for every s, $p(a \mid s) > 0$ only if a minimizes c. That is,

$$\sum_a p(a) c(a) = c^*$$

It follows that the consumer's ex-ante anticipatory utility is

$$\begin{aligned}
& \textstyle\sum_s p(s) \sum_t p(t \mid s) \sum_y p(y) v(t,y) - c^* \\
= \; & \textstyle\sum_t p(t) \sum_y p(y) v(t,y) - c^* \\
= \; & \textstyle\sum_t p(t) \sum_y \left(\sum_a p(a) p(y \mid a)\right) v(t,y) - c^* \\
= \; & \textstyle\sum_t p(t) \sum_a p(a) \sum_y \left[p(y \mid a) v(t,y) - c(a)\right] \\
\leq \; & \textstyle\sum_t p(t) \max_a \left[\sum_y p(y \mid a) v(t,y) - c(a)\right]
\end{aligned}$$

The final expression is the ex-ante anticipatory utility induced by the narrative N^a coupled with no information. It follows that the maximal anticipatory utility from $N^{\emptyset}$ can be replicated by the narrative N^a (coupled with fully uninformative signals). ■

Proposition 2

Assume the contrary — i.e., suppose there is a media strategy that induces the same $(p(a \mid t))_{t,a}$ as in the rational-expectations benchmark, yet outperforms it.

We first show that N^a is the only narrative that can be part of the strategy. The proof of Proposition 1 showed that N^t can never outperform the benchmark; and N^* cannot do so by definition. Now consider $N^{\emptyset}$. Under this narrative, the consumer will assign probability one to $\arg\min_a c(a)$ for every t. By assumption, this is also the consumer's behavior under rational expectations, but this contradicts the definition of regularity. This leaves N^a as the only possible narrative.

By regularity, $p(a \mid t)$ assigns probability one to a distinct action for each t. Let $t(a)$ be the unique state for which a is played under p. Since $t = t(a)$ whenever $p(t,s,a) > 0$, it follows that

$$p(y \mid a) = p(y \mid t, a)$$

for every (t,a) in the support of p. Consequently,

$$p_{N^a}(t, y \mid s, a) = p(t, y \mid s, a)$$

and therefore, the consumer's anticipatory utility under p and N^a is equal to the rational-expectations benchmark, a contradiction. ■

Proposition 3

The proof proceeds stepwise.

Step 1: *Without loss of generality, each narrative is coupled with a unique* q.

Assume the contrary — i.e., M contains two pairs (q, N) and (q', N) with $q' < q$. This means that the signal function given by q' Blackwell-dominates the signal function given by q (recall that $\Pr(s = 1 \mid t = 1) = 1$ under both functions). Any consumer type c who compares the two pairs will weakly prefer (q', N). The reason is that consumers take the objective distribution p as given. Since both pairs share the same narrative N, they both induce the same $p_N(y \mid t, a)$. This reduces the comparison between the pairs to a standard comparison between signal functions by an expected-utility maximizer.

Consider consumer types c who choose (q, N) from M. They must be indifferent between this pair and (q', N). Except for a zero-measure set of types, this can only be the case if the consumers take a constant action given each of the pairs (if their subjectively optimal action were state-contingent, then q and q' would induce different ex-ante expected utility). Moreover, this must be the same constant action since otherwise, only a particular (zero measure) type would be ex-ante indifferent between (q, N) and (q', N). Now suppose we remove (q, N) from the menu. Then, since (q', N) was optimal for these consumer types under M, this will continue to be the case and all these consumers will therefore choose (q', N). This will have no effect on the aggregate consumer strategy because (q', N) and (q, N) induce the same choices by consumers who chose (q, N) from M. Therefore, the switch by these consumers from (q, N) to (q', N) has no effect on how other consumer types evaluate any media strategy. It follows that without loss of generality, we can remove (q, N) from the menu. $\square$

Step 2: *Under the optimal menu, a positive measure of consumer types play* $a = 1$ *with positive probability.*

Assume that under the optimal menu, all consumer types play $a = 0$ with certainty. Then, regardless of the media strategy they choose, their anticipatory utility is 0. This is obviously the case for consumer types who choose N^* or N^a, because these narratives induce the correct belief that $a = 0$ causes $y = 0$ with certainty.

As to types who choose N^t, they estimate the conditional probability

$$p(y=1 \mid t) = \sum_a p(a=1 \mid t) \cdot \frac{2-t}{2} = 0$$

for every t. Therefore, these types earn zero anticipatory utility as well.

Finally, types who choose $N^\emptyset$ form the correct belief that $p(y=1)=0$ (because $a=0$ with probability one by assumption, and $p(y=1 \mid a=0) = 0$). It follows that all types earn zero anticipatory utility. However, if the monopolist offers the singleton menu consisting of the media strategy $(0, N^*)$, every type $c < \frac{1}{2}$ will earn $\frac{1}{4} - \frac{c}{2} > 0$, a contradiction.

Step 3: *Interval structure of types' choices*

By Step 1, we can assume that for each feasible narrative N there is at most one q such that $(q, N) \in M$. Because the narratives N^t and $N^\emptyset$ omit a as a cause of y, any consumer who chooses a media strategy that includes one of these narratives will always play $a=0$. Furthermore, whenever a consumer chooses a media strategy that includes N^* or N^a, he will play $a=0$ in response to $s=0$, since $\pi_s = 0$. Therefore, by Step 2, M must include a pair (q, N) such that $N \in \{N^*, N^a\}$, and there is a positive measure of consumer types who select this pair and play $a=1$ in response to $s=1$.

How consumer types rank (q^*, N^*) and (q^a, N^a) when playing $a=s$ in response to both pairs

Suppose that M includes both (q^*, N^*) and (q^a, N^a) such that for each of these pairs, there is a positive measure of consumer types who choose it and play $a=1$ in response to $s=1$. The ex-ante anticipatory utility that these pairs induce for a consumer of type c is:

$$\begin{aligned} U_c(q^*, N^*) &= \frac{1}{2} p(y=1 \mid t=1, a=1) - \frac{1+q^*}{2} c \\ &= \frac{1}{2} \cdot \frac{1}{2} [2-1] - \frac{1+q^*}{2} c \\ &= \frac{1}{4} - \frac{1+q^*}{2} c \end{aligned}$$

and

$$\begin{aligned} U_c(q^a, N^a) &= \frac{1}{2}p(y=1 \mid a=1) - \frac{1+q^a}{2}c \\ &= \frac{1}{2}\cdot\frac{1}{2}[2 - p(t=1 \mid a=1)] - \frac{1+q^a}{2}c \end{aligned}$$

It is immediate that $U_c(q^*, N^*) > U_c(q^a, N^a)$ only if $q^* < q^a$. Therefore, $q^a > 0$ in this case. Consequently, if $U_c(q^*, N^*) > U_c(q^a, N^a)$, then $U_{c'}(q^*, N^*) > U_{c'}(q^a, N^a)$ for every $c' > c$. It follows that if both (q^*, N^*) and (q^a, N^a) are in M and induce $a = 1$ in response to $s = 1$, then the set of types who choose (q^*, N^*) lies above the set of types who choose (q^a, N^a).

Showing that M includes (q^a, N^a) without loss of generality
Suppose that M does not include (q^a, N^a). Then, M includes a pair (q^*, N^*) such that a positive measure of consumer types choose this pair and play $a = 1$ in response to $s = 1$. (The reason is that consumers always play $a = 0$ in response to N^t or $N^\emptyset$, as well as in response to N^* when $s = 0$.) Now add (q^*, N^a) to the menu. It is evident that $U_c(q^*, N^a) \geq U_c(q^*, N^*)$ for every c. Therefore, if playing $a = s$ is optimal given (q^*, N^*), then it is also optimal given (q^*, N^a).

Moreover, if types who previously chose (q^*, N^*) and played $a = s$ switch to (q^*, N^a) and thus continue to play $a = s$, this switch does not change the joint aggregate distribution $p(t, a)$ because (q^*, N^a) and (q^*, N^*) share the same signal function and induce the same consumer strategy.

Finally, consider types who previously chose a media strategy that induces $a = 0$ for all s now switch to (q^*, N^a). By revealed preferences, the switch improves their own anticipatory utility, hence they must play $a = s$ (because if they play $a = 0$, their anticipatory utility is 0). At the same time, the switch does not affect $p(t = 1 \mid a = 1)$ because this probability is equal to

$$\frac{\frac{1}{2}\left(m(q^*, N^a) + m(q^*, N^*)\right)}{\frac{1}{2}\left(m(q^*, N^a) + m(q^*, N^*)\right) + \frac{1}{2}q^*\left(m(q^*, N^a) + m(q^*, N^*)\right)} = \frac{1}{1+q^*}$$

where $m(q^*, N^a) + m(q^*, N^*)$ is the total mass of types who choose either (q^*, N^*) or (q^*, N^a) (which is precisely the mass of types who choose $a = s$). As a result, the switch does not affect $U_c(q^*, N^a)$ for any c. By definition, it also does not affect $U_c(q^*, N^*)$.

It follows that we can sustain an equilibrium with weakly higher aggregate anticipatory utility when (q^*, N^a) is added to the menu. Thus, the menu will contain some pair (q^a, N^a) that induces $a = 1$ in response to $s = 1$.

The set of types who choose (q^a, N^a) *and play* $a = s$

We now establish that there is $c^* > 0$ such that all types in $[0, c^*)$ choose (q^a, N^a) and play $a = s$. To see why, suppose first that (q^*, N^*) is in M and that there is a positive measure of consumers who select this pair and play $a = s$. Then, as we showed, the set of types who select (q^a, N^a) over (q^*, N^*) and play $a = s$ lies to the left of the set of types who choose (q^*, N^*) and play $a = s$.

Now suppose (q^a, N^a) is the only media strategy in M that induces $a = s$. If type c prefers (q^a, N^a) to a media strategy that induces him to always play $a = 0$, then so does every $c' < c$.

Thus, the set of types who prefer (q^a, N^a) and play $a = 1$ in response to $s = 1$ is at the low end of $[0, 1]$.

The set of types who always play $a = 0$

Suppose first that M includes (q^*, N^*) such that a positive measure of types choose this pair and play $a = s$. The payoff from this choice is $\frac{1}{4} - \frac{1}{2}c$, which is negative for $c > \frac{1}{2}$. Thus, such types will respond to the same pair by always playing $a = 0$.

Now suppose that the only media strategy that induces $a = 1$ with positive probability is (q^a, N^a). But then,

$$U_c(q^a, N^a) = \frac{1}{2} \cdot \frac{1}{2}[2 - \frac{1}{1 + q^a}] - \frac{1 + q^a}{2}c$$

which is negative for $c \approx 1$. Hence, such types will respond to (q^a, N^a) by always playing $a = 0$.

Clearly, every type c who chooses N^t or $N^\emptyset$ always plays $a = 0$.

Finally, if type c prefers a media strategy that induces him to always play $a = 0$, then so does type $c' > c$. Therefore, there must be $c^{**} < 1$ such that all types $c > c^{**}$ choose a media strategy that induces $a = 0$ for all $s = 0$.

To conclude this step, there are two cutoffs, $0 < c^* \leq c^{**} < 1$, such that: (i) all types below c^* choose (q^a, N^a) and play $a = s$; (ii) all types between c^* and c^{**} choose (q^*, N^*) and play $a = s$; and (iii) all types above c^{**} play $a = 0$ with probability one. $\square$

Step 4: *The menu contains exactly one of the narratives* N^t *or* $N^\emptyset$

First, observe that the menu need not include both N^t and $N^\emptyset$. The reason is that both narratives induce $a = 0$ with probability one, such that they only potentially differ in the subjective ex-ante expected value of ty that they induce. In particular, they exert the same externality on other types. Therefore, the menu will include only the one that yields the higher payoff.

Second, suppose that M contains neither N^t nor $N^\emptyset$. Then, types above c^{**} will select the narratives N^* or N^a and always play $a = 0$, thus obtaining a payoff of 0. Suppose we add N^t or $N^\emptyset$ coupled with no information. By Step 2, $a = 1$ with positive probability, and therefore, both narratives will induce strictly positive payoff for types above c^{**}. We need to examine the possibility that lower types will switch from (q^a, N^a) or (q^*, N^*) to the new media strategy. However, if a type $c < c^{**}$ deviates in this direction, then so does every $c' \in (c, c^{**})$. Consider two cases.

Case 1: N^* is not in M. In this case, the deviation does not change $p(t = 1 \mid a = 1)$ because this quantity is not affected by increasing the share of consumers who always play $a = 0$, and the set of consumers who play $a = s$ all induce the same $\Pr(a \mid t)$. Therefore, it does not affect $U_c(q^a, N^a)$ for any c. By revealed preference, the deviation improves the ex-ante payoff of the deviating types. It follows that there is an unambiguous increase in aggregate consumer payoffs.

Case 1: N^* is in M. In this case, the set of deviating types is some interval $[c^{***}, c^{**}]$. As a result, since $q^* < q^a$, this deviation lowers $p(t = 1 \mid a = 1)$ and therefore increases $U_c(q^a, N^a)$ for any c. Moreover, it has no effect on $U_c(q^*, N^*)$ by definition. By revealed preference, the deviation improves the ex-ante payoff of the deviating types. It follows that the deviation increases aggregate consumer payoffs, even after taking into account the equilibrium effects of this deviation due to the data externality. □

It remains to show that $q^a > 0$ if $c^* = c^{**}$ — i.e., the only narrative that induces $a = 1$ with positive probability is N^a. We know from the homogenous case that for every $c < \frac{1}{2}$, $q^a > 0$ attains higher utility than $q^a = 0$ (and recall that the utility from $q^a = 0$ for $c > \frac{1}{2}$ cannot be positive). Moreover, since $q^a > 0$ generates a higher overall probability of $a = 1$, it exerts a positive externality than $q = 0$ on consumers who choose $N^\emptyset$ (it has no externality on consumers who choose N^t). Therefore, deviating from $q^a = 0$ to $q^a > 0$ is strictly profitable for the media. ■

Proposition 4

The proof proceeds stepwise, taking the characterization in Proposition 3 as a starting point.

Step 1: $c^* = c^{**}$

Assume that $c^{**} > c^*$. The payoffs induced by (q^*, N^*) and (q^a, N^a) at some c are

$$\begin{aligned} U_c(q^*, N^*) &= \frac{1}{4} - \frac{1+q^*}{2}c \\ U_a(q^a, N^a) &= \frac{1}{4}\left[2 - p(t=1 \mid a=1)\right] - \frac{1+q^a}{2}c \end{aligned}$$

In the proof of Step 3 of Proposition 3, we showed that $q^a > q^*$. Since $c \sim U[0,1]$, we can write

$$p(t=1 \mid a=1) = \frac{c^{**}}{c^{**} + c^* q^a + (c^{**} - c^*)q^*} = \frac{c^{**}}{c^{**}(1+q^*) + c^*(q^a - q^*)}$$

At c^*, the indifference between (q^*, N^*) and (q^a, N^a) can be written as follows:

$$\frac{1}{2}c^*(q^a - q^*) = \frac{1}{4}\left[1 - \frac{c^{**}}{c^{**}(1+q^*) + c^*(q^a - q^*)}\right]$$

Observe that if we slightly raise c^* and lower q^a such that q^a is still above q^* and $c^*(q^a - q^*)$ remains unchanged, then the indifference condition continues to hold, as long as we keep c^{**} fixed. In this way, $p(t=1 \mid a=1)$ remains unchanged. This modified consumer strategy is an equilibrium and it is strictly profitable for the media. To see why, note first that c^{**} is unchanged because by construction, $p(a=1)$ and $p(a=1 \mid t=1)$ are both unchanged, hence the payoff from N^t or $N^{\emptyset}$ is unchanged. Since the payoff from (q^*, N^*) is by definition invariant to $(p(a \mid t))$, the indifference at c^{**} continues to hold. Thus, the set of types who always play $a = 0$ and their utility are unaffected. Now consider the inframarginal types $c < c^*$. These types are now better off thanks to the decrease in q^a, and since $p(a=1 \mid t=1)$ is unchanged. The types who chose and continue to choose (q^*, N^*) are unaffected by definition. Therefore, the new equilibrium is an improvement, a contradiction.

What this step establishes is that we can restrict attention to menus M and consumer strategies that take either of the two following forms:

(i) $M = \{\{q^a, N^a), (q^t, N^t)\}$, all consumer types in $[0, c^*]$ choose (q^a, N^a) and

play $a = s$, and all consumer types $c > c^*$ choose (q^t, N^t) and play $a = 0$; and (ii) $M = \{\{q^a, N^a), (q^\emptyset, N^\emptyset)\}$, all consumer types in $[0, c^*]$ choose (q^a, N^a) and play $a = s$, and all consumer types $c > c^*$ choose $(q^\emptyset, N^\emptyset)$ and play $a = 0$. $\square$

Step 2: *Completing the characterization when M includes N^t*
Aggregate utility under $M = \{\{q^a, N^a), (q^t, N^t)\}$ is

$$\int_0^{c^*} U_c(q^a, N^a)dc + \int_{c^*}^1 U_c(q^t, N^t)dc$$

where

$$U_c(q^a, N^a) = \frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2}c$$

and

$$\begin{aligned} U_c(q^t, N^t) &= p(ty = 1) \\ &= p(t = 1) \cdot p(y = 1 \mid t = 1) \\ &= \frac{1}{2} \cdot p(a = 1 \mid t = 1) \cdot \frac{1}{2}(2 - 1) \\ &= \frac{1}{4}c^* \end{aligned}$$

Thus, the objective function can be written as

$$\begin{aligned} & \int_0^{c^*} \left\{\frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2}c\right\} dc + (1 - c^*) \cdot \frac{1}{4}c^* \\ = \quad & c^* \cdot \frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2} \cdot \frac{1}{2}(c^*)^2 + (1 - c^*) \cdot \frac{1}{4}c^* \end{aligned}$$

The cutoff c^* satisfies

$$\frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2}c^* = \frac{1}{4}c^*$$

Plugging this equation into the objective function, we obtain

$$\frac{2q^a + 1}{(2q^a + 3)^2}$$

The optimal value of q^a is $\frac{1}{2}$, yielding an aggregate utility of $\frac{1}{8}$. $\square$

Step 3: *Completing the characterization when M includes $N^\emptyset$*

Aggregate utility under $M = \{\{q^a, N^a), (q^\emptyset, N^\emptyset)\}$ is

$$\int_0^{c^*} U_c(q^a, N^a)dc + \int_{c^*}^1 U_c(q^\emptyset, N^\emptyset)dc$$

where

$$U_c(q^a, N^a) = \frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2}c$$

and

$$\begin{aligned}
U_c(q^\emptyset, N^\emptyset) &= p(t=1)\cdot p(y=1) \\
&= p(t=1)\cdot[p(t=1)\cdot p(y=1 \mid t=1) + p(t=0)\cdot p(y=1 \mid t=0)] \\
&= \frac{1}{2}\cdot[\frac{1}{2}\cdot p(a=1 \mid t=1)\cdot\frac{1}{2}(2-1) + \frac{1}{2}\cdot p(a=1 \mid t=0)\cdot\frac{1}{2}(2-0)] \\
&= \frac{1}{2}\cdot[\frac{1}{2}\cdot c^*\cdot\frac{1}{2}(2-1) + \frac{1}{2}\cdot c^* q^a\cdot\frac{1}{2}(2-0)] \\
&= \frac{1}{2}\cdot[\frac{1}{4}c^* + \frac{1}{2}c^* q^a] \\
&= \frac{c^*}{4}[\frac{1}{2} + q^a]
\end{aligned}$$

Thus, the objective function can be written as

$$\begin{aligned}
&\int_0^{c^*}\left\{\frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2}c\right\}dc + (1-c^*)\cdot\frac{c^*}{4}[\frac{1}{2} + q^a] \\
= \;& c^*\cdot\frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2}\cdot\frac{1}{2}(c^*)^2 + (1-c^*)\cdot\frac{c^*}{4}[\frac{1}{2} + q^a]
\end{aligned}$$

The cutoff c^* satisfies

$$\frac{1}{4}\left[2 - \frac{1}{1+q^a}\right] - \frac{1+q^a}{2}c^* = \frac{c^*}{4}[\frac{1}{2} + q^a]$$

Plugging this equation into the objective function, we obtain

$$\frac{3}{4}\left(2q^a + 1\right)^2 \frac{2q^a + 3}{\left(6q^a + 5\right)^2 \left(q^a + 1\right)}$$

This expression is monotonically increasing in q^a, hence the optimal value of q^a is 1, yielding an aggregate utility of approximately 0.139. □

Since the menu characterized by Step 3 yields a higher payoff than the one characterized by Step 2, the optimal menu includes the denial narrative, and sets $q^a = 1$. ■

Proposition 5

First, we establish that without loss of generality, I_c is the perfectly informative signal function for every c. The reason is that the maximization of type c's anticipatory utility takes p as given without taking into account the effect of the behavior induced by (I_c, N_c) on p_N. Therefore, U_c is effectively the maximum of functions that are linear in beliefs, hence it is convex in posterior beliefs. It follows that a fully informative signal maximizes U_c (as in the rational-expectations benchmark). It is the unique maximizer if it induces $a = s$.

Second, we show that all consumers play $a = 0$ when $t = 0$. This holds under N^t or $N^\emptyset$ because under these narratives, a has no causal effect on y. Under N^a or N^*, optimal provision of information implies that when $t = 0$ the consumer knows that $ty = 0$, and therefore finds $a = 0$ optimal.

An immediate consequence of the previous step is that $p(t = 0 \mid a = 1) = 0$. such that the formulas for U_c under N^* and N^a coincide. Thus, from now on, we will take it for granted that the only narrative that can induce $a = 1$ with positive probability is N^*. Let us denote by σ the fraction of consumers who play $a = 1$ when $t = 1$.

Third, we will show that N^t weakly outperforms $N^\emptyset$ for every consumer type. To see why, let us write down the anticipatory utility under each of these narratives. The anticipatory utility under N^t is

$$p(t = 1)p(y = 1 \mid t = 1) = \frac{1}{2} \cdot \sigma \cdot \frac{1}{2}(2 - 1) = \frac{1}{4}\sigma$$

The anticipatory utility under $N^\emptyset$ is

$$\begin{aligned} p(t = 1)p(y = 1) &= \tfrac{1}{2} \cdot \left[\tfrac{1}{2} \cdot p(y = 1 \mid t = 1) + \tfrac{1}{2} \cdot p(y = 1 \mid t = 0)\right] \\ &= \tfrac{1}{4}p(y = 1 \mid t = 1) \\ &= \tfrac{1}{8}\sigma \end{aligned}$$

Note that $p(y = 1 \mid t = 0) = 0$ because all consumers play $a = 0$ when $t = 0$.

Thus, the only narratives we need to consider are N^* and N^t. Moreover, we can assume that any consumer who adopts N^* will play $a = 1$ when $t = 1$, because otherwise he would get zero payoffs, which is below the payoff he can get from N^t. A consumer of type c will prefer N^* if $\frac{1}{4} - \frac{c}{2} > \frac{1}{4}\sigma$. There is a cutoff $\bar{c}$ characterized by $\frac{1}{4} - \frac{\bar{c}}{2} = \frac{1}{4}F(\bar{c})$, such that all $c < \bar{c}$ choose N^* and play $a = t$, while all $c > \bar{c}$ choose N^t and always play $a = 0$. ■

Proposition 6

Suppose first $N_{nr} \in \{N^t, N^\emptyset\}$. Then, type nr's evaluation of the pair (q_{nr}, N_{nr}) is independent of q_{nr}, and he always plays $a = 0$ in response to this pair. As to type r's evaluation of the same pair, it is independent of N_{nr} because this type applies N^*. Thus, type r's evaluation of (q_{nr}, N_{nr}) only depends on q_{nr}. Moreover, it is decreasing in q. It follows that if the media collapses the menu $\{(q_r, N_r), (q_{nr}, N_{nr})\}$ into a singleton $\{(0, N_{nr})\}$, it will weakly raise consumers' aggregate anticipatory utility.

Now suppose $N_{nr} = N^a$. First, it cannot be the case that type nr responds to (q_{nr}, N_{nr}) by always playing $a = 0$. The reason is that under this strategy, $U_{nr}(q_{nr}, N^a) = 0$. This cannot be optimal, because the singleton menu $\{(0, N^t)\}$ would outperform it: Since type r responds to this pair by playing $a = t$, $U_{nr}(0, N^t) = \frac{1}{4}\lambda > 0$ — i.e., the singleton menu maximizes U_r while inducing $U_{nr} > 0$.

We can thus take it for granted that type nr plays $a = s$. If $(q_r, N_r) \neq (q_{nr}, N_{nr})$, then in order for type r to favor (q_r, N_r) over (q_{nr}, N_{nr}), he must find the former pair more informative — i.e., $q_r < q_{nr}$. Therefore, $q_r < 1$, such that type r plays $a = s$ in response to (q_r, N_r).

Then,

$$U_r(q_r, N_r) = \frac{1}{4} - \frac{1 + q_r}{2}c$$

and

$$\begin{aligned}
U_r(q_{nr}, N^a) &= \frac{1}{4}\left[2 - p(t = 1 \mid a = 1)\right] - \frac{1 + q_{nr}}{2}c \\
&= \frac{1}{4}\left[2 - \frac{p(t = 1)p(a = 1 \mid t = 1)}{p(t = 1)p(a = 1 \mid t = 1) + p(t = 0)p(a = 1 \mid t = 0)}\right] \\
&\quad - \frac{1 + q_{nr}}{2}c \\
&= \frac{1}{4}\left[2 - \frac{1}{1 + \lambda q_r + (1 - \lambda)q_{nr}}\right] - \frac{1 + q_{nr}}{2}c
\end{aligned}$$

Crucially, $p(t = 1 \mid a = 1)$ is based on the aggregate distribution, which makes use of both types' strategies. Denote $\bar{q} = \lambda q_r + (1 - \lambda)q_{nr}$. It follows that the aggregate anticipatory utility is

$$\lambda\frac{1}{4}+(1-\lambda)\frac{1}{4}\left[2-\frac{1}{1+\bar{q}}\right]-\frac{1+\bar{q}}{2}c \tag{15}$$

This is exactly the same aggregate utility that would be obtained from a singleton menu $\{(\bar{q},N^a)\}$, where both types respond to the pair by playing $a=s$. therefore, there is no loss of generality in restricting ourselves to singleton menus. ■

Proposition 7

Suppose λ is close to 0 — i.e., the population consists almost entirely of non-rational consumers. Then, N^a is part of an optimal strategy. This follows from continuity relative to the $\lambda=0$ case. The derivative of (15) with respect to $\bar{q}$ is

$$\frac{1-\lambda}{4(1+\bar{q})^2}-\frac{c}{2} \tag{16}$$

The optimal $\bar{q}$ is given by the first-order condition.

Now suppose λ close to 1 — i.e., the population consists almost entirely of rational consumers. Suppose that N^a is part of an optimal strategy. As $\lambda\to 1$, (16) becomes negative, hence it is optimal to set $\bar{q}=0$ — i.e., offering a fully informative signal. However, when $\bar{q}=0$, N^a offers the same utility for non-rational consumers as N^*, hence it cannot outperform $(0,N^*)$. The only remaining media strategies we need to check are $(0,N^t)$ and $(0,N^\emptyset)$ *do* outperform it, note that

$$U_{nr}(0,N^t)=\frac{1}{4}\lambda>\frac{1}{8}\lambda=U_{nr}(0,N^\emptyset)$$

Obviously, since $\lambda\approx 1$, $U_{nr}(0,N^t)>U_r(0,N^t)=\frac{1}{4}-\frac{1}{2}c$. It follows that the optimal media strategy when λ is close to 1 is $(0,N^t)$. ■

Proposition 8

First, observe that for every feasible strategy (I,N), the ex-ante subjective expectation of $w(t)$ is

$$\sum_s p(s)\sum_{t'} p_N(t'\mid s)w(t')$$

Recall that for every feasible narrative N, $p_N(t'\mid s)\equiv p(t'\mid s)$. Therefore, the above expression reduces to

$$\sum_{t'} p(t')w(t') = Ew(t)$$

regardless of (I, N). Therefore, we can regard $Ew(t)$ as a constant in the media's objective function, and focus on the v term. Thus, from now on, we conveniently set $w(t) = 0$ for all t — this without loss of generality.

Consider the narrative N^a. In this case,

$$\begin{aligned} U_{I,N^a}(s,a) &= \sum_t p(t \mid s) \sum_y p(y \mid a)v(a,y) \\ &= \sum_y p(y \mid a)v(a,y) \end{aligned}$$

We can see that I is irrelevant for the consumer's anticipatory utility from action a. It follows that his ex-ante anticipatory utility can be written as

$$\begin{aligned} \sum_a p(a) \sum_y p(y \mid a)v(a,y) \\ &= \sum_a p(a) \sum_y \left(\sum_t p(t \mid a)p(y \mid t,a) \right) v(a,y) \\ &= \sum_t p(t) \sum_a p(a \mid t) \sum_y p(y \mid t,a)v(a,y) \\ &\leq \sum_t p(t) \max_a \sum_y p(y \mid t,a)v(a,y) \end{aligned}$$

The final expression is the consumer's maximal ex-ante anticipatory utility according to the true narrative N^*. Therefore, N^a cannot be part of a media strategy that outperforms the strategy of providing complete information and the true narrative.

Now consider the narrative $N^{\emptyset}$. In this case,

$$\begin{aligned} U_{I,N^{\emptyset}}(s,a) &= \sum_t p(t \mid s) \sum_y p(y)v(a,y) \\ &= \sum_y p(y)v(a,y) \end{aligned}$$

Here, too, we can see that I is irrelevant for the consumer's anticipatory utility from action a. It follows that his ex-ante anticipatory utility can be written as

$$\begin{aligned}
&\sum_a p(a) \sum_y p(y) v(a,y) \\
&= \sum_a p(a) \sum_y \left(\sum_t p(t) p(y \mid t) \right) v(a,y) \\
&= \sum_a p(a) \sum_t p(t) \sum_y p_{N^t}(y \mid t, a) v(a,y)
\end{aligned}$$

This is equal to the ex-ante anticipatory utility from the mixture over actions $(p(a))$, when the media conveys the narrative N^t and provides no information. It follows that the maximal anticipatory utility from $N^{\emptyset}$ can be replicated by the narrative N^t (coupled with fully uninformative signals). ■

Proposition 9

Consider the term $v(t,a)$. As we have observed, $p_N(t,s,a) \equiv p(t,s,a)$ for every feasible narrative N. Therefore,

$$\sum_s p(s) \sum_a p(a \mid s) E_N v(t, a \mid s, a) = E_{N^*}(v(t,a))$$

Now turn to the term $w(y)$. The ex-ante expectation of this term according to some feasible (I, N) is

$$\sum_y p_N(y) w(y)$$

We will now show that $p_N(y) \equiv p_{N^*}(y)$ for every feasible false narrative. First, observe that

$$\begin{aligned}
p_N(y) &= \sum_t p(t) \sum_s p(s \mid t) \sum_a p(a \mid s) p_N(y \mid t, a) \\
&= \sum_{t,a} p(t,a) p_N(y \mid t, a)
\end{aligned}$$

Let us now write this expression for each of the three feasible false narratives. For N^a,

$$\sum_{t,a} p(t,a) p(y \mid a) = \sum_a p(a) p(y \mid a) = p(y)$$

For N^t,

$$\sum_t p(t,a) p(y \mid t) = \sum_t p(t) p(y \mid t) = p(y)$$

Finally, for $N^{\emptyset}$,

$$\sum_{t,a} p(t,a)p(y) = p(y)$$

It follows that both terms of u are undistorted by any false narrative. Therefore, the media cannot outperform the true narrative (coupled with full information). ■