News Media as Suppliers of Narratives (and Information)*

Kfir Eliaz and Ran Spiegler†

September 30, 2024

Abstract

We present a model in which news media shape beliefs by providing information (signals about an exogenous state) and narratives (models of what determines outcomes). To amplify consumers' engagement, the media maximize their anticipatory utility. We characterize the optimal monopolisitic media strategy under various classes of separable consumer preferences, and demonstrate the synergy between false narratives and biased information. Consumer heterogeneity gives rise to a novel menu-design problem due to an "equilibrium data externality" among consumers. The optimal menu features multiple narratives and creates polarized beliefs and choices. These effects also arise in a competitive media market model.

*Spiegler acknowledges financial support from Leverhulme Trust grant RPG-2023-120. Eliaz acknowledges financial support from ISF grant 455/24. We thank Nageeb Ali, Alex Clyde, Daniel Gottlieb, Nathan Hancart, Shira Dvir-Gvirtsman, Jesse Shapiro, Adam Szeidl, Heidi Thysen, and seminar and conference audiences for helpful comments.

†Eliaz: School of Economics, Tel-Aviv University and NHH. 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.

"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)

"The masses have never thirsted after truth. Whoever can supply them with illusions is easily their master." (Gustave Le Bon)

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.*

The term "narrative" has multiple meanings in the context of news reporting. We conceive of narratives as qualitative explanations of what causes outcomes of interest. For example, while many exogenous variables can be reported, the media often selects only some of them as relevant for the outcome and therefore worthy of reporting. Another example is how the media shapes popular perceptions about what determines material success: Is it one's personal choices, or is it external circumstances beyond one's control?[1]

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. Media consumers use the narrative to interpret empirical regularities and form beliefs about the mapping from states and actions to outcomes. A false narrative is a misspecified causal 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 this phenomenon is driven in large part by 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

[1] This is similar to Glenn Loury's (2020) distinction between "development" and "bias" narratives. See Iyengar (1990) for evidence on how the media influences popular perceptions regarding the role of personal agency and external factors in escaping poverty.

providing complete and objective information, the market will provide it. Even if consumers have heterogeneous preferences, they all want more informative news.

Studies across several disciplines (psychology, political science and communication) have provided evidence that consumer demand for news media reflects non-instrumental attitudes to beliefs (e.g., Hart et al. (2009), Van der Meer et al. (2020), Taber and Lodge (2006)). These findings have inspired models of media bias in which beliefs enter directly into consumers' utility function (see Prat and Strömberg (2013) and Gentzkow et al. (2015) for surveys).

Our approach to non-instrumental demand for news is based on the idea that people are more likely to follow a news outlet when it helps them arrive at "desirable beliefs". For example, in the context of sporting events or military conflicts, consumers want to believe that their side will win (if the right action is taken).[2] Likewise, in the context of ideological debates, consumers want to believe they are "on the right side of history" (i.e., that posterity will prove them right — again, if appropriate actions are taken).[3] In the context of reporting on social issues (police brutality, climate change), they would like to believe in the ability of policy reforms ("defunding the police", switching to green energy) to improve social welfare. In the context of business reporting, retail investors want to believe they can "beat the market", and entrepreneurs want to believe they will be the next Jeff Besos. We assume that in all these contexts, consumers are attracted to news outlets that cultivate such hopeful beliefs. Accordingly, we propose a model in which news media aim to maximize consumers' *anticipatory utility* — i.e., their expected indirect utility from their posterior beliefs.

However, under the conventional assumption that news media only supply information, this objective cannot give rise to media bias. The reason is that under rational expectations, maximizing ex-ante anticipatory utility is indistinguishable from maximizing conventional indirect utility of a Bayesian rational consumer, where full information provision is known to be optimal. Thus, even when we assume non-instrumental demand for information (based on anticipatory utility), the standard view of the media as mere information providers cannot generate media bias.

[2] In discussing the popularity of patriotic coverage of the war in Afganistan and Iraq, a New York Times story (Ruthenberg (2003)) quotes MSNBC's president Erik Sorenson: "After Sept. 11 the country wants more optimism and benefit of the doubt...It's about being positive as opposed to being negative."

[3] E.g., see Chopra et al. (2023).

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 synergy between these two instruments: They complement each other in producing the hopeful 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 over states, actions and outcomes. A monopolistic media outlet commits ex-ante to a "media strategy", which consists of: (i) a Blackwell experiment (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. The "fatalistic" narrative postulates that only the state matters for the outcome.[4] Finally, the "denial" narrative asserts that neither the state nor the action cause the outcome (implicitly attributing outcomes to unspecified exogenous factors).

The representative consumer's strategy is a stochastic mapping from signals to actions. We interpret the strategy as the long-run aggregate behavior of many identical consumers, each making a one-shot decision. Together with the media strategy, it induces a long-run empirical joint distribution over states, actions and outcomes.

A narrative produces a subjective conditional belief over outcomes, by "fitting" it to the long-run joint 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 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

[4] The empowering and fatalistic narratives are the analogues of Loury's (2020) development and bias narratives, mentioned in footnote 1.

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 media's problem is to find a strategy and an equilibrium consumer strategy that maximize the consumer's ex-ante expected anticipatory utility. Incorporating equilibrium responses into the choice of a media strategy is in the spirit of the information-design literature (e.g., Kamenica and Gentzkow (2011), Bergemann and Morris (2019)). However, in standard models, equilibrium effects arise in multi-agent settings with payoff externalities. In contrast, in our model equilibrium effects arise because false narratives induce misspecified beliefs.

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 for a specification in which the consumer is an aspiring entrepreneur who considers a costly investment and dreams about making it big. An outcome in this example indicates whether the entrepreneur is the first to develop a new product, and the state of Nature indicates whether there are positive returns from being the first. Objectively, these two variables are negatively correlated. This specification, aptly titled "*the American dream*", is 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).

This feature of the optimal media strategy is robust in the following sense: For any action-separable utility function, if a media strategy gives higher anticipatory utility than the rational-expectations 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 synergy between false narratives and biased information.*

We then analyze the model under alternative classes of separable utility functions. When consumer utility 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 unintended consequences that the false narrative (coupled with noisy signals) neglects. When the utility function is separable in the outcome, the media cannot outperform the rational-expectations benchmark. These results cement one of the main insights of our paper: When demand for news media is driven by anticipatory utility, false narratives are an integral part of media bias.

Section 4 introduces preference heterogeneity (in the form of diverse investment costs) into our "American dream" example. We now envisage our monopolistic media provider as a gatekeeper or platform that restricts the entry of news outlets. Formally, the platform chooses a menu of media strategies, aiming to maximize aggregate anticipatory utility. For tractability, we restrict media strategies to report good news in the good state. Each consumer type chooses the media strategy that maximizes his own anticipatory utility.

At first glance, it may appear that incentive-compatibility is moot in this model, because all parties have a common objective: Maximizing consumers' anticipatory utility. However, this is not the case because of an "*equilibrium data externality*" that exists among consumer types. When evaluating a combination of a Blackwell experiment and a narrative, a consumer's conditional belief over states is generated by the specific Blackwell experiment. However, his conditional belief over outcomes is determined by how the narrative interprets the *aggregate* distribution over relevant variables, which reflects the choices of *all* consumer types. Although consumers are separate individuals with idiosyncratic preferences, they all rely on the same aggregate data to form beliefs over outcomes given the narratives they adopt. Consequently, changes in the behavior of one segment of the consumer population can change how another segment evaluates media strategies. Dealing with this externality in the context of a menu design problem is a methodological novelty of our paper, and one of our motivations for introducing heterogeneity in the first place.

The optimal menu has the following structure. Low-cost consumers choose the empowering narrative coupled with biased information (as in the homogenous case) and make the costly investment whenever they receive a good signal. Intermediate-cost types choose the true narrative coupled with a more informative signal. High-cost consumers choose a false narrative that neglects the action as a cause of outcomes (coupled with an arbitrary signal), and never make the investment.

One of the first two segments in the above-described structure may be crowded out by the optimal menu. In particular, the false narratives that cater to the extremes of the consumer distribution can chip away at the intermediate-cost segment (which adopts the true narrative) in a way that — thanks to the equilibrium data externality — only reinforces this "poaching" effect. Using a parametric example, we show that this effect can be stark: Consumer types below a cost threshold receive no information and always invest (egged on by the empowering narrative), while types above the threshold select the denial narrative and never invest. Thus, a heterogeneous population of consumers trying to make sense of the same aggregate data can end up holding highly polarized beliefs and taking opposite actions based on no information, just because they select different false narratives from the menu.

Finally, we 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 internalizing the equilibrium data externality. In the essentially unique equilibrium, only the true and fatalistic narratives prevail, where the former narrative is coupled with full information.[5] Thus, while perfect competition under-performs relative to monopoly in terms of consumers' aggregate anticipatory utility (because news outlets fail to incorporate the equilibrium data externality), it can provide more accurate information. However, competition does not eradicate wrong beliefs due to false narratives.

2 A Model

We begin by introducing the primitives of our model. There are four relevant variables, all taking finitely many values: A *state* of Nature t, an *action* a taken by a representative consumer, a *signal* s that the consumer observes before taking the action, and an *outcome* y. The state t is drawn from some exogenous distribution. The outcome y is determined according to some exogenous distribution conditional on a and t. The consumer has a vNM utility function $u(t, a, y)$.

A monopolistic news media outlet (referred to as "the media") commits ex-ante to a pair (I, N), where I is a *signal function*, which is a Blackwell experiment assigning a distribution over signals s to each state t; and N is a narrative, which is a subset of the two direct causes of y. The four possible

[5] This result does not rely on restricting the domain of feasible Blackwell experiments.

narratives are (with slight abuse of notation): The *true narrative* $N^* = \{t, a\}$; the *"empowering" narrative* $N^a = \{a\}$; the *"fatalistic" narrative* $N^t = \{t\}$; and the *"denial" narrative* $N^\emptyset = \emptyset$, which implicitly attributes y to other, unspecified exogenous factors.

The consumer's strategy is a (possibly stochastic) mapping from signals s to actions a. We think of this strategy as a description of long-run behavioral patterns by an infinite sequence of individual consumers, each one making a one-shot choice of action against the background of historical observations of all four variables. The consumer views these observations through the prism of the narrative N, as we will describe below. The long-run distribution p induced by the two parties' strategies 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 exogenous; the second term is given by the media's signal function I; and the third term is given by the consumer's strategy. The factorization reflects the causal structure underlying the data-generating process, which can be described by the following directed acyclic graph (DAG):

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

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. For example, the link $s \rightarrow a$ means that s is a direct cause of a. The DAG represents N^* by including the links $t \rightarrow y$ and $a \rightarrow y$. The three false narratives $N^a, N^t, N^\emptyset$ can be represented by DAGs that omit at least one of these links into y, while maintaining the true causal relations among t, s, a. For example, N^a omits the link $t \rightarrow y$, producing the DAG $t \rightarrow s \rightarrow a \rightarrow y$.

Given an objective full-support distribution p and the pair (I, N), the consumer forms the following belief over t and y conditional on the signal realization s and an action a:

$$\tilde{p}(t, y \mid s, a) = p_I(t \mid s)p_N(y \mid t, a) \tag{2}$$

where $p_I(t \mid s)$ is the objective posterior probability of t conditional on s, which is induced by the signal function I via Bayes' rule; and $p_N(y \mid t, a)$ is the perceived probability of y conditional on t and a, which is shaped by the narrative N. Specifically,

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

The interpretation is that the narrative N makes sense of the long-run distribution p by imposing a particular explanation for what causes variation in outcomes. The belief $p_N(y \mid t, a)$ is a systematic, narrative-based distortion of the objective conditional outcome distribution. Thus, the media affects the consumer's beliefs via two channels: (i) the signal function given by I, which determines the consumer's conditional belief over states; and (ii) the narrative N, which determines the consumer's conditional belief over outcomes.

More concretely, our interpretation of the second channel is that in addition to the signal s, the media also provides the statistical data described by $p_N(y \mid t, a)$ and frames it as a causal quantity. For example, when peddling the empowering narrative N^a, the media quotes statistical data about the historical correlation between a and y and pitches it as a causal effect of a on y.

Importantly, when the narrative N is false, $p_N(y \mid t, a)$ is not invariant to the the consumer's strategy, namely the long-run consumer average behavior given by $(p(a \mid s))_{a,s}$. To see why, elaborate $p_N(y \mid t, a)$ for each of the false narratives:

$$p_{N^a}(y \mid t, a) = \sum_{s',t'} p(s' \mid a) p(t' \mid s') p(y \mid t', a) \tag{3}$$

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

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

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. 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 consumer strategy $(p(a \mid s))_{a,s}$ is an ε-equilibrium if, whenever $p(a \mid s) > \varepsilon$, a maximizes*

$$V_{I,N}(s,a) = \sum_{t,y} p_I(t \mid s) p_N(y \mid t,a) u(t,a,y) \tag{6}$$

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

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

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

$$U(I,N) = \sum_t p(t) \sum_s p(s \mid t) \sum_a p(a \mid s) V_{I,N}(s,a) \tag{7}$$

subject to the constraint that the consumer strategy $(p(a \mid s))_{a,s}$ is an equilibrium. The media's objective function is the consumer's *expected anticipatory utility*. The interpretation is that anticipatory utility drives 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.

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

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

This is a conventional indirect utility function. Since it is a maximum over linear functions of 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*.

A revelation principle

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 *Without loss of optimality, we can let the set of signals coincide with the set of feasible actions, and restrict attention 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 holds 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. By expressions (3)-(5), this means that $p_N(y \mid t, a)$ remains unchanged as well, regardless of how t and a are jointly distributed with s. This enables the standard proof to go through. The revelation principle simplifies our analysis.

Discussion

We close this section with two comments on the interpretation of our model.

Non-instrumental demand for news. Our model assumes that consumers' demand for news is entirely non-instrumental. We make this assumption for several reasons. First, it obviously enables a sharper analysis. Second, as we saw, the distinction between instrumental and non-instrumental demand is irrelevant when the media is restricted to the true narrative. Thus, demand for news in our model already has a heavy dose of rationality. Third, in a previous version of this paper (Eliaz and Spiegler (2024)), we also considered a variant on our model with a mixed population of consumers, some of whom know the true model N^*. These consumers have conventionally instrumental demand for news; they evaluate media strategies purely on the basis of their signal functions. The media cannot discriminate between consumers and therefore offers a menu of media strategies. However, we showed that the optimal menu is a singleton, which is structurally the same as the optimal media strategy in the present model. Finally, we believe that a model in which consumers evaluate media strategies according to a weighted average of material and anticipatory utility (as in Brunnermeier and Parker (2005)) would deliver similar qualitative results while making the analysis considerably less transparent.

The media's anticipation of equilibrium effects. 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. At any rate, our methodology is in essence the same as in the multi-agent information design literature (e.g., Bergemann and Morris (2019)), which evaluates information structures according to agents' equilibrium responses. And as in that literature, our media can select among equilibria when its strategy induces multiple equilibria. The key difference is that the equilibrium notion in our model deviates from rational expectations.

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 a career decision or a dietary choice, in which case y represents earnings or health outcomes, respectively. The data that the consumer relies on to form beliefs 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.

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, involving a utility function that is separable in a. We then show that the qualitative features of the optimal strategy in this example hold for any specification that shares this separability. Finally, we extend our analysis to other classes of separable utility functions.

3.1 "The American Dream"

Let all four variables take values in $\{0, 1\}$. The exogenous components of the data-generating process are $p(t = 1) = \frac{1}{2}$ and $p(y = 1 \mid t, a) = a \cdot f_t$, where $f_0 > f_1 > 0$. The consumer's payoff function is $u(a, t, y) = ty - ca$, where $c \in (0, f_1)$. The action a represents a private decision whether to engage in a costly economic activity. The outcome y indicates whether the activity is successful in the sense of attaining some objective. The state t represents the returns from attaining the objective. High returns are associated with lower chances of a successful outcome, reflecting background equilibrium effects.[6]

For a concrete story, the consumer is an aspiring entrepreneur who decides whether to develop a new product. The outcome $y = 1$ represents being the first to succeed. The state $t = 1$ means there is demand for the product — in which case, more competitors flock to the market, thus lowering the entrepreneur's chances of being the first.

In an alternative story, the consumer is a high school student (or his parent) who decides whether to exert costly effort at school (private tutoring, extracurricular activities). A successful outcome means being admitted to a prestigious

[6] A more elaborate version of our example would model these forces explicitly, incorporating the contribution of media consumers' decisions to the equilibrium effects. Since this would add complexity without altering the main qualitative insight, we chose not to do so. In this sense, we perform a partial equilibrium analysis.

college. The state represents the college wage premium. A higher premium makes colleges more selective (hence the negative correlation between t and y).

Under both stories, the media provides information about the fundamentals represented by t, as well as a narrative about what drives the outcome y. The narrative determines whether people attribute personal material outcomes to internal factors under their control or to external factors beyond their control.

By the revelation principle, we can restrict attention to binary signals and an equilibrium in which the consumer always plays $a = s$. Denote $q_t = p(s = 1 \mid t)$.

Rational-expectations benchmark
Suppose the media offers the true narrative N^*. As we saw in Section 2, 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 $p(y = 1 \mid t = 1, a) = af_1$. Since $c < f_1$, the consumer plays $a = 1$. It follows that the rational-expectations benchmark in this example is $\frac{1}{2}(f_1 - c)$.

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 $a = 0$. This means that in any equilibrium, $a = 0$ with certainty for every t. Since $y = 0$ whenever $a = 0$, it follows that $p(y = 1) = 0$. 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 acknowledges a as a cause of y.

The empowering narrative
Under the narrative N^a,

$$p_{N^a}(ty = 1 \mid a, s) = p(t = 1 \mid s)p(y = 1 \mid a) \tag{8}$$

Observe that although the consumer believes that only a causes y, he cares about t because his net payoff is positive only when $ty = 1$.

The consumer's subjective payoff from $a = 0$ is zero regardless of s, because $p(y = 1 \mid a = 0) = 0$. Let us now turn to his payoff from $a = 1$ for each s. Applying the revelation principle (i.e., $a = s$ with probability one under p),

$$p(y = 1 \mid a = 1) = \sum_t p(t \mid a = 1)f_t = \sum_t p(t \mid s = 1)f_t \tag{9}$$

Plugging $p(t=1 \mid s=1) = q_1/(q_1+q_0)$ and $p(t=1 \mid s=0) = (1-q_1)/(2-q_1-q_0)$ and (9) in (8), we obtain

$$p_{N^a}(ty=1 \mid s=1, a=1) = \frac{q_1}{q_1+q_0} \cdot \left[f_0 - \frac{q_1}{q_1+q_0}(f_0-f_1) \right] \tag{10}$$

and

$$p_{N^a}(ty=1 \mid s=0, a=1) = \frac{1-q_1}{2-q_1-q_0} \cdot \left[f_0 - \frac{q_1}{q_1+q_0}(f_0-f_1) \right] \tag{11}$$

In order for the consumer's strategy to be an equilibrium, we need (10) and (11) to be weakly above and below c, respectively. Suppose these constraints hold. Then, when $s=0$, the consumer plays $a=0$ and gets zero payoffs. The consumer's anticipatory utility is thus

$$p(s=1) \cdot [p_{N^a}(ty=1 \mid s=1, a=1) - c]$$

which is equal to

$$\frac{q_1+q_0}{2} \cdot \left[\frac{q_1}{q_1+q_0} \cdot \left(f_0 - \frac{q_1}{q_1+q_0}(f_0-f_1) \right) - c \right] \tag{12}$$

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 outperforms 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. The following claim simplifies the problem.

Claim 1 *Under N^a, 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 simple proof of this claim (like all proofs in this paper) is in the Appendix. The claim reduces the consumer's anticipatory utility into

$$\frac{1}{2}\left[f_0 - \frac{1}{1+q_0}(f_0-f_1) - c(1+q_0) \right] \tag{13}$$

Note that $q_1 = 1$ also implies that (11) is zero, 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. When $c < f_0 - f_1$, we obtain

$$q_0 = \min\left\{1, \sqrt{\frac{f_0 - f_1}{c}} - 1\right\} > 0 \tag{14}$$

Since we saw that plugging $q_0 = 0$ in (13) reproduces the rational-expectations benchmark, it follows that the unique solution (14) strictly outperforms it.

Thus, as long as $c < f_0 - f_1$, 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. When $c \geq f_0 - f_1$, the media cannot outperform the rational-expectations benchmark.

In terms of the interpretation we offered for this example, the false narrative N^a claims that attainment of a career or business objective depends entirely on one's initiative. The accompanying signal function has an optimistic bias, claiming that returns from attaining the objective are high even when they are not. Thus, on one hand the media exaggerates the attractiveness of the external environment, while on the other hand it suppresses — via the empowering narrative — the negative effect that good fundamentals have on the chances of a successful outcome. Thus, we find it apt to refer to the media as peddling "the American dream" in this example.[7]

The synergy between false narratives and biased signals

Biased information is necessary for N^a to beat the rational-expectations benchmark. Suppose the media provides full information. This means that t and s are perfectly correlated ($s \equiv t$). The revelation principle means that $a \equiv s$ on the equilibrium path, such that a and t are perfectly correlated, too. But this means that omitting t as an explanatory variable for y does not lead to erroneous beliefs: $p(y \mid a)$ coincides with $p(y \mid t, a)$. In turn, this implies that the consumer effectively has rational expectations and perfectly monitors t, which gives the rational-expectations benchmark. Therefore, incomplete information is necessary for N^a to enhance the consumer's anticipatory utility.

The reason that the combination of N^a and biased information outperforms the benchmark is that it produces a *correlation-neglect* effect. As expression (8) makes explicit, the consumer believes that t and y are independent conditional

[7] The political-economics implications of popular perceptions of the role of personal choices in life outcomes have been studied by Piketty (1995) and Alesina and Angeletos (2005).

on (s, a). In reality, t and y are *negatively* correlated. By neglecting this correlation, the consumer attains a more optimistic belief about the product ty conditional on $s = a = 1$. However, this effect is non-null only when $p(a = 1 \mid t = 0) > 0$, which only happens when information is biased.

3.2 Generalizing the Example

The "American dream" example has two noteworthy features. First, the empowering narrative emerges as optimal. Second, it distorts consumer behavior away from the rational-expectations benchmark. We now show that both features hold more generally when u is *action-separable* — i.e., it takes the form $u(t, a, y) = v(t, y) - c(a)$.

Proposition 1 *Suppose u is action-separable. 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 media strategies that beat 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 is replicable 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 u is action-separable and the environment is regular. If an 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.[8]

Regularity plays a key role in the result. To see why, consider the payoff specification of Section 3.1, and modify the data-generating process by assuming $p(y = 1 \mid t, a) = 1-t$ 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^a (or, equivalently, $N^{\emptyset}$) 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.

3.3 Other Separable Utility Specifications

In this section we examine alternative specifications of $u(t, a, y)$. Let us begin with an example.

"Whac-a-Mole"

Impose the following structure on the exogenous components of the data-generating process: $p(t = 1) = \frac{1}{2}$, and $p(y = 1 \mid t, a) = \beta(1 - a) + (1 - \beta)t$, where $\beta \in (\frac{1}{3}, 1)$. The consumer's payoff function is $u(a, t, y) = \mathbf{1}[a = y]$.

We offer the following story behind this specification. The action a represents public allocation of domestic-security resources to one sector of criminal activity or another. The outcome y represents which sector ends up being active. The state t is an early indicator (e.g., a lagged realization) of y. Public policy is successful if it allocates the policing effort to the eventually active sector. However, criminal activity exhibits a "whac-a-mole" property: When

[8] Proposition 2 does not claim that the media necessarily employs biased signals. Thus, we cannot rule out the possibility that it is optimal for the media to accompany N^a with full information, anticipating that the consumer's subjective best-reply will involve mixing (the revelation principle does not guarantee that $a = s$ in *all* equilibria).

police cracks down on one sector of activity, criminals partly divert their activity to the other sector. This explains the negative correlation between a and y. In this context, the media reports on the indicators of criminal activity, and conveys a narrative about what ultimately determines the active sector. Consumer choice represents public support for a certain policy (e.g., voting for a political party that runs on this policy).

As in Section 3, the revelation principle allows us to focus on binary signals that take the form of action recommendations, which are followed in equilibrium. We continue to use the notation $q_t = p(s = 1 \mid t)$.

Claim 2 *The optimal media strategy in the whac-a-mole example consists of the fatalistic narrative N^t and the Blackwell experiment*

$$(q_0, q_1) = \left(\frac{3\beta - 1}{4\beta}, \frac{\beta + 1}{4\beta}\right)$$

The false narrative N^t that emerges in this example attributes all the variation in y to t (even though it is partly due to variation in a). In this sense, the narrative ignores the whac-a-mole effect. This enables the consumer to be more optimistic about the success of policies, but only when the narrative is accompanied by *noisy* signals: In each state, the media sends the wrong signal with positive probability $(3\beta - 1)/4\beta$.

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

Proposition 3 *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 Proposition 1 (it can also be shown that the analogue of Proposition 2 holds in this case).

Finally, consider utility functions that are separable in y.

Proposition 4 *Suppose that $u(t, a, y) = v(t, a) + w(y)$. No media strategy outperforms the rational-expectations benchmark.*

This case is degenerate in the sense that it never gives rise to false narratives.

4 Heterogeneous Consumers

In this section we extend the model by introducing consumer preference heterogeneity. Accordingly, the supply side consists of multiple media strategies that consumers can choose from, each according to his preferences. We analyze two market structures. In Section 4.1, we consider a monopolistic media platform acting as a gatekeeper that restricts the entry of media providers (each represented by a distinct media strategy). 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.

This extension introduces a methodological innovation. 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 are subjected to an "equilibrium data externality" from other consumers. This externality changes the formulation and analysis of the monopolistic and competitive models of the media market. A key 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 under these market structures.

4.1 Monopoly

In this version of the model, a monopolistic media platform commits ex-ante to a menu M of pairs (I, N). The set of consumer types is $C = [0, 1]$. Types are distributed according to a continuous and strictly increasing *cdf* G with full support. Let u_c be type c's payoff function. Each type c selects a pair $(I_c, N_c) \in M$ and a signal-dependent action $a_c(s)$ to maximize his ex-ante anticipatory utility. The platform's objective is to maximize consumers' *aggregate* ex-ante anticipatory utility.

The platform faces a "second-degree discrimination" problem, which arises because it cannot prevent consumers from freely choosing their favorite media strategy on the menu. What makes the problem non-standard is the equilibrium data externality described above.

The menu design problem

To formally describe the design problem, we begin with how consumers evaluate alternatives. Fix some profile of consumer types' media-strategy choices and signal-dependent actions, $(I_c, N_c, (a_c(s)))_{c\in C}$. Aggregate consumer behavior is given by $(p(a \mid t))_{a,t}$, where

$$p(a \mid t) = \int_c \sum_s p_{I_c}(s \mid t) \cdot 1[a_c(s) = a] dG(c) \tag{15}$$

and $(p_{I_c}(s \mid t))_{s,t}$ is the Blackwell experiment given by I_c. Denote $\mathbf{a} \equiv (a(s))_s$. Given $(p(a \mid t))_{a,t}$, consumer type c's ex-ante evaluation of any $(I, N, \mathbf{a})$ is:

$$U_c(I, N, \mathbf{a}) = \sum_s p_I(s) \sum_{t,y} p_I(t \mid s) p_N(y \mid t, a(s)) u_c(t, a(s), y) \tag{16}$$

In this formula, $p_I(s)$ and $p_I(t \mid s)$ are induced by the objective prior probability $p(t)$ and the Blackwell experiment given by I. The conditional probability $p_N(y \mid t, a)$ is as defined in Section 2, based on $p(t, a, y) = p(t)p(a \mid t)p(y \mid t, a)$, with $p(a \mid t)$ representing *aggregate* consumer behavior as in (15). Thus, although different consumer types may select different media outlets, they do not live in isolated islands; they all belong to the same society, and the news media they consume offer narratives that interpret the *same* aggregate data that arises from the choices of all consumers. It follows that the anticipatory payoff that some type c gets from his choice of triplet $(I_c, N_c, (a_c(s)))$ is affected by the choices made by all the other types since these determine the joint aggregate distribution $p(t, a, y)$.

Effectively, the platform's problem is to design a menu of (I, N) pairs, such that consumer types select items from this menu together with signal-dependent actions. An optimal menu maximizes consumers' aggregate anticipatory utility such that consumers' choices and actions satisfy some constraints. Formally, the platform chooses a profile of *triplets* $(I_c, N_c, \mathbf{a}_c)_{c\in C}$ to maximize

$$\int_c U_c(I_c, N_c, \mathbf{a}_c) dG(c)$$

subject to the constraints that for every $c \in C$: (i) the triplet $(I_c, N_c, \mathbf{a}_c)$ maximizes U_c over the set $\{I_c, N_c, \mathbf{a}_c\}_{c\in C}$; and (ii) $\mathbf{a}_c$ maximizes $U_c(I_c, N_c, \mathbf{a})$

given (I_c, N_c).

To see the "equilibrium data externality" that the menu design problem reflects, suppose some consumer types change their choice of triplet. If this change involves different signal-dependent actions, it can affect the aggregate distribution $p(a \mid t)$, which in turn may affect the anticipatory payoff of types who did not change their choice.

Revisiting the "American dream"

Complete characterization of this menu-design problem is beyond the scope of this paper. Here we make do with applying it to the "American dream" example of Section 3.1, extending it by introducing consumer heterogeneity. Specifically, we identify consumer types with the cost parameter c.[9]

We restrict the domain of feasible information strategies: Signals are binary, $s \in \{0, 1\}$, and the set of feasible signal functions satisfy $\Pr(s = 1 \mid t = 1) = 1$. The restriction entailed no loss of generality in the representative-consumer case of Section 3.1. This is no longer the case here. 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.

The need for this domain restriction arises from two non-standard sources of complexity. First, since different types can select different narratives having non-linear effects on their beliefs, there is no obvious single-crossing-like argument that would impose order on the incentive constraints. Second, the equilibrium data externality is *global*: When we change the (I, N) that one interval of types selects, this potentially affects the evaluation of all items on the menu by *all* types, however distant. Therefore, we cannot reduce the problem to checking local incentive constraints.

Thus, in what follows, 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 under I is thus $p_q(t = 1 \mid s) = s/(1 + q)$. In particular, when the consumer observes the signal $s = 0$, he infers that $t = 0$ and therefore $ty = 0$ with probability one. Hence, we can take it for granted that all consumer types play $a = 0$ and earn zero payoffs when receiving the signal $s = 0$. This simplifies

[9] Although we only analyze optimal menu design for this example, we chose to present the general problem first, because we believe this makes its logic more transparent.

$U_c(q, N, \mathbf{a})$ into

$$\begin{aligned}
U_c(q, N, \mathbf{a}) &= p_q(s=1) \cdot [p_q(t=1|s=1)p_N(y=1|t=1, a(s=1)) - ca(s=1)] \\
&= \frac{1+q}{2} \cdot \left[\frac{1}{1+q} p_N(y=1|t=1, a(s=1)) - ca(s=1)\right] \\
&= \frac{1}{2} p_N(y=1|t=1, a(s=1)) - \frac{c(1+q)}{2} a(s=1) \qquad (17)
\end{aligned}$$

Likewise, consumers' aggregate state-dependent behavior can be simplified into

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

We can now restate the platform's problem: Choose a profile $(q_c, N_c, a_c(1))_{c \in C}$ that maximizes

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

subject to the constraints that for every c, $U_c(q_c, N_c, a_c(1)) \geq U_c(q_{c'}, N_{c'}, a_{c'}(1))$ for every $c' \in C$; and that $a_c(1)$ maximizes U_c given (q_c, N_c). The latter constraint can be written as follows:

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

Proposition 5 *The platform maximizes its objective function with a menu that has the following structure. There exist $c^{**} \in (0,1)$ and $c^* \in [0, c^{**}]$ such that:*

(i) All consumer types in $[0, c^*]$ choose (q^a, N^a) with $q^a > 0$ and play $a \equiv s$.

(ii) All consumer types in $(c^*, c^{**}]$ choose (q^*, N^*) and play $a \equiv s$. Moreover, if $c^* = 0$, then $q^* = 0$; and if $c^* > 0$, then $q^* < q^a$.

(iii) All consumer types in $(c^{**}, 1]$ choose the same narrative $N \in \{N^t, N^\emptyset\}$, coupled with an arbitrary q, and always play $a = 0$. In particular, if $c^* = 0$, then $N = N^t$.

There are a few noteworthy differences from the homogenous case of Section 3. 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.[10] 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 $a \rightarrow y$ is coupled with a unique signal function. The reason is that thanks to our restricted domain of signal functions, different media strategies that share the same narrative are Blackwell-ordered. A media strategy that involves a Blackwell-dominated signal function will never be chosen.

More specifically, the menu includes *at least* one of the narratives that include the link $a \rightarrow y$, and *exactly* one that does not. If N^a is on the menu, it is coupled with biased information. If N^* is on the menu, it is coupled with more precise information (perfectly precise when N^a is not on the menu). The narrative on the menu without the $a \rightarrow y$ link generates the action $a = 0$ with certainty. Thus, we have a proliferation of narratives, which lead to polarized beliefs and behavior.

Second, in the homogenous case, market coverage is partial: Consumer types $c > \max\{f_1, f_0 - f_1\}$ receive zero payoffs and 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 equilibrium data externality. High-c types "free ride" on low-c types, who play $a = 1$ with positive probability. Whether N^t or $N^\emptyset$ prevail in the high-c range depends on whether $p(t = 1 \mid a = 1)f_1$ is larger than $p(t = 0 \mid a = 1)f_0$. While $f_0 > f_1$ by assumption, $p(t = 1 \mid a = 1)$ is above $p(t = 0 \mid a = 1)$ unless every consumer type that ever plays $a = 1$ receives entirely uninformative signals. Therefore, in general the comparison between these two narratives is ambiguous.

At first glance, it might seem obvious that adding N^t or $N^\emptyset$ to the menu is optimal, because it improves the welfare of high-c consumer types. However, the equilibrium data externality complicates the argument: Mid-c types may react to the addition by switching away from a narrative that includes the link $a \rightarrow y$, which induces $a = 1$ and thus creates the positive externality that makes N^t or $N^\emptyset$ attractive in the first place. This switch changes aggregate behavioral patterns and therefore affects this very externality. This feedback effect is what makes the proof of Proposition 5 non-standard. Nevertheless, we show that the addition of N^t or $N^\emptyset$ is profitable, despite the switching of mid-c

[10]In principle, the menu-design problem can be defined for a homogenous consumer population, where identical consumers can select different pairs (I, N). Nevertheless, it can be shown that in the "American dream" example, the degenerate menu of Section 3.1 is optimal.

types.

Just as the narratives N^t or $N^\emptyset$ attract the upper part of the range of types that would otherwise choose N^*, the narrative N^a (when it is on the menu) attracts the lower part. As a result, the fraction of consumers who adopt the true narrative shrinks. The following example shows that it may all but disappear.

Claim 3 *Let $f_0 = 1$, $f_1 = \frac{1}{2}$, and $c \sim U[0,1]$. There is an optimal menu consisting 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 $a = 0$.*

Under this menu, the media never provides *any* information to *any* consumer, and only false narratives prevail. Consumer behavior is highly polarized: High-c consumers always play $a = 0$ whereas low-c consumers 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.

The absence of any information provision in Claim 3 is not a robust feature of the optimal menu. To see why, suppose $f_1 > f_0 - f_1$, and assume that G is almost entirely concentrated around some $c \in (f_0 - f_1, f_1)$. This is a perturbation of a homogenous-population model in which $(0, N^*)$ is the optimal media strategy. It can be shown that in the perturbed case, $c^* = 0$ — i.e., N^a is not on the menu. The reason is that there are too few low-c types to make N^a profitable (they would find N^a superior to N^* only when there are enough consumers who adopt N^a and play $a = 1$ at $t = 0$). As a result, $(0, N^*)$ is the only media strategy on the menu that generates $a = 1$. High-c types adopt the narrative N^t. Without loss of optimality, both items on the menu involve fully informative signals.

Thus, Proposition 5 does not imply a clear-cut conclusion regarding the prevalence of media bias: That will depend on the consumer type distribution. The robust feature of the optimal menu is the proliferation of narratives, including false ones, which generate polarized beliefs and actions.

4.2 Perfect Competition

Let us now consider a competitive media market, in which every media firm is small and therefore cannot affect aggregate consumer behavior.

Definition 2 (Competitive equilibrium) *A profile $(I_c, N_c, \mathbf{a}_c)_{c\in C}$ is a competitive equilibrium if for every $c \in C$, $(I_c, N_c, \mathbf{a}_c)$ maximizes U_c over all possible triples $(I, N, \mathbf{a})$; where U_c is defined as in (16) and calculated taking as given the aggregate distribution $(p(a \mid t)_{a,t}$ that is induced by $(\mathbf{a}_c)_{c\in C}$.*

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

As in the previous sub-section, let us apply this definition to the "American dream" example. As before, consumer types are identified by their cost parameter. Unlike the previous sub-section, here 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.

Proposition 6 *There is an essentially unique competitive equilibrium in the "American dream" setting. Specifically, there is $\bar{c} \in (0, 1)$ uniquely given by the equation $\bar{c} = f_1(1 - G(\bar{c}))$, such that: (i) for every $c < \bar{c}$, 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 profile of beliefs and actions. When a consumer chooses N^t, 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.

The optimal menu has the same structure as in the monopoly case when $c^* = 0$. Consumers who always play $a = 0$ select the narrative N^t. The reason is that $N^\emptyset$ beats N^t when $p(y = 1) > p(y = 1 \mid t = 1)$ — which can only happen if $p(a = 1 \mid t = 0) > 0$. However, this is never the case in competitive equilibrium, because consumers who sometimes play $a = 1$ are perfectly informed and therefore play $a \equiv t$.

Revisit our numerical example from the previous sub-section, where $f_0 = 1$, $f_1 = \frac{1}{2}$, and $c \sim U[0, 1]$. In this case, we have $\bar{c} = \frac{1}{3}$, which is above the cutoff $c^* = \frac{3}{11}$ of the monopoly case. Thus, under this specification, competition improves informativeness for *all* consumer types.

5 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. (2024) 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; and (*iii*) a new conception of a competitive media market.[11]

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 Strömberg (2013) and Gentzkow et al. (2015) provide comprehensive reviews. Our paper contributes to a theoretical strand in this literature that tries to explain media bias as a demand-based phenomenon arising from consumers' non-instrumental demand for information. The basic idea is that consumers derive intrinsic utility from beliefs or from the news they consume, independently of their effect on decisions. This idea draws on findings in disciplines outside economics. For example, a meta-study by Hart et al. (2009) finds that when participants are faced with a choice between information that supports their prior beliefs and information that may challenge it, they exhibit a preference for the former. Within the context of news media, Van der Meer et al. (2020) find evidence that participants are more likely to view news that confirm their prior beliefs than news that oppose them.

Mullainathan and Shleifer (2005) attempt to model this phenomenon. They

[11] 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).

formalize both 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 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 belief updating. This dissociation is a limitation of existing models we are aware of. We believe that even when people appear to behave as if they dislike a clash between news and their *prior* beliefs, this may in fact reflect their prediction that the news will lead to undesirable *posterior* beliefs.

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 among the first models (along with Herrera and Sethi (2022)) in which the hedonic aspect of media consumers' beliefs is fully integrated with Bayesian updating. Eliaz and Spiegler (2006) is a precedent for this aspect of our model. In that paper, we studied demand for information — represented by prior-dependent preferences over Blackwell experiments — driven by maximization of expected utility from (correctly specified) Bayesian posterior beliefs. Since that model allows for non-convex utility from beliefs, it accommodates demand for information that is non-increasing in Blackwell informativeness. Lipnowski and Mathevet (2018) examine optimal information provision for agents with such preferences.

The assumption that news consumers seek hopeful narratives may appear to be at odds with the common notion that consumers are attracted to negative news and that news media exhibit a "negativity bias" (e.g., see Robertson et al. (2023)). We believe, however, that the two ideas are orthogonal. First, what often attracts consumers to negative news is their element of drama or sensationalism (e.g., a collapsing bridge). Second, when we measure negativity of a news piece by the prevalence of "negative words", we may fail to capture its message that bad outcomes are a consequence of wrong decisions (which a false narrative like N^a in our model conveys). Finally, it is not clear that media consumers invariably regard bad things that happen to other people as bad news.

Our assumption of Bayesian updating rules out non-Bayesian responses to information due to motivated reasoning. Taber and Lodge (2006) show that when subjects are confronted with information that questions their prior beliefs, they try to discredit it. Thaler (2023) studies experimentally the supply of information to agents whose belief updating exhibits motivated reasoning.

The idea that misspecified models can be used to manipulate agents' beliefs has been studied in other contexts. Eliaz et al. (2021a) analyze 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) characterize 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) study 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. Szeidl and Szucs (2024) present a model in which the sender can use "propaganda" to alter the receiver's perception of the sender's motives. 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. Working paper.

[2] Alesina, A., & Angeletos, G. (2005). Fairness and redistribution. *American economic review* 95, 960-980.

[3] Ambuehl S., & Thysen, H. (2023). Competing causal interpretations: An experimental study. Working paper.

[4] Andre, P., Roth, C., Haaland, I., & Wohlfart, J. (2021). Inflation narratives. Working paper.

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

[6] Bergemann, D., & Morris, S. (2016). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics* 11(2), 487-522.

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

[8] Brunnermeier, M., & Parker, J. (2005). Optimal expectations. *American Economic Review* 95, 1092-1118.

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

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

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

[12] Eliaz, K., Galperti, S., and Spiegler, R. (2024). False narratives and political mobilization. *Journal of the European Economic Association*, forthcoming.

[13] Eliaz, K., & Spiegler, R. (2006). Can anticipatory feelings explain anomalous choices of information sources? *Games and Economic Behavior* 56(1), 87-104.

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

[15] Eliaz, K., & Spiegler, R. (2024). News media as suppliers of narratives (and information). arXiv preprint arXiv:2403.09155.

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

[17] Eliaz, K., Spiegler, R., and Weiss, Y. (2021b). Cheating with models. *American Economic Review*: *Insights* 3(4), 417-434.

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

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

[20] 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.

[21] Hart, W., Albarracin, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin* 135(4), 555-588.

[22] Herrera, H., & R. Sethi (2022). Identity-based elections. CEPR Working Paper.

[23] Iyengar, S. (1990). Framing responsibility for political issues: The case of poverty. *Political Behavior* 12(1), 19-40.

[24] Kamenica, E., & Gentzkow, M. (2011) Bayesian persuasion. *American Economic Review* 101(6), 2590-2615.

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

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

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

[28] 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.

[29] Macaulay, A., & Song, W. (2023). News media, inflation, and sentiment. *AEA Papers and Proceedings* 113, 172-176.

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

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

[32] Piketty, T. (1995). Social mobility and redistributive politics. The Quarterly journal of economics 110, 551-584.

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

[34] Robertson, C. E., Pröllochs, N., Schwarzenegger, K., Pärnamets, P., Van Bavel, J. J., & Feuerriegel, S. (2023). Negativity drives online news consumption. *Nature Human Behaviour* 7(5), 812-822.

[35] Rutenberg, Jim (2003). A nation at war: The news media; cable's war coverage suggests a new "Fox effect" on television journalism. *The New York Times.* April 16.

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

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

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

[39] Szeidl, A., & Szucs, F. (2024), The political economy of alternative realities. Working paper.

[40] Taber, C., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50(3), 755–69.

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

[42] Van der Meer, T.G.L.A., Hameleers, M., & Kroon, A.C. (2020). Crafting our own biased media diets: The effects of confirmation, source, and negativity bias on selective attendance to online news, *Mass Communication and Society* 23(6), 937-967.

Appendix: Proofs

Claim 1

Rewrite (12) as

$$\frac{q_1}{2}\left[f_0 - \frac{q_1}{q_1+q_0}(f_0-f_1) - \frac{q_1+q_0}{q_1}c\right]$$

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 $q_1 \geq q_0$. Now rewrite the expression as

$$\frac{q_1}{2}\left[f_0 - \frac{1}{1+\frac{q_0}{q_1}}(f_0-f_1) - (1+\frac{q_0}{q_1})c\right]$$

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

Proposition 1

Denote $\min_a c(a) = c^*$. Under the narratives N^t and $N^\emptyset$, the consumer believes that a has no causal effect on y. Therefore, for every s, he will only mix over actions that minimize c. Moreover, without loss of optimality, I is completely uninformative, such that a is independent of t (i.e., $p(a \mid t) = p(a)$ for every a, t).

Under N^t (coupled with no information), any equilibrium induces

$$\begin{array}{rcl} U(I, N^t) & = & \sum_t p(t) \sum_y p(y \mid t) v(t,y) - c^* \\ & = & \sum_t p(t) \sum_{a'} p(a') p(y \mid t, a') v(t,y) - c^* \end{array}$$

Since $c(a') = c^*$ whenever $p(a') > 0$, $U(I, N^t)$ can be rewritten as

$$\sum_t p(t) \sum_{a'} p(a') \left[\sum_y p(y \mid t, a') v(t,y) - c(a') \right]$$

which is by definition weakly below

$$\sum_t p(t) \max_a \left[\sum_y p(y \mid t, a) v(t,y) - c(a) \right]$$

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

Now turn to $N^{\emptyset}$ (coupled with no information). Any equilibrium induces

$$\begin{array}{rcl} U(I, N^{\emptyset}) & = & \sum_t p(t) \sum_y p(y) v(t,y) - c^* \\ & = & \sum_t p(t) \sum_y \left(\sum_{a'} p(a') p(y \mid a') \right) v(t,y) - c^* \\ & = & \sum_t p(t) \sum_{a'} p(a') \sum_y \left[p(y \mid a') v(t,y) - c(a') \right] \end{array}$$

This last expression is by definition weakly below

$$\max_a \sum_t p(t) \left[\sum_y p(y \mid a) v(t,y) - c(a) \right] \tag{18}$$

Note that throughout this calculation,

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

because a is independent of t. This expression involves entirely exogenous quantities — i.e., it is invariant to the consumer's strategy. Yet, it is not well-defined if $p(a) = 0$, which makes (18) potentially ill-defined. However, any full-support perturbed strategy for an uninformed consumer will ensure that (18) is well-defined. Moreover, it will be 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. ■

Claim 2

Consider the narrative N^*. As before, 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 establish later that the narratives N^a and $N^{\emptyset}$ 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 $p(t=1 \mid s=1) = q_1/(q_0+q_1)$, $p(t=1 \mid s=0) = (1-q_1)/(2-q_0-q_1)$,

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

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 $q_0 = (3\beta-1)/4\beta$ and $q_1 = (\beta+1)/4\beta$. 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 $(1+\beta)^2/8\beta$, which is greater than

$1-\beta$. ■

Proposition 3

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,

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

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 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,

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

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

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

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 4

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

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

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). $\blacksquare$

Proposition 5

We say $(q, N) \in M$ is *redundant* if there exists $(q', N') \in M$ that every consumer type finds weakly preferable. It can be shown that there is no loss of optimality in focusing on menus without redundant strategies. Unlike standard menu-design models, it is not entirely obvious that this is true, because of equilibrium data externalities. We omit the proof because it would make the proof more tedious without adding insight.

The proof proceeds stepwise.

Step 1: *Under any menu, every consumer type $c > 0$ chooses $a = 0$ with certainty in response to the signal $s = 0$. Moreover, every consumer type $c > 0$ who chooses the narratives N^t* or $N^{\emptyset}$ plays $a = 0$ for every s.

By our restriction on the set of signal functions, $\Pr(t = 0 \mid s = 0) = 1$ under any media strategy. Therefore, the consumer understands that $ty = 0$ with probability one, regardless of a. As a result, any consumer type with $c > 0$ will optimally choose $a = 0$, regardless of the narrative he adopted.

A consumer type $c > 0$ who chooses N^t or $N^{\emptyset}$ believes that a has no effect on y. Therefore, he prefers not to incur the cost c of playing $a = 1$. $\square$

Step 2: *Without loss of optimality, 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 both pairs share the same narrative N, hence 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. Therefore, (q, N) is redundant, contradicting our assumption that M does not contain redundant media strategies. $\square$

Step 3: *Under any optimal menu, a positive measure of consumer types play $a \equiv s$. All these types necessarily choose N^a* or N^*.

Assume that under an optimal menu, almost all consumer types play $a = 0$ with certainty. Then, regardless of the media strategy they choose, their anticipatory utility is zero. 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 $p(y = 1 \mid t) = p(a = 1 \mid t) \cdot f_t = 0$ 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 platform offers the singleton menu consisting of the media strategy $(0, N^*)$, every type $c < f_1$ will earn $\frac{1}{2}(f_1 - c) > 0$, a contradiction.

It follows that under an optimal menu, a positive measure of consumer types sometimes play $a = 1$. By Step 1, this means that these consumer types play $a = s$ for every s and cannot choose N^t or $N^{\emptyset}$. □

In preparation for the next steps, we present expressions for the ex-ante anticipatory utility (derived from (17)) that a consumer type c obtains from the pairs (q, N^*) and (q, N^a) when he responds to them by playing $a \equiv s$:

$$\begin{aligned} U_c(q, N^*) &= \frac{1}{2}p(y = 1 \mid t = 1, a = 1) - \frac{1+q}{2}c \\ &= \frac{1}{2}f_1 - \frac{1+q}{2}c \end{aligned} \tag{19}$$

and

$$\begin{aligned} U_c(q, N^a) &= \frac{1}{2}p(y = 1 \mid a = 1) - \frac{1+q}{2}c \\ &= \frac{1}{2}[p(t = 0 \mid a = 1)f_0 + p(t = 1 \mid a = 1)f_1] - \frac{1+q}{2}c \end{aligned} \tag{20}$$

Since $f_0 > f_1$, it is immediate that $U_c(q, N^a) \geq U_c(q, N^*)$.

Step 4: Any *optimal menu induces a cutoff* $c^{**} \in (0, 1)$ *such that all types* $c > c^{**}$ *always plays* $a = 0$, *while every* $c < c^{**}$ *chooses* N^* *or* N^a *and plays* $a \equiv s$.

By Step 3, M includes the narratives N^* or N^a. Suppose $(q^*, N^*) \in M$. If type c chooses this pair and always plays $a = 0$, he obtains zero anticipatory utility. If he plays $a \equiv s$, he obtains $\frac{1}{2}(f_1 - c)$, which is negative for $c > f_1$. Thus, types

$c > f_1$ will respond to (q^*, N^*) by always playing $a = 0$. Suppose $(q^a, N^a) \in M$. If type c chooses this pair and always plays $a = 0$, he obtains zero anticipatory utility. Note that (20) is negative when $c \approx 1$ (because $f_1 < f_0 \leq 1$ and $p(t = 1 \mid a = 1) > 0$). Hence, types $c \approx 1$ will respond to (q^a, N^a) by always playing $a = 0$. Finally, suppose M includes N^t or $N^\emptyset$. Then, by Step 1, every type c that chooses one of these narratives always plays $a = 0$ and obtains a payoff that is independent of the value of c.

We have thus established that if a consumer chooses a pair (q, N) that induces him to always play $a = 0$, he obtains a payoff that is independent of his value of c. In contrast, the payoff from choosing any pair that induces playing $a \equiv s$ is decreasing in c. It follows that if type c prefers a media strategy that induces him to always play $a = 0$, then so does type $c' > c$.

We can conclude that there exists a cutoff $c^{**} < 1$ such that all types $c > c^{**}$ choose a media strategy that induces always playing $a = 0$, while all types $c < c^{**}$ choose a media strategy that involves N^* or N^a and induces playing $a \equiv s$. Furthermore, by Step 3, $c^{**} > 0$. $\square$

Step 5: *Suppose an optimal menu M includes both N^a and N^*, and each is selected by a positive measure of consumers. Then, there is $c^* \in (0, c^{**})$, such that every $c < c^*$ chooses (q^a, N^a), whereas every $c \in (c^*, c^{**})$ chooses (q^*, N^*). Furthermore, $q^a > q^*$.*

By the definition of c^{**}, every $c < c^{**}$ chooses (q^*, N^*) or (q^a, N^a) and plays $a \equiv s$. Note that if $q^a = q^* = 0$, then $U_c(q, N^*) = U_c(q, N^a)$ for every c. Also, recall that both (q^*, N^*) and (q^a, N^a) induce zero payoffs when the consumer always plays $a = 0$. This contradicts our restriction to menus without redundancies.

Suppose that $q^a = q^* > 0$. Therefore, $p(t = 0 \mid a = 1) > 0$. Since $f_0 > f_1$, it follows from (19) and (20) that $U_c(q, N^a) > U_c(q, N^*)$ for every c. It follows that no consumer type $c < c^{**}$ will choose (q^*, N^*), a contradiction.

Now suppose $q^* \neq q^a$. Then, q^* or q^a are strictly positive, hence $p(t = 0 \mid a = 1) > 0$. Since $f_0 > f_1$, it follows that $U_c(q^*, N^*) \geq U_c(q^a, N^a)$ only if $q^a > q^*$. Therefore, we must have $q^a > q^*$ in order to have a positive measure of consumers who choose (q^*, N^*). Note that 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 chosen by a positive measure of consumers, then the set of types who choose (q^*, N^*) lies above the set of types who choose (q^a, N^a). $\square$

Step 6: *An optimal menu includes exactly one narrative $N \in \{N^t, N^\emptyset\}$, which* *is chosen by all consumer types $c > c^{**}$.*

First, observe that M 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 (we will later see that this externality is in fact null, but this is immaterial at this stage of the proof). Because we rule out redundancies, the menu will include only one of these two narratives — 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 zero payoffs. We now show that adding N^t or $N^{\emptyset}$ coupled with fully uninformative signals is profitable for the platform. Since $c^{**} > 0$, a positive fraction of consumers plays $a = 1$ at $t = 1$, and therefore $p(y = 1 \mid t = 1) > 0$, such that both N^t and $N^{\emptyset}$ 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. Note that if a type $c < c^{**}$ deviates in this direction, then so will every $c' \in (c, c^{**})$. Consider two cases.

Case 1: M includes exactly one $N \in \{N^*, N^a\}$. In this case, the deviation does not change $p(t = 1 \mid a = 1)$, for two reasons. First, this conditional probability is not affected by the share of consumers who always play $a = 0$. Second, consumer types who play $a \equiv s$ all induce the same $\Pr(a \mid t)$, since they all choose the same pair (q, N). Therefore, it does not affect $U_c(q, N)$ 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, even after taking into account the equilibrium effects of this deviation due to the equilibrium data externality.

Case 2: M includes both N^* and N^a. By Step 5, if a type c that chooses N^a belongs to the set of deviating types, then every type that chooses N^* also belongs to that set. Therefore, the deviation increases the fraction of types who select (q^a, N^a) within the population of consumers who ever play $a = 1$. Since by Step 5 $q^a > q^*$, it follows that the deviation raises $p(t = 0 \mid a = 1)$, 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 equilibrium data externality. $\square$

Step 7: *If $c^* = 0$, then $q^* = 0$, and types above c^{**} choose N^t. If $c^* = c^{**}$ and there is no alternative optimal menu that would induce $c^* = 0$, then $q^a > 0$.*

Suppose $c^* = 0$ and yet $q^* > 0$. Then, by (19)-(20), $U_c(q, N^a) > U_c(q, N^*)$ for every c. If the platform replaces (q^*, N^*) with (q^*, N^a), it increases the payoff of every type that selected the original pair and now selects the new pair.

If types above c^{**} (who formerly always played $a = 0$) now switch to the new pair, then by revealed preferences their payoff increases. At the same time, they do not affect $p(t \mid a = 1)$, and therefore do not affect $U_c(q^*, N^a)$ for any c. Finally, they exert a positive externality on types above c^{**} who do not switch, because it increases $p(a = 1 \mid t)$ at any t.

If types who formerly chose (q^*, N^*) now switch to a pair that induces always playing $a = 0$, then again by revealed preferences, the payoff of every type who selects such a pair increases (including types who originally chose a pair that induces always playing $a = 0$). However, this switch does not affect $p(t \mid a = 1)$ (since all consumers who ever play $a = 1$ face the same signal function given by q^*), and therefore exerts no externality on types who now select (q^*, N^a).

It follows that replacing (q^*, N^*) with (q^*, N^a) is profitable for the platform. Let us now calculate the anticipatory utility from N^t and $N^\emptyset$ for any type c when $c^* = 0$:

$$
\begin{aligned}
U_c(q, N^t) &= p(t = 1)p(a = 1 \mid t = 1)f_1 = \frac{1}{2}G(c^{**})f_1 \\
U_c(q, N^\emptyset) &= p(t = 1)\sum_t p(t)p(a = 1 \mid t)f_t = \frac{1}{4}G(c^{**})f_1
\end{aligned}
$$

where the last equality follows from the fact that $q^* = 0$, such that $p(a = 1 \mid t = 0) = 0$. Therefore, $U_c(q, N^t) > U_c(q, N^\emptyset)$.

Now suppose $c^* = c^{**}$. If $q^a = 0$, then the pair $(0, N^a)$ is equivalent to $(0, N^*)$. Therefore, we can replicate M with an equivalent menu that induces $c^* = 0$ and sets $q^* = 0$. $\square$

This completes the proof. $\blacksquare$

Claim 3

The proof proceeds stepwise, taking the characterization in Proposition 5 as a starting point.

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

Assume that $c^{**} > c^* > 0$. 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 5 of Proposition 5, 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 action profile 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 infra-marginal 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 $a = 0$. □

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

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

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 $\frac{1}{8}$. □

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

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

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 $q^a = 1$.

The only remaining case is $c^* = 0$. By Proposition 5, this means that all types $c < c^{**}$ choose the media strategy $(0, N^*)$. However, recall that this pair is equivalent to $(0, N^a)$ for all types. Steps 2 and established that this pair is inferior to $(1, N^a)$. ■

Proposition 6

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. Thus, U_c is effectively the maximum of linear functions of beliefs, hence 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 $a \equiv s$.

Second, by Step 1 in the proof of Proposition 5, all consumers play $a = 0$ when $t = 0$. This means 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 show that N^t weakly outperforms $N^\emptyset$ for every consumer type. The anticipatory utility under N^t is

$$p(t = 1)p(y = 1 \mid t = 1) = \frac{1}{2}\sigma f_1$$

The anticipatory utility under $N^\emptyset$ is

$$\begin{aligned} p(t = 1)p(y = 1) &= \tfrac{1}{2} \cdot \left[\tfrac{1}{2}p(y = 1 \mid t = 1) + \tfrac{1}{2}p(y = 1 \mid t = 0)\right] \\ &= \tfrac{1}{4}p(y = 1 \mid t = 1) \\ &= \tfrac{1}{4}\sigma f_1 \end{aligned}$$

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. A consumer of type c will prefer N^* if $\frac{1}{2}(f_1 - c) > \frac{1}{2}\sigma f_1$. Therefore, there is a unique cutoff $\bar{c}$, such that all $c < \bar{c}$ choose N^* and play $a = t$, while all $c > \bar{c}$ choose N^t and always play $a = 0$. Plugging $\sigma = G(\bar{c})$, we obtain the implicit equation for $\bar{c}$. ■