Narrative-Driven Fluctuations in Sentiment: Evidence Linking Traditional and Social Media*

Alistair Macaulay
University of Oxford

Wenting Song
Bank of Canada

June 23, 2023

Abstract

We study the empirical importance of narratives by linking narratives in newspapers to the sentiment of social media users. First, we model narratives as directed acyclic graphs and show how exposure to different narratives can affect expectations in an otherwise-standard macroeconomic model. We then measure competing narratives in news media reports on the US yield curve inversion in 2019, using techniques in natural language processing. Linking these narratives to data from Twitter, we show that exposure to the narrative of an imminent recession is associated with a more pessimistic sentiment, while exposure to a more neutral narrative implies no such change in sentiment. In addition, we find that narratives are contagious: their effects spread in the social network, even to those who are indirectly exposed.

JEL: D8, E3, G1

Keywords: Media, Twitter, Narratives, Textual Analysis

*Macaulay (alistair.macaulay@economics.ox.ac.uk): University of Oxford. Song (wsong@bank-banque-canada.ca): Bank of Canada. We thank Julian Ashwin, Michael Ehrmann, Yuriy Gorodnichenko, P.J. Lambert, Nora Lamersdorf, John Leahy, Michael McMahon, Kristoffer Nimark, Pablo Ottonello, Christopher Roth, Robert Shiller, Ran Spiegler, Leif Anders Thorsrud, Laura Veldkamp, and participants at various seminars and conferences for helpful comments and suggestions. We also thank Andrew Usher for his help with the Twitter data. Alexa Kaminski provided excellent research assistance. Macaulay acknowledges financial support from the John Fell OUP Research Fund. The views expressed herein are those of the authors and not necessarily those of the Bank of Canada. First version: June 29, 2022.

1. Introduction

Information provided by the news media can have substantial effects on household beliefs (see, for example, Chahrour, Nimark and Pitschner, 2021). However, alongside the factual information, news stories also provide a narrative (Shiller, 2017), which may describe forces that have led to the economic event or interpret what it may mean for the readers. Do those narratives affect beliefs beyond the effect of the information being reported? If they do, then the responses of households and firms to macroeconomic shocks may depend on the narratives that are popular in the media at the time and not just on fundamentals.

In this paper, we study the importance of narratives by linking articles in traditional news media to engagement with the articles on social media. Motivated by a model formalizing the potential role of narratives in general equilibrium, we capture competing narratives in traditional news media using natural language processing. We trace the influence of those narratives by comparing the sentiment of Twitter users before and after engaging with a particular narrative. Focusing on the 2019 yield curve inversion in the US, we provide direct evidence that exposure to a narrative associating the inversion with an imminent recession causes users to display a more pessimistic sentiment. Exposure to an alternative narrative claiming the yield curve has lost its predictive power has no such effect on sentiment. Since all articles in our sample report on the same event, the difference between these responses stems from the different narratives rather than reactions to the underlying event.

The 2019 yield curve inversion provides an ideal laboratory to assess the effects of narratives on sentiment. First, yield curve inversions are a popular recession indicator in the US, but there is a history of false positives (Bauer and Mertens, 2018). As a result, different narratives circulated in the media simultaneously, offering different interpretations of what the inversion meant for the macroeconomic outlook. This allows us to compare different narratives in the cross-section, and therefore to separate the effect of narratives from the information about the event itself. Second, the inversion was brief. The precise timing of the yield curve inversion was driven by very short-term volatility in financial markets, making it plausibly exogenous with respect to other macroeconomic news and monetary policy. This allows us to use a high-frequency event study approach to isolate the effect of narratives on sentiment.

We begin by developing a theoretical framework to guide our empirical exercise. In a textbook consumption-and-saving problem, we formalize household narratives as *directed acyclic graphs* (DAGs), as in Eliaz and Spiegler (2020) and Andre, Haaland, Roth and Wohlfart (2022b). DAGs are network representations of simple structural models, which have a natural interpretation as "causal" stories. We then embed these households with heterogeneous narratives in a simple New Keynesian model. Specifically, we consider two competing narratives that mirror those seen in news media around the time of the yield curve inversion: one in which a popular recession indicator such as the yield curve is related to future changes in output, and another in which it is not. In equilibrium, the two narratives can generate different responses of expectations to shocks. As a result, the dynamics of aggregate output depend on the distribution of narratives across households.

Importantly, we also use the model to derive an equivalence result. While there are several possible ways to relate the yield curve to output changes in narratives—as a *shock* affecting future income or as a *signal* of other variables—the resulting DAGs imply the same expectations in each case. This equivalence result implies that we do not need to distinguish between different directions of causation in these narratives in the data. It is sufficient to identify only whether the yield curve inversion is associated with output changes in an article or not.

This result means that standard topic models from natural language processing, which capture groupings of words which tend to appear together, are capable of measuring the aspects of narratives that are relevant for expectations in text. Motivated by this, we use topic models to measure narratives in news articles about the 2019 yield curve inversion. We uncover two competing narratives in major news outlets' coverage, which correspond closely to the narratives in the model:[1] a "recession" narrative that links the inverted yield curve to an imminent recession and a "nonrecession" narrative that does not. We also obtain estimates of how strongly each article makes use of each narrative.

We then study the effects of these narratives on readers who are exposed to them. To do this, we link narratives in newspapers to social network data from Twitter, creating a novel data set that combines narratives in newspapers, Twitter users who are exposed, tweets of these users, and tweets of their followers. We use retweeting activities on Twitter

[1] This is a result, not an assumption: the topic model generates topics, which we interpret as narratives, without our guidance.

to trace whether a consumer has engaged with news articles containing certain narratives. We find that tweets posted by users exposed to the recession narrative become significantly more negative after the exposure, while tweets posted by users exposed to the more neutral narrative display no such changes. The magnitude of the sentiment decline is enough to offset the effect of a positive release of the jobs report, another closely watched macroeconomic indicator.

This differential effect of the two narratives is robust to a range of checks and tests. In particular, we find that before the day of exposure, there is no difference in the sentiment of readers in each group, and no trends in their sentiment over time. The drop in sentiment following engagement with a recessionary narrative is persistent, remaining significant 30 days after the retweet.

Finally, we further leverage the network structure of Twitter to assess the hypothesis in Shiller (2017) that narratives are contagious, spreading between people like a virus. We find that the sentiment effects of each yield curve narrative are present not just among those who engage with the articles directly, but also among their followers. Again, being exposed to a recession narrative leads to declines in sentiment, while exposure to the nonrecession narrative has no effect. The effect of the recession narrative is approximately 40% smaller on followers than on the original sample, suggesting a substantial but not perfect contagion of the narrative.

Related literature Our paper relates to four strands of the literature. First, we contribute to the emerging literature on narratives in economics, pioneered by Shiller (2017).[2] Shiller (2017, 2020) shows that perennial economic narratives spread across the economy in a viral way. The power of these narratives may come especially from collective memory and recall of rare disasters (Goetzmann, Kim and Shiller, 2022). Our paper provides direct evidence that narratives, once they have spread in the media, go on to affect the sentiment of those exposed to them. The viral spread of narratives, combined with the substantial effects on sentiment we find, could therefore generate epidemiological dynamics in expectations and sentiment. Burnside, Eichenbaum and Rebelo (2016), Flynn and Sastry (2022), and Carroll and Wang

[2]Also see the body of work that highlights importance of political narratives, which includes, for example, Gentzkow, Shapiro and Sinkinson (2014), Levy (2021), Bianchi, Kung and Cram (2021), and Eliaz, Galperti and Spiegler (2022).

(2023) show that such dynamics have important consequences for aggregate fluctuations, and indeed the latter two propose narratives as a potential source of these effects.[3] The evidence we document of the effects of narratives on sentiment, therefore, forms an important link in the transmission of narratives to macroeconomic fluctuations.

Theoretically, our model builds on the recent literature formalizing narratives as DAGs in microeconomic theory (Spiegler, 2016, 2020; Eliaz and Spiegler, 2020), taking this to a dynamic general equilibrium context. Empirically, we develop a text-based measure of competing narratives that is directly connected to the theoretical framework, and link this to rich social media microdata for assessing the impacts on sentiments. This complements semantics-based approaches that attempt to capture causal directions in textual narratives (e.g. Ash, Gauthier and Widmer, 2021; Goetzmann et al., 2022), and experimental evidence on household responses to narratives (Andre et al., 2022b; Kendall and Charles, 2022). Closely related to our methodology of topic models, Larsen and Thorsrud (2019) study the effects of narratives on business cycle fluctuations, defining narratives as prominent topics in a corpus of newspaper articles. We instead capture narratives as news media's competing interpretations of the *same* underlying economic event, which allows us to separate the effect of the narrative from the effect of information on the event.

Second, we contribute to the literature studying the macroeconomic implications of news media. Empirically, several recent papers have, like us, used text data to study the economic effects of news reporting (see, for example, Calomiris and Mamaysky, 2019; Bybee, Kelly, Manela and Xiu, 2020; Nyman, Kapadia and Tuckett, 2021). Other work in this literature has focused on the effects of selective news reporting, which affects the economy by influencing the information sets of agents (Nimark, 2014; Chahrour et al., 2021; Bui, Huo, Levchenko and Pandalai-Nayar, 2022). We extend this literature by studying the effect of the narratives provided in media articles, beyond the effects of the information provided.

Third, our approach to measuring the effects of media narratives involves linking news articles with behavior on social media. The key advantage of this is that we can therefore study reactions to narratives at the individual level, which enables us to compare the effects of different narratives about the same economic event. This therefore contributes to the

[3] See also the large theoretical literature on sentiments in macroeconomics, surveyed in Angeletos and Lian (2016). Recent empirical contributions include Angeletos, Collard and Dellas (2018), Levchenko and Pandalai-Nayar (2020), and Lagerborg, Pappa and Ravn (2022).

literature that uses social network data to study the effects of policy (see, for example, Bailey, Cao, Kuchler and Stroebel, 2018; Gorodnichenko, Pham and Talavera, 2021; Bianchi et al., 2021; Matveev and Ruge-Murcia, 2021; Haldane, Macaulay and McMahon, 2021; Ehrmann and Wabitsch, 2022). In particular, we extend this literature by linking traditional and social media data, which could also be a useful method for further research.[4]

Finally, narratives provide a way for individuals to interpret economic news and translate that into expectations. We therefore also relate to the broad literature on belief formation. Empirically, a large literature documents evidence of deviations by households and firms from full-information rational expectations (see Coibion, Gorodnichenko and Kamdar, 2018, for a comprehensive survey). Previous literature points to inattention (Sims, 2003; Mankiw and Reis, 2002), personal experiences (Malmendier and Nagel, 2016), salience (Cavallo, Cruces and Perez-Truglia, 2017), heuristics (Bordalo, Gennaioli and Shleifer, 2018), wishful thinking (Caplin and Leahy, 2019), among others, as important drivers of individuals' expectations. We provide empirical evidence on the importance of narratives, particularly in the context of the yield curve.[5]

Outline The rest of the paper proceeds as follows: in Section 2, we present our theoretical framework that connects narratives with expectations; in Section 3, we use the model to derive results that inform the measurement of narratives and their effects; in Section 4, we describe our data and text analysis methodology; in Section 5, we conduct our main empirical analysis on the narratives surrounding the yield curve inversion; in Section 6, we study the contagion of those narratives; Section 7 concludes.

2. Model

We now develop a simple model in which household expectations are formed using narratives, which may or may not involve variables such as the slope of the yield curve. The purpose of the model is twofold: it guides our strategy for measuring narratives in newspapers, and it informs the interpretation of our empirical results.

[4]See, for example, the companion paper Macaulay and Song (2023), which applies a similar method to study how media and households interpret inflation.

[5]For other work on beliefs and the yield curve, see e.g., Bauer and Chernov (2021), Bauer, Pflueger and Sunderam (2022), Leombroni, Vedolin, Venter and Whelan (2021).

2.1. Defining Narratives

We begin by stating the precise definition of narratives we will use in this paper. The key feature of this definition, common to the definition in many English dictionaries, is that a narrative involves a *causal* account of how variables or events relate to each other.[6] This prominent role for causality can be seen, for example, in the "perennial economic narratives" highlighted by Shiller (2020), which include "Labor-Saving Machines Replace Many Jobs" and "The Wage-Price Spiral and Evil Labor Unions."

To capture this aspect of narratives, we follow Eliaz and Spiegler (2020) and Andre et al. (2022b) in formalizing narratives as *directed acyclic graphs* (DAGs).[7] A given DAG, or narrative, is characterized by a series of causal relationships between variables.

***Definition* 1** (narrative as a DAG)**.** *A narrative is defined as a DAG consisting of:*

1. *a set of nodes $\mathcal{N}$, where each element is a real-valued economic variable; and*
2. *a set of links $\mathcal{L}$, which define the directed causal links between nodes.*

such that the links $\mathcal{L}$ are acyclic: the graph contains no directed path from a node back to itself.

As well as increasingly being used to capture narratives in economics, DAGs are common in computer science and statistics (Koller and Friedman, 2009; Pearl, 2009). They have recently been used to analyse identification in applied econometrics (Hünermund and Bareinboim, 2023), and to aid the solution of heterogeneous-agent models in macroeconomics (Auclert, Bardóczy, Rognlie and Straub, 2021).

2.2. Environment

Time is discrete. The economy consists of households, firms, and a central bank. The focus of the model are households, who form expectations using narratives, or DAGs. Our results do not rely on whether the true model can be expressed as a DAG: we only require that households use DAGs to form expectations.

[6] See the discussions in (among others) Eliaz and Spiegler (2020), Shiller (2020), Andre et al. (2022b), and Goetzmann et al. (2022). The latter, for example, state that "Narratives often elicit causal relationships between a sequence of events connecting one state to another with things that happen in between."

[7] For a thorough review of the use of DAGs in modeling expectations, see Spiegler (2020).

2.3. Households

A continuum of households receive real income, y_t, and can save or borrow in one-period bonds with a real interest rate of r_t. Household i chooses consumption, c_{it}, to maximize the expected present value of CRRA utility. Since this problem is standard, we begin with the consumption function log-linearized about a steady state with zero asset holdings (see e.g., Bilbiie, 2019):

$$c_{it} = (1-\beta)\sum_{s=0}^{\infty}\beta^s\,\mathbb{E}_{it}\,y_{t+s} - \sigma\beta\sum_{s=0}^{\infty}\beta^s\,\mathbb{E}_{it}\,r_{t+s}, \tag{1}$$

where $\beta \in [0,1]$ is the discount factor, and $\sigma > 0$ is the elasticity of intertemporal substitution. The operator $\mathbb{E}_{it}$ denotes the expectations of household i in period t. It is here that narratives enter the model.

Households are members of large families, which redistribute wealth between members every period. This ensures that any heterogeneity in narratives does not lead to heterogeneity in wealth. Although interesting, this interaction between narratives and wealth is beyond the scope of this paper. We assume that households act as if all family members use the same narrative as they do. This means that household i does not need to adjust their consumption function (1) due to this intra-family redistribution.[8]

Information. Households have full information on current and past realizations of y_t, r_t, and z_t.

Narratives. To form expectations of future income and interest rates, households combine their information set with a belief about the evolution of both variables. That belief comes from a narrative.

We begin by considering just two possible narratives, which are common in discussions of yield curve inversions and illustrate how narratives may affect equilibrium outcomes. We widen the analysis to more narratives in Section 3 below. The two DAGs are drawn in Figure 1 and defined in Definition 2. For simplicity, they both abstract from links between interest

[8] Alternatively, we could have a large household forming a consensus forecast by averaging over many household members, who use heterogeneous narratives, and then choosing consumption based on these consensus expectations. Since the model is linear, that setup gives the same aggregate consumption as the one presented here.

rates r_t and other variables. The key difference between them is that in the first ("baseline") narrative, output only depends on lagged output, while in the second ("z") narrative, it can also depend on an extra variable, z_t. In our empirical analysis, z_t will reflect the slope of the yield curve. Under the z narrative, changes in output affect z_t, implying it could be a signal of imminent recessions.

Figure 1: DAG representations of baseline and z narrative

(a) Baseline narrative **(b)** z-narrative

***Definition* 2** (baseline and z narratives)**.** *Let nRm denote a directed link from node n to node m. The baseline and z narratives consist of the set of nodes, $\mathcal{N} = \{r_s, y_s, z_s\}_{s=t}^{\infty}$; and a set of links $\mathcal{L}^k$, where $k \in \{b, z\}$ denotes the baseline and z narrative respectively:*

$$\mathcal{L}^b = \{y_s R y_{s+1}\} \tag{2}$$

$$\mathcal{L}^z = \{y_s R y_{s+1}, y_s R z_s, y_{s+1} R z_s\} \tag{3}$$

To go from narratives to expectations, we note that each narrative dictates which information households should condition on when forming expectations.[9] Specifically, using the baseline narrative, expectations of real income one period ahead are not conditioned on z_t, because it is independent of y_{t+1} in the narrative:[10]

$$\mathbb{E}_t^b(y_{t+1}|\mathcal{I}_t) = \int y_{t+1} p(y_{t+1}|y_t) dy_{t+1}. \tag{4}$$

[9]Formally, any DAG implies a set of conditional independence assumptions about the nodes $\mathcal{N}$ (Spiegler, 2020).

[10]In both narratives, r_{t+1} is independent of all other variables, so $\mathbb{E}_{it}(r_{t+1}|\mathcal{I}_t) = \int r_{t+1} p(r_{t+1}) dr_{t+1} = 0$ in all time periods: households act as if r_t is i.i.d. This is not critical for the immediate results in this subsection, but it allows us to solve for general equilibrium analytically in Section 2.5.

In contrast, using the z narrative, the same expectation is conditioned on z_t:

$$\mathbb{E}_t^z(y_{t+1}|\mathcal{I}_t) = \int y_{t+1} p(y_{t+1}|y_t, z_t) dy_{t+1}. \tag{5}$$

Since households have access to the full history of data on each variable, they are able to estimate the true likelihood functions $p(.|.)$ involved in their narrative. The likelihoods in equations (4) and (5) are, therefore, determined in equilibrium and are only subject to bias from incorrect independence assumptions encoded in each narrative.

This approach to mapping narratives to expectations directly follows Eliaz and Spiegler (2020). Households accurately fit their own narrative to data, but they do not estimate the alternative narrative, or engage in comparisons between the narratives at any point. The question of how and why particular narratives spread is discussed extensively in (among others) Shiller (2020), and we abstract from those mechanisms here.

2.4. Closing the model

Since the households are the focus of this model, we keep the rest of the model simple, embedding the households in a reduced-form variant of the textbook New Keynesian model in Galí (2008). Inflation π_t is determined by a simple Phillips curve with myopic firms, and interest rates r_t are chosen according to a Taylor rule:[11]

$$\pi_t = \kappa \cdot mc_t + v_t^\pi, \tag{6}$$

$$r_t = \phi \cdot \pi_t + v_t^r, \tag{7}$$

where mc_t denotes a firm's marginal costs, $v_t^\pi \sim N(0, \sigma_\pi^2)$ and $v_t^r \sim N(0, \sigma_r^2)$ are i.i.d. shocks, and $\kappa > 0$ and $\phi > 1$ are parameters related to the slope of the Phillips curve and the Taylor rule, respectively.

We specify that marginal costs are increasing in output, and that all income from production flows equally to households, so real income, y_t, is equal to real output. This specification could be microfounded, for example, by adding a labor supply choice to the household problem, and assuming production takes place using labor as the only input.

[11] Typically this rule would be specified in terms of nominal, rather than real, interest rates. However, as with r_t, π_t is not involved in household narratives, and so $\mathbb{E}_{it}\, \pi_{t+1} = 0$ and real and nominal interest rates coincide.

In addition, we also allow the extra variable, z_t, to potentially affect marginal costs with a lag. This captures the possibility outlined in the z narrative that z_t signals future output. The marginal cost process is

$$mc_t = y_t + \mu z_{t-1}, \tag{8}$$

where the parameter μ determines the effect of z_t on marginal costs. This effect could come from z_t signaling changes in future productivity or financial frictions.

For the goods market to clear, output must equal aggregate consumption each period. Letting λ denote the proportion of households using the z narrative, the market clearing condition is

$$y_t = (1 - \lambda)c_t^b + \lambda c_t^z, \tag{9}$$

where c_t^b and c_t^z denote the consumption of households using the baseline and z narratives, respectively.

Finally, we specify a process for z_t

$$z_t = \chi y_t + v_t^z, \tag{10}$$

where χ is a parameter and $v_t^z \sim N(0, \sigma_z^2)$ is an i.i.d. shock.

Notice that with this model specification, neither baseline nor z narrative paints a full picture of the economy, because both ignore the roles of interest rates and inflation in determining output. However, the only source of persistence is z_{t-1}. This means that a household with a complete understanding of the economy (i.e., with rational expectations) would only condition expectations of future output on z_t, and not on anything else. Despite the fact that their simple narrative misses many relationships in the model, households using the z narrative do therefore condition on the relevant information when forming their expectations of future output.

2.5. Narrative equilibrium

***Definition* 3** (narrative equilibrium)**.** *Given a distribution of households across narratives λ, the endogenous state z_{t-1}, and shocks v_t^π, v_t^r, v_t^z, a narrative equilibrium consists of $c_t^b, c_t^z, \pi_t, r_t, y_t, z_t$, and expectations $\mathbb{E}_{it}^b(r_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_{it}^b(y_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_{it}^z(r_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_{it}^z(y_{t+s}|\mathcal{I}_t)$, such that:*

1. *Given prices and expectations, households choose c_t^b, c_t^z according to (1);*
2. *Inflation π_t and the interest rate r_t are determined according to equations (6) and (7);*
3. *Marginal costs mc_t are determined according to equation (8);*
4. *z_t is determined according to (10);*
5. *The goods market clears according to (9);*
6. *Expectations are determined according to (4) and (5), where the likelihood functions $p(\cdot|\cdot)$ are consistent with the relevant true likelihoods.*

Since households form expectations by fitting misspecified models (narratives) to long histories of data, this is an example of the Constrained-Rational Expectations Equilibrium introduced by Molavi (2019). Solving for the likelihoods used in each narrative in this equilibrium involves a system of nonlinear equations with no general analytic solution. However, Proposition 1 considers a special case in which the system can be studied analytically.

***Proposition* 1.** *If $(1-\beta)\kappa\phi\sigma < 1$ and $\mu\chi \in (-1,1)$, then in the limit as $\sigma_v^2 \to 0$ there exists a unique stable equilibrium. In that equilibrium*

$$\mathbb{E}_t^z\, y_{t+1} = \mathbb{E}_t^b\, y_{t+1} + \mathcal{G} v_t^z \tag{11}$$

where $\mathcal{G}$ is a combination of model parameters defined in Appendix A, such that

$$\mathcal{G} \begin{cases} = 0 & \textit{if } \mu = 0 \\ \neq 0 & \textit{if } \mu \neq 0 \end{cases} \tag{12}$$

In addition,

$$\frac{\partial y_t}{\partial \lambda} = \mathcal{H} v_t^z \tag{13}$$

where $\mathcal{H}$ is a combination of model parameters defined in Appendix A, such that

$$\mathcal{H} \begin{cases} = 0 & \text{if } \mathcal{G} = 0 \\ > 0 & \text{if } \mathcal{G} > 0 \\ < 0 & \text{if } \mathcal{G} < 0 \end{cases} \tag{14}$$

Proof. Appendix A. □

The first two restrictions for this special case are weak restrictions on the parameter space.[12] The third is what makes it possible to solve for the equilibrium analytically. Since households with the baseline narrative estimate the distribution of y_t conditional on y_{t-1} only, their estimates are subject to an omitted variable bias which depends on the relative variances of z_{t-1} and y_{t-1}. Considering $\sigma_v^2 \to 0$ removes this variance ratio and simplifies the equilibrium substantially. Economically, this special case is the limit as exogenous shocks to the extra variable become small.

Proposition 1 shows that it is theoretically ambiguous whether narratives generate different expectations. In the case without a fundamental channel from z_t to y_{t+1} (if $\mu = 0$), the two narratives deliver identical expectations in equilibrium. In this environment, v_t^z would amount to a non-fundamental sunspot, and equation (12) verifies that this is ruled out by the narratives fitted to data in equilibrium.

However, in the case with a fundamental connection between z_t and y_{t+1} (if $\mu \neq 0$), the unique equilibrium features expectations that differ across narratives whenever a shock to z_t occurs. This is because households using the z narrative condition their expected income on realized z_t, and so react to that shock beyond its impact on y_t. Households using different narratives form different expectations, which lead to different consumption decisions. As a direct consequence, changes in the distribution of narratives across households affect output

[12] At a quarterly frequency β is typically very close to 1, and common estimates of σ are typically around 0.5 (Havránek, 2015), so the first restriction will be satisfied as long as the Phillips curve and Taylor rule are not extremely steep. The second restriction is that when combining equations (8) and (10), lagged output does not have more powerful effects on mc_t than current y_t.

whenever there is a shock to z. The resulting changes in output are larger if a greater proportion of households hold the z narrative (i.e., if λ is high).

Proposition 1, therefore, shows that cases exist where different narratives to have differential effects on expectations. If they do, the spread of one particular narrative over another can have consequences for aggregate consumption and output. However, the model does not guarantee that narratives have any such effects. We therefore turn to data to distinguish whether or not narratives affect households in practise.

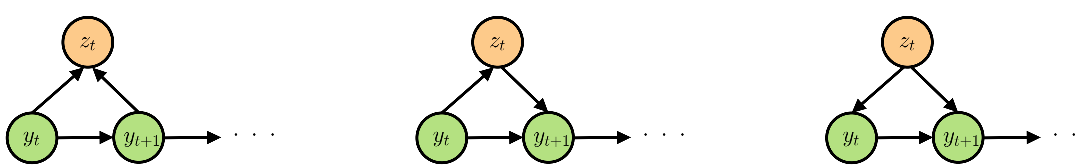
3. Taking the Model to the Data

Before we begin our empirical analysis, we first derive two results from the model that help to connect it into the data that we study in Section 4.2.

3.1. Expanding the set of z narratives

The z narrative in Figure 1 and Definition 2 is just one of several possible ways to construct a narrative in which the extra variable z_t is related to changes in output. When we turn to news articles in which z_t is the inversion of the yield curve, there could be other narratives present. We therefore expand the set of narratives considered in the model to include two other possibilities, displayed in Figure 2.[13]

Figure 2: DAG representations of expanded z-narratives



(a) “Signal” **(b)** “Channel” **(c)** “Shock”

Panel (a) shows the original z narrative from Section 2, in which the yield curve inversion is caused by the fact a recession is coming. Panels (b) and (c) show alternative narratives, in which z_t is still related to current and future output, but with different causal mechanisms. In Panel (b), z_t is a channel through which y_t affects y_{t+1}, and in Panel (c), z_t causes changes

[13]The nodes r_t, r_{t+1} have been omitted from Figure 2 for brevity, as they are not connected to any other node.

in output. This final case could capture a narrative in which a yield curve inversion causes a recession, potentially because of the way banks or other agents react to it. The full set of these z narratives is defined formally as follows:

***Definition* 4** (z narratives)**.** *The expanded set of z narratives are DAGs consisting of:*

1. *the set of nodes,* $\mathcal{N} = \{r_s, y_s, z_s\}_{s=t}^{\infty}$*; and*

2. *one of the sets of links* $\mathcal{L}_a$*,* $\mathcal{L}_b$*, or* $\mathcal{L}_c$*, where:*

   (a) $\mathcal{L}_a = \mathcal{L}^b \cup \{y_s R z_s, y_{s+1} R z_s\}$;

   (b) $\mathcal{L}_b = \mathcal{L}^b \cup \{y_s R z_s, z_s R y_{s+1}\}$;

   (c) $\mathcal{L}_c = \mathcal{L}^b \cup \{z_s R y_s, z_s R y_{s+1}\}$;

These are distinct narratives, with individual assumptions about causal mechanisms. However, Proposition 2 shows that they imply exactly the same expectations in all states of the world.

***Proposition* 2** (observational equivalence of z narratives)**.** *For the three z narratives* $\{a, b, c\}$ *in Definition 4:*

$$\mathbb{E}_t^a(x_{t+h}|\mathcal{I}_t) = \mathbb{E}_t^b(x_{t+h}|\mathcal{I}_t) = \mathbb{E}_t^c(x_{t+h}|\mathcal{I}_t)$$

for any variable x *at any horizon* h*, and an information set* $\mathcal{I}_t$ *consisting of any realizations of* $\{r_s, y_s, z_s\}_{s=0}^{t}$*.*

Proof. Appendix A. □

This equivalence occurs because, despite their different causal mechanisms, the three z narratives share the same set of conditional independence assumptions. The different varieties of z narrative are therefore observationally equivalent for expectations, and therefore for actions. Formally, this is a consequence of the fact they all have a property known as "perfection" (Verma and Pearl, 1990).

This is a key lesson for our empirical analysis. Proposition 2 implies that we do not need to distinguish between these varieties of z narrative to capture the effect of narratives on expectations. Rather, it is sufficient to categorize narratives based on whether the yield curve inversion is associated with changes in output or not. This motivates our use of off-the-shelf "bag-of-words" models from natural language processing, which uncover whether words appear together in a text, but discard the semantic relationship in which those words are linked. Our goal will be to separate the set of z narratives from the baseline narrative, in which the yield curve inversion is thought to be irrelevant. For simplicity, from here we use "the z narrative" to refer to any narrative in the set in Definition 4.

3.2. Expectations and sentiment

Our empirical analysis will focus on a measure of tweet sentiment, rather than the expectations that drive dynamics in the model. However, this sentiment measure is connected to household expectations, both in our theoretical framework and in the broader literature.

This connection in the model is demonstrated in Lemma 1, which shows that all expectations of interest are proportional to a single variable: the household's expectation of one-period ahead output.

***Lemma* 1** (common factor of expectations)**.** *Household expectations under narrative beliefs can be written:*

$$\mathbb{E}_t^k\, x_{t+s} = \Gamma_{k,s}\, \mathbb{E}_t^k\, y_{t+1} \tag{15}$$

Where $\Gamma_{k,s}$ *is a constant independent of variable realizations,* $s > 0$*, and* $k \in \{b, z\}$*.*

Proof. Appendix A. □

This is a natural consequence of the recursive structure of DAGs: once the household forms an expectation for output one period in the future, all other expectations then follow from that. This common factor behind expectations can, therefore, be thought of as the household's overall level of optimism or pessimism, which is what our empirical measure of tweet sentiment aims to capture. Similarly, the well-known sentiment index from the Michigan Survey of Consumers is an overall level of optimism or pessimism derived by combining many different expectations (see Lagerborg et al., 2022, for a recent application). Theoretical models in which a single sentiment-like factor drives expectations across many variables have also been successful in explaining a range of features of macroeconomic, survey, and financial data (Kamdar, 2019; Molavi, 2019; Molavi, Tahbaz-Salehi and Vedolin, 2021; Andre, Pizzinelli, Roth and Wohlfart, 2022a).

Notably, this differs from the definition of sentiments in e.g., Angeletos and La'O (2013) or Acharya, Benhabib and Huo (2021), where sentiments are self-fulfilling beliefs orthogonal to macroeconomic fundamentals. In our model, sentiment is the common factor determining all expectations, so it will naturally be correlated with fundamentals whenever expectations react at all to the state of the economy. Our empirical measure of sentiments are similarly allowed to be influenced by fundamentals, consistent with this feature of the model.

Overall, there are three important observations from this theoretical framework that guide our empirical analysis. First, it is theoretically ambiguous whether narratives have differential effects on expectations or not. The importance of narratives is, therefore, an empirical question. Second, families of z narratives are observationally equivalent. This allows us to use bag-of-words models to distinguish z narratives from baseline narratives. Lastly, the model makes it explicit that any narrative DAG is associated with particular likelihood functions, obtained by fitting the narrative to data. The two parts are both integral to the formation of expectations, and our empirical analysis will capture the combined effects of both.

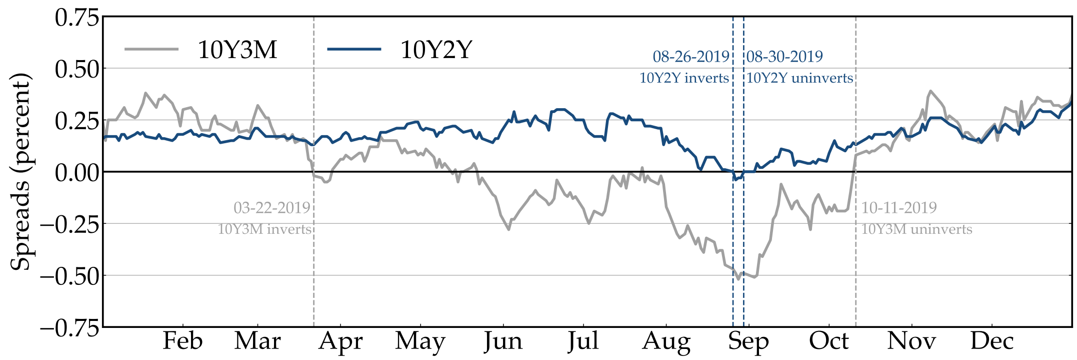
4. Data and Methodology

4.1. Background

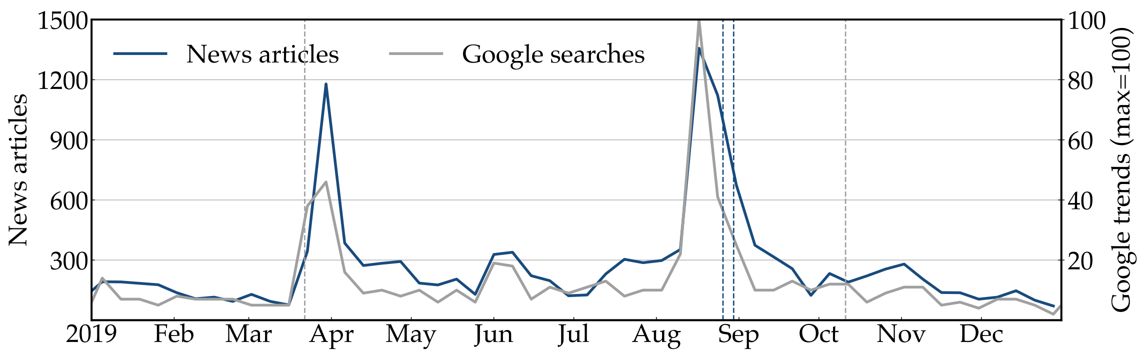
The theoretical framework suggests that narratives may have an effect on household sentiment. We now set out to study their empirical importance. To separate narratives from the underlying event being reported, we focus on an episode of yield curve inversion in 2019 and compare the sentiment effects of different narratives about that one event.

Yield curve inversions have been a closely-watched indicator of upcoming recessions in the U.S. since Harvey (1988) documented their predictive power from the 1960s to the 1980s. Figure B.1 in the Appendix shows that the spread between the 10-year and 2-year Treasury bond yields has turned negative within 12 months before every recession in the US for the past 40 years. However, despite this track record, there have also been false-positive signals, such as 1966. The spread between long-term and short-term treasury yields is influenced by

Figure 3: Timeline of the yield curve inversion episode



(a) Treasury spreads



(b) Media coverage and Google searches for “yield curve”

Notes: Panel (a) shows the spread between 10-year treasury yield and 3-month treasury yield (“10Y3M”) and the spread between 10-year treasury yield and 2-year treasury yield (“10Y2Y”) in 2019. Dates when the spreads first turn negative and revert back to positive are annotated. Panel (b) shows the number of news articles from Factiva containing the term “yield curve” and the Google search frequency in 2019. Google search frequency for the term “yield curve” has been scaled so the maximum value is 100.

investors’ expectations of monetary policy and risk factors, along with other factors, and does not predict a recession with certainty (as emphasized, for example, in Bauer and Mertens, 2018).

When the yield curve inverted in 2019, it received substantial attention from households and the media. Figure 3a plots the timeline of the inversion, showing that the most widely-watched 10-year-over-2-year (10Y2Y) term spread inverted on August 28 and un-inverted on August 30. Figure 3b shows that media coverage[14] and Google searches for the term “yield curve” spiked before and during the inversions of both the 10Y2Y term spread and the 10-

[14]We measure media coverage using weekly data from Factiva. We obtain the number of nonduplicate news articles containing the term “yield curve” and restrict articles to be in English and specific to the US.

year-over-3-month (10Y3M) term spread, with a peak of interest right before the inversion of the 10Y2Y spread.

Against the backdrop of a booming labor market and the longest expansion in US history, the inversion received several different interpretations in the media. The first interpretation is that a recession is looming. An example of such a recession narrative is Cristina Alesci's article for CNN[15]:

> Navarro is wrong on two fronts: The inversion did happen, and it's not a good sign for the economy. Although the inversion was brief and small, major banks took note of it. [...] Yield curve inversions often signal recessions, which is why economic prognosticators pay so much attention to them.

which draws on the track record yield curve inversion to predict a recession and paints a negative picture on the economic outlook. Notably, the argument draws on both the "signal" narrative in Figure 2 ("inversions often signal recessions") and the "shock" narrative ("major banks took note of it"). This highlights the intuition for Proposition 2: both of these narratives imply readers should update their expectations towards believing a recession is likely. It also underlines the importance of Proposition 2 for our empirical exercise, as it implies we do not need to disentangle these often-combined narratives to estimate the effects of the narrative on expectations.

The second common interpretation is that the yield curve inversion is no longer an informative signal. Peter Coy illustrates such a narrative for Bloomberg[16]:

> Well, guess what, folks? It's still rainbows and pots of gold out there. Contrary to what seems to have become the overnight conventional wisdom in politics, a recession before Election Day 2020 remains a less than 50-50 proposition.

which goes on to explain that the long end of the yield curve has been trending down because of low and stable inflation and the strong fundamentals of the economy, suggesting that recession concerns are overblown. This corresponds to the "baseline narrative" in section 2.

The articles by Cristina Alesci and Peter Coy are strong examples of each of these narratives. Some other media reports are less stark, presenting a more balanced view of the

[15] "Fact-checking Peter Navarro's claims that the yield curve is not inverted" by Cristina Alesci on August 19, 2019. Link to the article on CNN.

[16] "What a Yield-Curve Inversion Really Says About the U.S. Economy: A reliable recession indicator has lost some of its power to predict" by Peter Coy on August 22, 2019. Link to the article on Bloomberg.

yield curve inversion, with a mix between the two narratives. Our measurement of narratives detailed below is able to account for such mixed articles, as well as those that lay out a single narrative.

4.2. Data

4.2.1. Newspaper articles

To form the media corpus for our analysis, we collect news articles covering the inversion of the 10Y2Y spread. Our data source is Factiva, a news database, and news outlets' websites. To separate the effects of economic narratives from political narratives, we focus on news outlets classified as "centrist" by the Pew Research Center and exclude news aggregators such as Google News.[17] The 10 news outlets included in our sample in listed in Table 1.

Table 1: Media outlets and coverage on the yield curve inversion

Outlet	**Ideology placement**	**Twitter handle**	**# base tweets**	**# articles**
MSNBC	Liberal/Center	msnbc	4	1
CNN	Liberal/Center	cnn	8	4
NBC News	Center	nbcnews	4	1
CBS News	Center	cbsnews	3	3
Bloomberg	Center	business	143	68
ABC News	Center	abc	1	1
USA Today	Center	usatoday	1	1
Yahoo News	Center	yahoonews	3	3
Wall Street Journal	Center	wsj	9	6
Fox News	Conservative/Center	foxbusiness	0	0
Total			176	88

Notes: Media outlets with centrist political leaning and their coverage of the yield curve inversion. Data on media outlets' political placement is from (Jurkowitz et al., 2020), which determines the political ideology of an outlet by surveying the political leaning of its audience. The twitter handles of news outlets are hand searched. The tweets and articles on the yield curve are collected as described in Section 4.2.

During the event window of August 19 to September 13, 2019 (one week before the inversion and two weeks after the un-inversion, respectively)[18], we search for tweets by news

[17] Jurkowitz, Mitchell, Shearer and Walker (2020) determine the political bias of a media outlets by surveying the political ideology of its audience.

[18] Although the yield curve was inverted from August 26 to August 30, media coverage and Google search trends in Figure 3b suggest that the interests in the yield curve rose before the actual inversion and stayed elevated after the un-inversion. Therefore, we expand the search window for news articles to one week before the inversion and two weeks after the un-inversion.

outlets which contains both "yield curve" and any of the stems from "invert", "invers", or "recession". These "base tweets" by news outlets contain URLs to their webpages containing the full-length news articles, which form the corpus from which we extract narratives. Table 1 shows that the search criteria lead to 176 base tweets, linking to 88 unique articles. The majority of these are from Bloomberg, who devoted many more articles to the yield curve inversion than other outlets. However, within Bloomberg there is a diverse range of journalists, who put forward a diverse range of narratives.

4.2.2. Twitter data

Our Twitter data consists of four parts. First, as described in the last subsection, we use outlet's base tweets to identify news articles related to the yield curve inversion. We use Twitter's API to query the full archive of tweets since the start of Twitter in 2006.

Second, when a user interacts with a tweet (by "quote retweeting", "retweeting", "replying" or "liking"), it leaves a trace to measure the exposure to narratives. Among the four methods of interaction, we focus on quote retweets, which require that a user writes additional text when retweeting. Importantly, for this method of interaction Twitter records a timestamp of precisely when the quote-retweet occurred, allowing us to construct narrow event time windows around the narrative exposure.[19] In addition, the commentaries added by quote retweeters makes it more plausible that the users have digested the new information contained in the articles. For each base tweet, we therefore obtain the users who quote retweeted, and the time that they did so. The Twitter API only provides the first 100 such users for each base tweet, but this limit only binds in 1 out of the 178 base tweets in our sample. Table 2a summarizes the retweeting activities of the base tweets on the yield curve. On average the base tweets in the sample have 9 quote retweets, and the 95 percentile has 28 quote retweets.

Third, we measure changes in users' tweet sentiment after they are exposed to a narrative by measuring the sentiment of their tweets on all subjects. For users who have quote retweeted any of the base tweets on the yield curve, we collect every tweet posted in a 1-month window around the quote retweet. Table 2b reports descriptive statistics of tweeting

[19] For likes, and retweets without additional commentary, Twitter only provides the time of the original tweet but not the time when the like or retweet occurred. This makes it infeasible to determine the time of the exposure.

Table 2: Descriptive statistics on base tweets and retweeting users

(a) Outlets' base tweets on the yield curve

	Mean	**SD**	**5th Pctl**	**Median**	**95th Pctl**	**Obs**
Quote retweet count	8.5	39.1	0	3	28.2	178
Retweet count	45.4	89.9	0	23	162.6	178
Reply count	8.8	25.0	0	4	25.3	178
Favorite count	67.4	120.6	0	35	235.8	178

(b) Quote retweeting users

	Mean	**SD**	**5th Pctl**	**Median**	**95th Pctl**	**Obs**
All quote retweeters						
# tweets	64.4	249.1	0.1	10.6	356.1	404
# outlets	3.5	2.5	1	3	8	404
# followers	3,562	14,720	13	523	11,120	404
Active quote rewteeters during event windows						
# tweets	73.6	276.1	0.2	12.1	279.6	324
# outlets	3.7	2.5	1	3	8	324
# followers	2,304	7,324	10	554	8,353	324

Notes: Panel (a) reports descriptive statistics of media outlets' tweets about the yield curve inversion between August 19 and September 13, 2019. The table reports descriptive statistics of the numbers of quote retweets, retweets, replies and favorites of media outlets' tweets. Panel (b) reports descriptive statistics of users' Twitter activity based on tweets one month before and one month after the quote retweets of the base tweets. The top panel includes the full sample. The number of tweets represent the daily average. The number of outlet appearing in a users timeline is counted over the sample period. The number of followers are reported as of our data-collection date of October 2021. The bottom panel includes users that enter our regression analysis. A user is active during the event window if the user has posted tweets both the day before and the day after the quote retweet.

activity for the users in our sample, which shows that the median user is active and posts around 10 tweets per day. We measure changes in sentiment in one-day windows surrounding the exposure, which requires a user to be active during the event windows and post at least one tweet in the days before and after the exposure. This restricts our sample to 324 unique users. Our analysis is at the retweet level. 17 users quote retweet more than once and appear in the sample with each retweet.

Lastly, we use the social network structure to study the contagion of narratives. Table 2b shows a large variation in the number of followers. The top 5% quote retweeters have more than 11,000 followers, while the bottom 5% quote retweeters have less than 13. For users that have quote retweeted a news article, we observe the list of their followers. We randomly sample 200 followers when the follower count exceeds the threshold. We then collect every tweet posted by these followers in the days surrounding their friends' quote retweet.

4.3. Methodology

In this section, we use tools from natural language processing to construct sentiment measures and capture narratives.

4.3.1. Measuring tweet sentiment

We measure the sentiment of a tweet using a naïve Bayes classifier trained specifically to analyze the colloquial language on Twitter (for more details see Appendix D).[20] The sentiment score measures the probability that a tweet conveys positive sentiment and is a uniform scale between 0 and 1. A score greater than 0.5 corresponds to positive sentiment, and a score less than 0.5 corresponds to negative sentiment. To validate the sentiment measure, we present in Appendix Table B.1 the top 5 positive and negative tweets related to the yield curve.

4.3.2. Measuring narratives with topic models

As the theoretical framework in Section 2 illustrates, the distinguishing feature between narratives is their network structures. CNN's "fact checking Navarro" presents a direct causal connection between the yield curve inversion and macroeconomic output, corresponding to a "z narrative". Bloomberg's "rainbows and pots of gold," on the other hand, dismisses the possibility of the inversion predicting an imminent recession. Under this "baseline narrative", the yield curve inversion is disconnected from output and incomes.

We extract these economic narratives from news articles using latent Dirichlet allocation (LDA) (Blei, Ng and Jordan, 2003, and see Appendix C for details).[21] LDA is a Bayesian factor model that uncovers topics in the articles and represents each article in terms of these topics. It reduces the dimensionality of the text from the entire corpus of articles to just K "topics", or groupings of words that tend to appear together. To uncover these topics,

[20] As recognized by Buehlmaier and Whited (2018), naïve Bayes is one of the oldest tools in natural language processing and has better out-of-sample performance in text-based tasks than alternative models (Friedman, Hastie, Tibshirani et al., 2001).

[21] Also see Hansen, McMahon and Prat (2018) for a discussion on LDA and its application in macroeconomics.

it replies on specialized vocabulary that are unique to each topic (for example, "risk" and "recession" versus "rainbow" and "pots of gold") to detect topics in an unsupervised way. Together with these estimated topics, LDA also estimates the loading of article d on topic k, $\theta(d,k) \in (0,1)$, which enables us to analyze both polarizing articles containing a single narrative and balanced articles with multiple narratives.

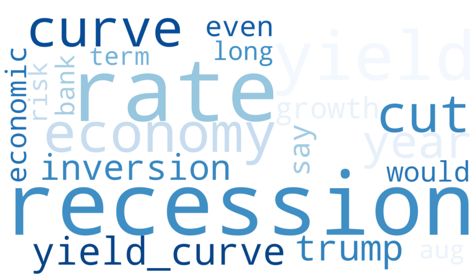
LDA belongs to a broader class of bag-of-words models, which represent individual words irrespective of its surroundings. "Yield curve inversion leads to recession" and "recession leads to yield curve inversion" would have identical representation, since they share word frequencies. It may be surprising, then, that we employ LDA to capture narratives, when the direction of causality is an essential part of a DAG. However, Proposition 2 shows that for the subset of *perfect* DAGs (that share the same conditional independence assumptions), the direction of the causality within a DAG does not affect how a narrative influences expectations. The important difference between narratives for fluctuations is whether phrases such as "yield curve" and "recession" are connected to each other—precisely what LDA is designed to capture—and not the direction of causality between these words. We therefore restrict our attention to DAGs that satisfy the assumptions in Proposition 2. This greatly simplifies the measurement challenge and allows us to capture narratives with simple and interpretable LDA models.[22]

To estimate LDA outputs, we specify uniform Dirichlet priors, as in previous studies using LDA (e.g., Hansen et al., 2018).[23] The remaining parameter that we need to specify is the number of topics K. Our algorithm increments the number of topics from 2 until a topic emerges that does not contain word "recession". LDA is a multi-membership model that allows a word to appear in multiple topics. Since most news articles start with introducing the yield curve inversion as a recession predictor regardless of the narrative, the multi-membership feature of LDA allows for the word "recession" to appear in multiple topics, even when it is not the main thrust of the narrative. We set $K = 5$, the smallest number of topics to ensure at least one topic does not contain the word "recession", which we label as the *nonrecession* narrative. Among the remaining estimated topics, we label the topic with

[22]Recent advances by Ash et al. (2021) and Goetzmann et al. (2022), among others, employ distributed representation of words to capture information embedded in word orderings and show great promises for capturing a broader set of narratives in which the direction of causation may matter.

[23]The pre-processing of texts includes removing stop words and numbers, lemmatizing, and representing the documents with a bigram model.

Figure 4: Economic narratives of the yield curve inversion: LDA outputs



(a) "Recession" narrative

(b) "nonrecession" narrative

(c) Other estimated topics

Notes: This figure reports topics estimated with the LDA model on articles about the yield curve, with $K = 5$ and symmetric Dirichlet priors. The size of a term represent the likelihood for it to appear in a topic. Raw values for this figure are reported in Appendix Table B.2.

the highest probability of the word "recession" appearing as the *recession* narrative.

To ensure that our results are not sensitive to the human labelling of the topics, in Appendix 4.3, we alternatively estimate topics using a guided LDA model, specifying a lexical prior for the first topic to contain the word "recession" rather than a uniform prior as in the baseline LDA. This method automatically detects two topics, one related to recession and one unrelated to it. Appendix Table B.3 shows results under automatic labelling are qualitatively similar as our main results in Table 3.

The estimated topics from the LDA are shown in Figure 4. They represent groupings of words that correspond to the theoretical definitions of the yield curve narratives in Section 2. The first topic in Panel (a) features the terms such as "recession," "yield curve," "economy" and "Trump," mapping naturally to a "recession" narrative, corresponding to the extraneous narrative in our theoretical framework. It discuss the economic policy by the Trump administration in conjunction with the yield curve inversion and recession risks. The second topic in Panel (b) contains a broader discussion of other factors affecting the economy and investment opportunities in the bond and stock markets. Since it does not directly connect the slope of the yield curve to a coming recession, we interpret it as a "nonrecession" narrative, corresponding to the baseline narrative in our theoretical framework. The remaining three estimated topics are reported in Panel (c) for completeness.

We verify the performance of the model in capturing the narratives conveyed in news articles. For Peter Coy's article discussed in Section 4.1 that argues the yield curve has lost its predictive power, the model estimates a loading of $\theta(\text{nonrecession}) = 0.96$ on the nonrecession narrative and $\theta(\text{recession}) = 0.01$ on the recession narrative. In contrast, for Cristina Alesci's article emphasizing the recession risks, the model estimates $\theta(\text{recession}) = 0.84$ and $\theta(\text{nonrecession}) = 0.05$.

Based on these LDA outputs, we construct two measures of the narratives conveyed in an article. The first measure is $\theta(d, k)$, the estimated loading of article d on narrative k, where k is either the recession narrative or the nonrecession narrative. The second measure, $\mathbb{1}(d, k)$, is a binary measure to capture articles which are heavily loaded on one particular narrative. We define $\mathbb{1}(d, k) \equiv \mathbb{1}(\theta(d, k) > \frac{1}{D} \sum_{d \in D} \theta(d, k))$, which takes the value 1 if the article loading exceeds the cross-sectional average loading of the narrative and 0 otherwise.

5. Narrative-Driven Fluctuations in Sentiment

5.1. Event-study specification

We now use these measures to test whether different narratives of the yield curve inversion affect consumer sentiment. Our empirical model is a high-frequency event-time regression. For Twitter user i exposed to news article d, we estimate

$$\Delta s_{id} = \alpha + \beta_r \cdot \mathbb{1}(d, \text{recession}) + \beta_{nr} \cdot \mathbb{1}(d, \text{nonrecession}) + \varepsilon_{id}. \tag{16}$$

The dependent variable, Δs_{id}, is the change in tweet sentiment 24 hours before and after the exposure to a narrative, where sentiment is measured with the naïve Bayes classifier described in Section 4.3. The exposure to a narrative is measured using quote retweeting activities on Twitter. We focus on the high-frequency changes in tweet sentiment 24 hours around the exposure to isolate the effect of the yield curve narratives from other macroeconomic events.

The timing is normalized so that the time when a Twitter user is exposed to a narrative is $t = 0$. Therefore, the time dimension of the baseline model in (16) is collapsed. The explanatory variables are the narratives conveyed in the retweeted article, measured using the indicator variables described in Section 4.3. We also consider an alternative specification using the continuous measure of narratives $\theta(d, k)$ (the loading of article d on narrative k). The parameters of interest are β_r and β_{nr}, which estimate the effects of recession and nonrecession narratives on tweet sentiment, respectively.

There are a number of challenges to interpret these coefficients as the effect of narratives on sentiment, which we address in turn. First is the concern for reverse causality. Retweeting decisions may be correlated with sentiment changes, particularly when users choose to retweet a narrative that fits with their pre-existing sentiment. In the following sections, we find no evidence of systematic pre-trends in sentiment among those who quote retweet either narrative. We also study the sentiment of the followers of retweeters, who are exposed to narratives because of the accounts they already follow, not because they chose to engage with the narrative.

Second, we assume that users who quote retweet an article have read the article and processed the narrative it contains. This may be unlikely for Twitter "bots" (i.e., automated accounts that mechanically tweet in response to certain prompts). Robustness checks to rule out likely bots generate slightly stronger results than the baseline in Table 3. This is consistent with bot-driven noise in retweeting creating measurement error that attenuates our results.

Lastly, Twitter users may read an article with a particular narrative but do not become convinced by the article's narrative. In Appendix E, we study this possibility that users stick with their pre-existing narratives rather than adopt a new narrative and find any resulting bias is likely to be small.

5.2. Results

Table 3 contains our main results from estimating variants of (16). Column 1 reports our baseline estimates of β_{r} and β_{nr}, displayed in basis points. Exposure to the recession narrative is associated with a significantly more pessimistic outlook. After a Twitter user is exposed to an article emphasizing the recession narrative, their tweets display a 1.3-basis-

Table 3: Effects of narratives on consumer sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment Changes					
Recession narrative						
$\mathbb{1}(d,k)$	-1.29**		-1.25**			
	(0.65)		(0.62)			
$\theta(d,k)$		-1.74**		-1.65**		
		(0.82)		(0.80)		
Nonrecession narrative						
$\mathbb{1}(d,k)$	-0.11				0.15	
	(0.47)				(0.46)	
$\theta(d,k)$		-0.28				0.03
		(0.64)				(0.63)
R^2	0.012	0.013	0.011	0.012	0.000	0.000
Observations	352	352	352	352	352	352

Notes: This table reports results from estimating variants of the baseline specification in (16). Column (1) reports β_r and β_{nr} in basis points from estimating the baseline specification

$$\Delta s_{id} = \alpha + \beta_r \cdot \mathbb{1}(d, \text{recession}) + \beta_{nr} \cdot \mathbb{1}(d, \text{nonrecession}) + \varepsilon_{id},$$

where Δs_{id} denotes changes in user i's tweet sentiment 24 hours before and after reading article d; and $\mathbb{1}(d,k)$ for $k \in \{\text{recession}, \text{nonrecession}\}$ denotes an indicator variable for whether the loading of article d on narrative k is above the cross-sectional mean. Tweet sentiment is measured with naïve Bayes classifier and an article's loading on a narrative is measured with the LDA model, as described in the main text. Column (2) reports β_r and β_{nr} from estimating $\Delta s_{id} = \alpha + \beta_r \cdot \theta(d, \text{recession}) + \beta_{nr} \cdot \theta(d, \text{nonrecession}) + \varepsilon_{id}$, where $\theta(d,k)$ denotes the loading of article d on narrative k. Columns (3) through (6) report β from estimating univariate models $\Delta s_{id} = \alpha + \beta \cdot x_{dk} + \varepsilon_{id}$, where x_{dk} is $\mathbb{1}(d, \text{recession})$, $\theta(d, \text{recession})$, $\mathbb{1}(d, \text{nonrecession})$, or $\theta(d, \text{nonrecession})$. Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

point more pessimistic sentiment. In contrast, exposure to the nonrecession narrative leads to no significant changes in sentiment. This is consistent with the nonrecession narrative arguing the yield curve inversion is not connected to output, as in the baseline narrative in the model. These results are robust to different measures of narratives, and univariate specifications including just one narrative at a time, as reported in Columns 2 to 6.

To interpret the economic significance, we re-express the estimates in terms of the standard deviations of average daily sentiment changes (5.98 basis points). Our baseline estimates from Table 3 indicate that the exposure to the recession narrative is associate with sentiment declines of 0.2–0.3 standard deviations. To put the numbers into perspective, we compare them with the effects of a major macroeconomic news release during the 3-

week sample period—the release of the August 2019 Jobs Report by the Bureau of Labor Statistics (BLS) on September 6, 2019. The news of 130,000 jobs added is associated with a daily average of 0.3 standard-deviation sentiment increase of Twitter users in our sample. The recession narrative, therefore, has a substantial effect on the users that are exposed, comparable in size to that of a BLS jobs report release.

Our estimates in Table 3 capture the effects of two components of each narrative. The first is its DAG structure. The recession narrative contains a causal link between the slope of the yield curve and future output, while the nonrecession narrative contains no such link. The second is the perceived magnitudes of those links. The model shows that both are necessary to map from a given information set to expectations. The interpretation given above is that households obtain the coefficients of their narrative by estimating the narrative on a long history of data, but they could equally be provided by the news articles themselves. This gives the same outcomes in the model, as long as newspapers are constrained to truthfully report the partial correlations between the variables in their narratives (as in Eliaz and Spiegler, 2020; Eliaz, Spiegler and Weiss, 2021).[24]

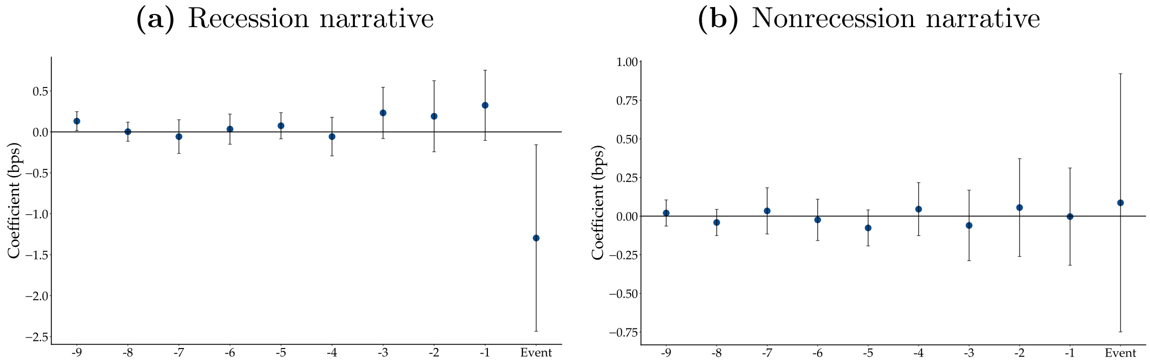
We conduct a number of additional robustness tests. First, a subset of quote retweeters resembles Twitter bots, posting hundreds of tweets every day. In Appendix Table B.4, we exclude users with the highest 5% posting activities. After removing these potential bots, our results remain robust, and the recession narrative displays a slightly stronger effect on the remaining users.

Second, financial and macro conditions can affect sentiments even in the narrow windows that we consider. In Appendix Table B.5, we control for market conditions including the S&P 500 and VIX. Our results remain robust, which suggest that narratives influence sentiment beyond information captured in the financial markets.

Finally, sentiments can differ systematically in certain days of the week. Relatedly, certain news outlets employ different editorial teams for weekdays and weekends. In Appendix Table B.6, we account for the potential seasonality in sentiments by including day-of-the-week controls. Again, our estimates are little changed.

[24]The U.S. Supreme Court applied principles, e.g., in *New York Times Co. v. Sullivan*, that the freedom of press is no broader than that of the general public, and newspapers can be held responsible for defamation, as in the recent case of *Dominion Voting Systems v. Fox News Network*.

Figure 5: Sentiment changes around narrative exposure



Notes: This figure reports regression coefficients and 90% confidence intervals from estimating $\Delta s_{id,t-h} = \alpha + \beta_h \cdot \mathbb{1}(d,k) + \varepsilon_{idh}$, where t denotes the event time when a user i interacts with base tweet containing article d; $\Delta s_{id,t-h} = s_{id,t-h+1} - s_{id,t-h}$ denotes daily sentiment changes h days before the event; and $\mathbb{1}(d,k)$ for $k \in \{\text{recession}, \text{nonrecession}\}$ denotes an indicator variable for whether the loading of article d on narrative k is above the cross-sectional mean. Panel (a) reports the estimates for the recession narrative, and Panel (b) reports the estimates for the nonrecession narrative, measured as described in the main text.

No evidence of pretrends Because newspaper subscription is not exogenously assigned, unobserved differences of Twitter users other than the exposure to narratives could drive both their sentiment changes and retweeting decisions. To ameliorate this concern, we show in Figure 5 that narratives are not associated with sentiment changes in the days before the exposure. For user i interacting with narrative d at event time t, we estimate

$$\Delta s_{id,t-h} = \alpha + \beta_h \cdot \mathbb{1}(d,k) + \varepsilon_{idh}, \tag{17}$$

where $\Delta s_{id,t-h} = s_{id,t-h+1} - s_{id,t-h}$ denotes daily sentiment changes h days before the event; and $\mathbb{1}(d,k)$ for $k \in \{\text{recession}, \text{nonrecession}\}$ denotes the narrative indicator described in Section 4.3.

The lack of pretrends suggests that there is no systematic relationship between sentiment and retweeting decisions until the exposure to a narrative. It also confirms that no "leaked" information has lead to anticipatory movements in sentiment before the inversion. The Federal Reserve's open market operations do not control the exact timing of the yield curve inversion, which makes it plausible that the inversion is unpredictable based on macroeconomic and financial information available prior to the event.

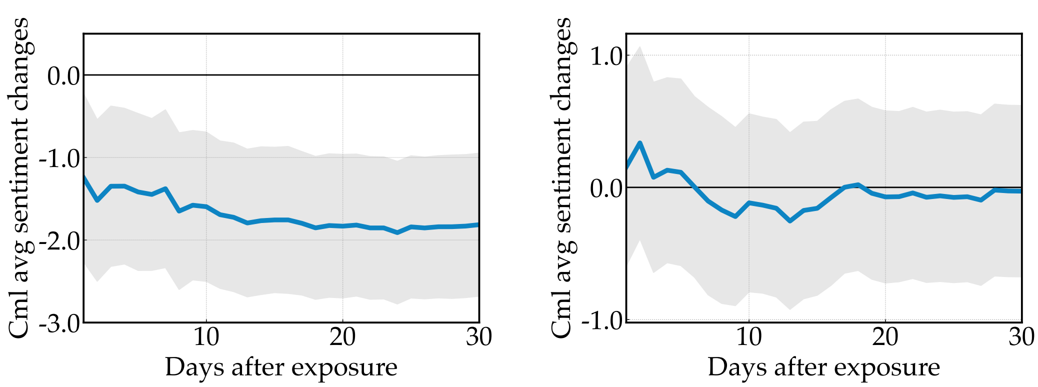
Even though there is no evidence of trends prior to the event, unobserved factors unrelated to narratives can still drive sentiment changes during the narrow event window. For example, someone having a bad day might seek out negative articles to retweet, which raises the concern of reverse causality in our event-time regression. We address this concern by studying the effects of narratives on those who "follow" the quote retweeters. These followers are indirectly exposed when scrolling through the tweets that appear in their timelines and therefore have not chosen to select into the sample. We describe the details in Section 6 and show that the recession narrative also has a significant negative effect on the sentiment of followers.

Persistent effects of narratives The effects of narratives on sentiment is persistent, remaining significant in the month after the exposure. For each narrative $k \in \{\text{recession}, \text{nonrecession}\}$ and horizon h, we estimate in the style of Jordà's 2005 local projections

$$\Delta_h s_{id} = \alpha + \beta_{kh} \cdot \mathbb{1}(d, k) + \varepsilon_{idh}, \tag{18}$$

where $\Delta_h s_{id}$ denotes the average change in consumer i's tweet sentiment between 1 day before and h days after the exposure to a narrative; and $\mathbb{1}(d, k)$ denotes the binary measure of whether the loading of an article d on a narrative k is above the cross-sectional mean. As before, we collapse the time dimension by normalizing the time when a consumer is exposed to a narrative to be $t = 0$.

Figure 6: Dynamic effects of narratives



(a) Recession narrative **(b)** Nonrecession narrative

Notes: Panels (a) and (b) report $\beta_{\text{recession,h}}$ and $\beta_{\text{nonrecession,h}}$ in basis points, respectively, from estimating local projection in (18): $\Delta_h s_{id} = \alpha + \beta_{kh} \cdot \mathbb{1}(d, k) + \varepsilon_{idh}$ for $k \in \{\text{recession}, \text{nonrecession}\}$, where $\Delta_h s_{id}$ denotes the average change in consumer i's tweet sentiment between 1 day before and h days after the exposure to a narrative; and $\mathbb{1}(d, k)$ is the binary measure of narrative. We estimate (18) separately for each horizon $h = 1, \cdots, 30$. Shaded areas represent 90% confidence intervals.

Figure 6 displays the results. Panel (a) shows that the negative effects of the recession narrative are persistent. In the month after reading the interpretation that the yield curve inversion signals an imminent recession, consumers become on average 15 basis points more pessimistic. Panel (b) shows that the exposure to the nonrecession narrative has no such effect.

Economic sentiment or general sentiment Finally, we decompose the content of user tweets to study the source of pessimism—is economic sentiment or general sentiment the driver of pessimism in response to yield-curve narratives? We sort tweets into economic tweets (containing keywords ∗economic∗ or ∗economy∗) and noneconomic, general tweets. Because tweets are short, simple keyword-based methods perform better than natural language models such as topic models (Antenucci, Cafarella, Levenstein, Ré and Shapiro, 2014). Appendix Figure B.2 shows that while exposure to the recession narrative leads to a sharp decline in economic-specific sentiment, pessimism spreads to general sentiment and has a lasting effect. Recessionary narratives, therefore, shape not only users' sentiment of the economic outlook but also their sentiment in other aspects of their everyday lives.

6. Contagion of Narratives

Studies in psychology, marketing, and other fields have suggested that narratives can spread from person to person, potentially going "viral".[25] In this section, we leverage the social network structure of Twitter to trace the contagion of narratives.

6.1. Effects on susceptible users

Shiller (2017) proposes a model in which narratives spread like a virus. The economy consists of three types of agents: susceptibles (those that can be influenced by the narrative), infectives (those that already believe in the narrative), and recovereds (those that have recovered from the narrative). We now focus on measuring the effects of narratives on susceptible users, those that are most likely to react to a new narrative. It should be noted that our estimates are semi-elasticities for a group of users and therefore reflect the intensive margin

[25] See for example Escalas (2007), Machill, Köhler and Waldhauser (2007), and McQuiggan, Rowe, Lee and Lester (2008).

Table 4: Limiting the number of outlets in user timelines

	(1)	(2)	(3)	(4)	(5)	(6)
			Tweet Sentiment Changes			
Recession narrative						
$\mathbb{1}(d,k)$	-1.74*		-1.74*			
	(0.99)		(0.96)			
$\theta(d,k)$		-2.34*		-2.23*		
		(1.26)		(1.23)		
Nonrecession narrative						
$\mathbb{1}(d,k)$	-0.01				0.29	
	(0.69)				(0.67)	
$\theta(d,k)$		-0.34				0.04
		(0.91)				(0.89)
R^2	0.014	0.015	0.014	0.014	0.001	0.000
Observations	227	227	227	227	227	227

Notes: This table reports results from estimating variants of the baseline specification in (16), restricting the sample to users whose Twitter timelines contain no more than 4 different news outlets in the 2-month window around their quote retweets. Column (1) reports β_r and β_{nr} from estimating the baseline specification

$$\Delta s_{id} = \alpha + \beta_r \cdot \mathbb{1}(d, \text{recession}) + \beta_{nr} \cdot \mathbb{1}(d, \text{nonrecession}) + \varepsilon_{id},$$

where Δs_{id} denotes changes in user i's tweet sentiment 24 hours around reading article d; and $\mathbb{1}(d,k)$ for $k \in \{\text{recession}, \text{nonrecession}\}$ denotes an indicator variable for whether the loading of article d on narrative k is above the cross-sectional mean. Tweet sentiment is measured with naïve Bayes classifier and an article's loading on a narrative is measured with the LDA model, as described in the main text. Column (2) reports β_r and β_{nr} from estimating $\Delta s_{id} = \alpha + \beta_r \cdot \theta(d, \text{recession}) + \beta_{nr} \cdot \theta(d, \text{nonrecession}) + \varepsilon_{id}$, where $\theta(d,k)$ denotes the loading of article d on narrative k. Columns (3) through (6) report β from estimating univariate models $\Delta s_{id} = \alpha + \beta \cdot x_{dk} + \varepsilon_{id}$, where x_{dk} is $\mathbb{1}(d, \text{recession})$, $\theta(d, \text{recession})$, $\mathbb{1}(d, \text{nonrecession})$, or $\theta(d, \text{nonrecession})$. Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

of how "lethal" a narrative is, rather than the extensive margin (the probability of being infected).

To focus on susceptible users, we limit our sample to users who retweet articles from a small number of news outlets only. This restriction is based on two assumptions. First, a Twitter user who is "infected" by a particular narrative will tend to retweet a large number of news outlets to promote their story. We rule out such users by restricting the maximum number of different news outlets to be 4, the mean number of outlets in the sample. Second, there are no "recovered" users, because the yield curve inversion is rare and brief, lasting only two days.

Table 4 shows that, as in our main exercise, the recession narrative leads to a decline in sentiment and the nonrecession narrative has no effect. However, the impact of recession narrative is approximately 50% stronger on susceptible users than on the full sample. We can alternatively interpret the results in Table 4 as a robustness check, ensuring that the effects are not driven by users who selectively retweet many articles with a particular narrative to promote that agenda, rather than processing the information contained in a narrative.

6.2. Contagious effects on followers

We can further trace the effects of a narrative as it spreads through the social network. In Section 5.2, we estimated the direct effects of a narrative on someone that engages with it. In this section, we estimate the spillover effect of a narrative on someone that is indirectly exposed through the social network.

When a user quote retweets an article, the article along with the added commentary are posted on this user's timeline. This post would appear on the Twitter home screen of anyone that follows this user. When the followers browse Twitter when the quote retweet is posted, they are therefore exposed to the narrative via the tweets of people they follow. Do the effects of the recession narrative survive as it spreads through the social network?

We collect the list of followers of a quote retweeter, sample 200 followers for a given quote retweeter, and compare the changes in their tweet sentiment around their friends' posting of a narrative. For there to be an indirect exposure, followers need to be active on Twitter during the days around the quote retweet. We therefore require them to have posted at least one tweet the day before and the day after their friends' quote retweet. This also allows us to estimate the sentiment of those followers, by analyzing the content of those tweets.

For follower j of quote retweeter i of article d published by news outlet n, we estimate the effects of an indirect exposure to a narrative

$$\Delta s_{j(i)d} = \alpha_n + \beta_r \cdot \mathbb{1}(d, \text{recession}) + \beta_{nr} \cdot \mathbb{1}(d, \text{nonrecession}) + \varepsilon_{jid}, \tag{19}$$

where $\Delta s_{j(i)d}$ denotes changes in follower j's tweet sentiment 24 hours before and after j's friend i retweets article d; α_n is an outlet fixed effect; and $\mathbb{1}(d, k)$ for $k \in \{\text{recession}, \text{nonrecession}\}$

Table 5: Spillover effects of narratives on followers of quote retweeters

	Tweet Sentiment Changes of Followers			
	(1)	(2)	(3)	(4)
	Equal-Weighted		Probability-Weighted	
Recession narrative				
$\mathbb{1}(d,k)$	-0.729***		-1.629***	
	(0.092)		(0.146)	
$\theta(d,k)$		-0.933**		-3.944
		(0.405)		(2.372)
Nonrecession narrative				
$\mathbb{1}(d,k)$	-0.314		0.174	
	(0.287)		(0.754)	
$\theta(d,k)$		-0.149		0.972
		(0.397)		(0.853)
Observations	2107	2107	2107	2107
R^2	0.002	0.001	0.001	0.003
FE	outlet	outlet	outlet	outlet
Double-clustered SE	yes	yes	yes	yes

Notes: This table reports estimates of β_r and β_{nr} (in basis points) from estimating variants of the regression in (19). In Columns (1) and (2), observations are equal weighted, and standard errors, reported in parenthesis, are double-clustered by date and quote retweet. In Columns (3) and (4), observations are weighted by $w_{j(i)} = N_i/n_i$, where N_i represent the total number of followers i has and $n_i = \max\{200, N_i\}$ represent the number of i's followers that are randomly sampled. Standard errors are double clustered by date and quote retweet and estimated using a robust sandwich estimator. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

denotes an indicator variable for whether the loading of article d on narrative k is above the cross-sectional mean. Standard errors are double clustered by date and quote retweet. As before, tweet sentiment is measured with naïve Bayes classifier and an article's loading on a narrative is measured with the LDA model.

Twitter "influencers" with many followers can have disproportional sway on their followers. To account for this possibility, we adopt an alternative probability weighting scheme that assigns more weight to quote retweeters with many followers. Each follower j of quote retweeter i is weighted by $w_{j(i)} = N_i/n_i$, where N_i represent the total number of followers i has and $n_i = \max\{200, N_i\}$ represent the number of i's followers that are randomly sampled. In the spirit of survey designs (see Skinner and Mason, 2012, for a comprehensive discussion), this weighting scheme takes into account that observations have different probability of being sampled: 1 follower of a user with 10 followers represent 1 person, while 1 follower of a user with 10,000 followers represent 50 people.

The first two columns in Table 5 report our baseline results. The main result in Column 1 shows that the recession narrative is contagious. An indirect exposure to the recession narrative reduces the followers' tweet sentiment by 0.8 basis points. The magnitude is only around 40% smaller than those who are directly exposed (Table 3). The nonrecession narrative, on the other hand, shows no sign of being viral. Results are little changed in Column 2 if we measure narrative using the continuous LDA loading rather than the binary measure.

The last two columns report the results under probability weighting. Allowing Twitter influencers to have differential sway on their followers, we find that the recession narrative makes exposed followers more pessimistic compared to the baseline case in which we limit the weight of influencers. The results suggest central nodes on Twitter with many followers play an important role in the contagion of viral narratives.

7. Conclusion

Narratives are increasingly seen as an important factor in how economic agents form their expectations, by both academics (Shiller, 2017, 2020) and policymakers (Schnabel, 2020). We provide evidence that exposure to particular narratives in the media does indeed have significant effects on consumer sentiment.

Formalizing narratives as directed acyclic graphs used to form expectations of the future, we show that the distribution of narratives across households can affect expectations and aggregate outcomes in an otherwise standard macroeconomic model. In addition, certain groups of narratives are observationally equivalent for expectations. In the context of the inversion of the U.S. yield curve in 2019, the distinguishing feature between a "recession" narrative and a "nonrecession" narrative is, therefore, whether there is a causal link made connecting the inverted yield curve with an upcoming recession. The direction of that link does not matter for expectations or choices.

Standard tools from topic modeling in natural language processing are well suited to making this distinction. We do this in a large corpus of articles from traditional news media, which is a key source of macroeconomic narratives (Andre et al., 2022b). Linking these articles with rich data on Twitter activity, we find that engaging with an article advancing a “recession” narrative causes a significant and persistent decline in the sentiment of that Twitter user, as embodied in their other activity on the social media site at the time. In contrast, engaging with a “nonrecession” narrative has no such effect on sentiment. This is precisely what would be predicted by models in which viral narratives affect aggregate behaviour by shifting expectations. It also suggests a powerful role for the media in influencing aggregate sentiment (highlighted, for example, in Nimark, 2014). Furthermore, we find that the sentiment effects of recession narratives are contagious, as hypothesized by Shiller (2017) and others: narrative-driven changes in sentiment transmit from those who engage with the particular news article to their Twitter followers.

Our approach using tools from natural language processing to extract relevant groups of narratives from text can be used in other settings. For example, while news media is an important source of narratives, similar techniques can be used to study economic narratives created by policymakers in monetary and fiscal policy statements. These data sources are naturally occurring, which means that our method can be deployed to track the evolution of narratives and their ongoing effects—potentially providing a useful input to discussions of macroeconomic policy.

References

Acharya, Sushant, Jess Benhabib, and Zhen Huo, "The Anatomy of Sentiment-Driven Fluctuations," *Journal of Economic Theory*, 2021, *195*, 105280.

Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart, "Subjective Models of the Macroeconomy: Evidence from Experts and a Representative Sample," *Review of Economic Studies*, 2022.

_ , Ingar Haaland, Christopher Roth, and Johannes Wohlfart, "Narratives about the Macroeconomy," Technical Report, University of Bonn and University of Cologne, Germany 2022.

Angeletos, George-Marios and Chen Lian, "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, Elsevier, January 2016, pp. 1065–1240.

_ and Jennifer La'O, "Sentiments," *Econometrica*, 2013, *81* (2), 739–779.

_ , Fabrice Collard, and Harris Dellas, "Quantifying Confidence," *Econometrica*, 2018, *86* (5), 1689–1726.

Antenucci, Dolan, Michael Cafarella, Margaret Levenstein, Christopher Ré, and Matthew D Shapiro, "Using social media to measure labor market flows," Technical Report, National Bureau of Economic Research 2014.

Ash, Elliott, Germain Gauthier, and Philine Widmer, "Text Semantics Capture Political and Economic Narratives," *arXiv preprint arXiv:2108.01720*, 2021.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, September 2021, *89* (5), 2375–2408.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel, "The Economic Effects of Social Networks: Evidence from the Housing Market," *Journal of Political Economy*, 2018, *126* (6), 2224–2276.

Bauer, Michael D. and Mikhail Chernov, “Interest Rate Skewness and Biased Beliefs,” Technical Report, National Bureau of Economic Research 2021.

Bauer, Michael D and Thomas M Mertens, “Information in the yield curve about future recessions,” *FRBSF Economic Letter*, 2018, *20*, 1–5.

Bauer, Michael D., Carolin Pflueger, and Adi Sunderam, “Perceptions about Monetary Policy,” Technical Report, National Bureau of Economic Research 2022.

Bianchi, Francesco, Howard Kung, and Roberto Gomez Cram, “Using Social Media to Identify the Effects of Congressional Partisanship on Asset Prices,” Technical Report, National Bureau of Economic Research 2021.

Bilbiie, Florin O., “The New Keynesian Cross,” *Journal of Monetary Economics*, 2019.

Blei, David M, Andrew Y Ng, and Michael I Jordan, “Latent Dirichlet Allocation,” *Journal of Machine Learning Research*, 2003, *3*, 993–1022.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, “Diagnostic Expectations and Credit Cycles,” *Journal of Finance*, 2018, *73* (1), 199–227.

Buehlmaier, Matthias MM and Toni M Whited, “Are Financial Constraints Priced? Evidence from Textual Analysis,” *Review of Financial Studies*, 2018, *31* (7), 2693–2728.

Bui, Ha, Zhen Huo, Andrei A. Levchenko, and Nitya Pandalai-Nayar, “Information Frictions and News Media in Global Value Chains,” *NBER Working Paper Series*, May 2022, *30033.*

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, “Understanding Booms and Busts in Housing Markets,” *Journal of Political Economy*, 2016, *124* (4), 1088–1147.

Bybee, Leland, Bryan T Kelly, Asaf Manela, and Dacheng Xiu, “The Structure of Economic News,” Technical Report, National Bureau of Economic Research 2020.

Calomiris, Charles W and Harry Mamaysky, “How News and its Context Drive Risk and Returns around the World,” *Journal of Financial Economics*, 2019, *133* (2), 299–336.

Caplin, Andrew and John V Leahy, "Wishful Thinking," Technical Report, National Bureau of Economic Research 2019.

Carroll, Christopher and Tao Wang, "Epidemiological Expectations," in Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, January 2023, pp. 779–806.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia, "Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 1–35.

Chahrour, Ryan, Kristoffer Nimark, and Stefan Pitschner, "Sectoral Media Focus and Aggregate Fluctuations," *American Economic Review*, 2021, *111* (12), 3872–3922.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar, "The Formation of Expectations, Inflation, and the Phillips Curve," *Journal of Economic Literature*, 2018, *56* (4), 1447–91.

Ehrmann, Michael and Alena Wabitsch, "Central Bank Communication with Non-Experts–A road to Nowhere?," *Journal of Monetary Economics*, 2022, *127*, 69–85.

Eliaz, Kfir and Ran Spiegler, "A Model of Competing Narratives," *American Economic Review*, 2020, *110* (12), 3786–3816.

_ , _ , and Yair Weiss, "Cheating with Models," *American Economic Review: Insights*, December 2021, *3* (4), 417–434.

_ , Simone Galperti, and Ran Spiegler, "False Narratives and Political Mobilization," *arXiv*, 2022, (2206.12621).

Escalas, Jennifer Edson, "Self-referencing and persuasion: Narrative transportation versus analytical elaboration," *Journal of Consumer Research*, 2007, *33* (4), 421–429.

Flynn, Joel P and Karthik Sastry, "The Macroeconomics of Narratives," *Manuscript*, 2022.

Friedman, Jerome, Trevor Hastie, Robert Tibshirani et al., *The Elements of Statistical Learning*, Vol. 1, Springer series in statistics New York, 2001.

Galí, Jordi, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton, USA and Oxford, UK: Princeton University Press, 2008.

Gentzkow, Matthew, Jesse M Shapiro, and Michael Sinkinson, "Competition and Ideological Diversity: Historical Evidence from US Newspapers," *American Economic Review*, 2014, *104* (10), 3073–3114.

Go, Alec, Richa Bhayani, and Lei Huang, "Twitter Sentiment Classification Using Distant Supervision," *CS224N project report, Stanford*, 2009, *1* (12), 2009.

Goetzmann, William N, Dasol Kim, and Robert J Shiller, "Crash Narratives," Technical Report, National Bureau of Economic Research 2022.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera, "Social Media, Sentiment and Public Opinions: Evidence from #Brexit and#USElection," *European Economic Review*, 2021, p. 103772.

Haldane, Andrew, Alistair Macaulay, and Michael McMahon, "The Three E's of Central-Bank Communication with the Public," 2021.

Hansen, Stephen, Michael McMahon, and Andrea Prat, "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach," *Quarterly Journal of Economics*, 2018, *133* (2), 801–870.

Harvey, Campbell R, "The Real Term Structure and Consumption Growth," *Journal of Financial Economics*, 1988, *22* (2), 305–333.

Havránek, Tomáš, "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting," *Journal of the European Economic Association*, 2015, *13* (6), 1180–1204.

Hoffman, Matthew, Francis Bach, and David Blei, "Online Learning for Latent Dirichlet Allocation," *Advances in Neural Information Processing Systems*, 2010, *23*, 856–864.

Hünermund, Paul and Elias Bareinboim, "Causal Inference and Data Fusion in Econometrics," *The Econometrics Journal*, March 2023.

Jordà, Òscar, "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 2005, *95* (1), 161–182.

Jurkowitz, Mark, Amy Mitchell, Elisa Shearer, and Mason Walker, "US Media Polarization and the 2020 Election: A Nation Divided," *Pew Research Center*, 2020, *24.*

Kamdar, Rupal, "The Inattentive Consumer: Sentiment and Expectations," *Manuscript*, 2019.

Kendall, Chad W. and Constantin Charles, "Causal Narratives," *NBER Working Paper Series*, August 2022, *30346.*

Koller, Daphne and Nir Friedman, *Probabilistic Graphical Models: Principles and Techniques* Adaptive Computation and Machine Learning, Cambridge, Mass. ; London: MIT Press, 2009.

Lagerborg, Andresa, Evi Pappa, and Morten O Ravn, "Sentimental Business Cycles," *The Review of Economic Studies*, 2022.

Larsen, Vegard and Leif Anders Thorsrud, "Business Cycle Narratives," 2019.

Leombroni, Matteo, Andrea Vedolin, Gyuri Venter, and Paul Whelan, "Central Bank Communication and the Yield Curve," *Journal of Financial Economics*, September 2021, *141* (3), 860–880.

Levchenko, Andrei A and Nitya Pandalai-Nayar, "TFP, News, and "Sentiments": The International Transmission of Business Cycles," *Journal of the European Economic Association*, 2020, *18* (1), 302–341.

Levy, Ro'ee, "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment," *American Economic Review*, 2021, *111* (3), 831–70.

Macaulay, Alistair and Wenting Song, "News Media, Inflation, and Sentiment," *AEA Papers and Proceedings*, May 2023, *113*, 172–176.

Machill, Marcel, Sebastian Köhler, and Markus Waldhauser, "The use of narrative structures in television news: An experiment in innovative forms of journalistic presentation," *European Journal of Communication*, 2007, *22* (2), 185–205.

Malmendier, Ulrike and Stefan Nagel, "Learning from Inflation Experiences," *Quarterly Journal of Economics*, 2016, *131* (1), 53–87.

Mankiw, N Gregory and Ricardo Reis, "Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve," *Quarterly Journal of Economics*, 2002, *117* (4), 1295–1328.

Matveev, Dmitry and Francisco Javier Ruge-Murcia, "Tariffs and the Exchange Rate: Evidence from Twitter," Technical Report, Bank of Canada Staff Working Paper 2021.

McQuiggan, Scott W, Jonathan P Rowe, Sunyoung Lee, and James C Lester, "Story-based learning: The impact of narrative on learning experiences and outcomes," in "Intelligent Tutoring Systems: 9th International Conference, ITS 2008, Montreal, Canada, June 23-27, 2008 Proceedings 9" Springer 2008, pp. 530–539.

Molavi, Pooya, "Macroeconomics with Learning and Misspecification : A General Theory and Applications," 2019.

_ , Alireza Tahbaz-Salehi, and Andrea Vedolin, "Model Complexity, Expectations, and Asset Prices," *Manuscript*, 2021.

Nimark, Kristoffer P, "Man-Bites-Dog Business Cycles," *American Economic Review*, 2014, *104* (8), 2320–67.

Nyman, Rickard, Sujit Kapadia, and David Tuckett, "News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment," *Journal of Economic Dynamics and Control*, 2021, *127*, 104119.

Pearl, Judea, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, 2009.

Ramos, Juan et al., "Using tf-idf to Determine Word Relevance in Document Queries," in "Proceedings of the first instructional conference on machine learning," Vol. 242 Citeseer 2003, pp. 29–48.

Rish, Irina et al., "An Empirical Study of the Naive Bayes Classifier," in "IJCAI 2001 workshop on empirical methods in artificial intelligence," Vol. 3 2001, pp. 41–46.

Schnabel, Isabel, "Narratives about the ECB's Monetary Policy–Reality or Fiction?," *Juristische Studiengesellschaft*, 2020.

Shiller, Robert J, "Narrative Economics," *American Economic Review*, 2017, *107* (42), 967–1004.

_ , *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*, Princeton University Press, 2020.

Sims, Christopher A, "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

Skinner, Chris and Ben Mason, "Weighting in the regression analysis of survey data with a cross-national application," *Canadian Journal of Statistics*, 2012, *40* (4), 697–711.

Spiegler, Ran, "Bayesian Networks and Boundedly Rational Expectations," *Quarterly Journal of Economics*, 2016, *131* (3), 1243–1290.

_ , "Behavioral Implications of Causal Misperceptions," *Annual Review of Economics*, 2020, *12*, 81–106.

Verma, Thomas and Judea Pearl, "Causal Networks: Semantics and Expressiveness," in "Machine Intelligence and Pattern Recognition," Vol. 9, Elsevier, 1990, pp. 69–76.

Appendices

A. Proofs

Proposition 1 This proof proceeds in several stages. First, we derive expressions for expectations under given (fixed) likelihoods $p(.|.)$. Second, we write down the equilibrium conditions of the model with those fixed narratives. Third, we solve for the likelihoods that make the narratives consistent with the equilibrium outcomes, and thus find two equilibria. Fourth, we show that under the parameter restrictions in the proposition, only one of these equilibria is stable. Fifth, we derive properties of expectations and output in this stable equilibrium.

Step 1: expressions for expectations. Since the model is log-linearized, the true data generating process for the vector $x_t = (y_t, r_t, z_t)'$ is a VAR(1). All shocks in this process have i.i.d. Normal distributions, so assuming that the initial state x_0 also has a multivariate Normal distribution, x_t is multivariate Normal in every t. All conditional distributions therefore imply conditional expectations which are linear in the conditioning variables.

In other words, the DAGs in Definition 2 can be written as if they reflect linear perceived laws of motion for each variable

$$y_t = h_y^k y_{t-1} + h_z^k z_{t-1} + e_t^y \tag{20}$$

$$r_t = e_t^r \tag{21}$$

$$z_t = f^k y_t + e_t^z \tag{22}$$

for $k \in \{b, z\}$, where e_t^y, e_t^r, e_t^z are all mean-zero shocks and $h_z^b, f^b = 0$.

Rolling forward one period and taking expectations, we obtain

$$\mathbb{E}_t^k\, y_{t+1} = h_y^k y_t + h_z^k z_t \tag{23}$$

$$\mathbb{E}_t^k\, r_{t+1} = 0 \tag{24}$$

$$\mathbb{E}_t^k\, z_{t+1} = f^k\, \mathbb{E}_t^k\, y_{t+1} = f^k h_y^k y_t + f^k h_z^k z_t \tag{25}$$

Next, we solve for consumption under each narrative. In this, it is useful to note that expectations at any horizon can be written as

***Lemma* 2** (rewriting expectations)**.** *With the perceived laws of motion defined in equations* (20)-(22)*, expectations are given by*

$$\mathbb{E}_t^k\, y_{t+s} = (h_y^k + h_z^k f^k)^{s-1}\, \mathbb{E}_t^k\, y_{t+1} \tag{26}$$

$$\mathbb{E}_t^k\, r_{t+s} = 0 \tag{27}$$

$$\mathbb{E}_t^k\, z_{t+s} = f^k (h_y^k + h_z^k f^k)^{s-1}\, \mathbb{E}_t^k\, y_{t+1} \tag{28}$$

for all $s \geq 1$.

Proof. From equations (20)-(22), we have

$$\begin{aligned}
\mathbb{E}_t^k\, y_{t+s} &= h_y^k\, \mathbb{E}_t^k\, y_{t+s-1} + h_z^k\, \mathbb{E}_t^k\, z_{t+s-1} + \mathbb{E}_t^k\, e_{t+s}^y \\
&= (h_y^k + h_z^k f^k)\, \mathbb{E}_t^k\, y_{t+s-1} \\
&= (h_y^k + h_z^k f^k)^{s-1}\, \mathbb{E}_t^k\, y_{t+1} && (29) \\
\mathbb{E}_t^k\, r_{t+s} &= \mathbb{E}_t^k\, e_{t+s}^r = 0 && (30) \\
\mathbb{E}_t^k\, z_{t+s} &= f^k\, \mathbb{E}_t^k\, y_{t+s} + \mathbb{E}_t^k\, e_{t+s}^z \\
&= f^k (h_y^k + h_z^k f^k)^{s-1}\, \mathbb{E}_t^k\, y_{t+1} && (31)
\end{aligned}$$

□

Substituting equations (26)-(28) into the consumption function (1), evaluating the infinite sums and then using equation (23) to substitute out for $\mathbb{E}_t^k\, y_{t+1}$, we obtain

$$c_t^k = (1 - \beta + h_y^k \psi^k) y_t - \beta \sigma r_t + h_z^k \psi^k z_t \tag{32}$$

where ψ^k is the elasticity of consumption to $\mathbb{E}_t^k\, y_{t+1}$ under narrative k

$$\psi^k = \frac{\beta(1-\beta)}{1 - \beta h_y^k + \beta h_z^k f^k} \tag{33}$$

Since $h_z^b, f^b = 0$ by assumption, we simplify notation by writing $h_z^z = h_z$ and $f^z = f$ from here.

Step 2: equilibrium conditions. Given these expressions for consumption, the equilibrium conditions of the model can be expressed as:

$$
\begin{aligned}
y_t &= (1-\lambda)c_t^b + \lambda c_t^z && (34)\\
&= \left(1-\beta+(1-\lambda)h_y^b\psi^b + \lambda h_y^z\psi^z\right) y_t - \beta\sigma r_t + \lambda h_z \psi^z z_t && (35)\\
r_t &= \kappa\phi y_t + \kappa\mu\phi z_{t-1} + \phi v_t^\pi + v_t^r && (36)\\
z_t &= \chi y_t + v_t^z && (37)
\end{aligned}
$$

Taking the narrative coefficients $f, h_y^b, h_y^z, h_z, \psi^b, \psi^z$ as given, we solve this system as standard to obtain

$$
\begin{aligned}
y_t &= -\frac{1}{\Lambda}\left(\beta\kappa\mu\phi\sigma z_{t-1} + \beta\phi\sigma v_t^\pi + \beta\sigma v_t^r - \lambda h_z\psi^z v_t^z\right) && (38)\\
r_t &= \frac{1}{\Lambda}\left(\kappa\mu\phi(\Lambda-\beta\kappa\phi\sigma)z_{t-1} + \phi(\Lambda-\beta\kappa\phi\sigma)v_t^\pi + (\Lambda-\beta\kappa\phi\sigma)v_t^r + \kappa\phi\lambda h_z\psi^z v_t^z\right) && (39)\\
z_t &= -\frac{1}{\Lambda}\left(\beta\kappa\mu\phi\sigma\chi z_{t-1} + \beta\phi\sigma\chi v_t^\pi + \beta\sigma\chi v_t^r - (\Lambda+\lambda h_z\psi^z\chi)v_t^z\right) && (40)
\end{aligned}
$$

where

$$
\Lambda = 1 + \beta\kappa\phi\sigma - (1-\beta+(1-\lambda)h_y^b\psi^b + \lambda h_y^z\psi^z) - \lambda h_z\psi^z\chi \tag{41}
$$

Step 3: consistency between narratives and outcomes. Matching coefficients between equations (20) and (38) for those with the z narrative, we obtain

$$
h_y^z = 0, \quad h_z = -\frac{\beta\kappa\mu\phi\sigma}{\Lambda} \tag{42}
$$

We now turn to f. We cannot simply match coefficients between equations (22) and (40) because $Cov(y_t, v_t^z) \neq 0$, so χ does not capture the full relationship between z_t and y_t. From Molavi (2019), the Constrained-Rational Expectations Equilibrium is obtained as the limit of least-squares learning. Without loss of generality we therefore assume households estimate equation (40) in a large sample using OLS, giving

$$
f = \frac{Cov(z_t, y_t)}{Var(y_t)} = \chi + \frac{\lambda h_z\psi^z}{\Lambda}\frac{\sigma_z^2}{Var(y_t)} \tag{43}
$$

Similarly, estimating equation (20) under the baseline narrative (with $h_z^b = 0$) gives

$$h_y^b = \frac{Cov(y_t, y_{t-1})}{Var(y_{t-1})} = -\frac{\beta\kappa\mu\phi\sigma}{\Lambda}\left(\chi + \frac{\lambda h_z \psi^z}{\Lambda}\frac{\sigma_z^2}{Var(y_{t-1})}\right) \tag{44}$$

Restricting attention to stable equilibria, we have that $Var(y_{t-1}) = Var(y_t)$, which means we can write

$$h_y^b = h_z f \tag{45}$$

In turn, combining equations (33), (44), and (45) we obtain

$$\psi^b = \psi^z = \frac{\beta(1-\beta)}{1-\beta h_z f} \tag{46}$$

In addition, equation (41) simplifies to

$$\Lambda = \beta(1+\kappa\phi\sigma) - \psi^z h_z((1-\lambda)f + \lambda\chi) \tag{47}$$

For a given variance of y_t, equations (42), (43), (46), and (47) form a system of equations in 4 unknowns: f, h_z, Λ, ψ^z.

At this point we turn to the special case, and take $\sigma_z^2 \to 0$. Equation (43) then reduces to $f = \chi$, and combining this with equations (46), and (47) yields

$$\Lambda = \beta(1+\kappa\phi\sigma) - \frac{\beta(1-\beta)\chi h_z}{1-\beta\chi h_z} \tag{48}$$

Solving equations (42) and (48) for h_z, Λ yields two solutions:

$$h_z = \frac{1+\kappa\phi\sigma(1-\beta\mu\chi) \pm \sqrt{\Omega}}{2\chi(1+\beta\kappa\phi\sigma)} \tag{49}$$

$$\Lambda = -\frac{\beta}{2}\left(-1-\kappa\phi\sigma(1-\beta\mu\chi) \pm \sqrt{\Omega}\right) \tag{50}$$

where

$$\Omega = 1 + \kappa\phi\sigma(2+2(2-\beta)\mu\chi) + (\kappa\phi\sigma)^2(1+\beta\mu\chi)^2 \tag{51}$$

Step 4: stability of equilibria. Only one of these solutions implies a stable equilibrium. To show this, we first substitute equation (42) into equation (40) to obtain:

$$z_t = h_z \chi z_{t-1} - \frac{1}{\Lambda}\left(\beta\phi\sigma\chi v_t^{\pi} + \beta\sigma\chi v_t^{r} - (\Lambda + \lambda h_z \psi^z \chi) v_t^{z}\right) \tag{52}$$

z_t is therefore only stable in equilibrium if $|h_z\chi| < 1$. We now show that given the parameter restrictions in Proposition 1, this is only true of one of the two solutions in equations (49) and (50).

First, consider the solution where $\sqrt{\Omega}$ enters positively. In this case $h_z\chi > 1$ whenever:

$$\sqrt{\Omega} > 2(1 + \beta\kappa\phi\sigma) - 1 - \kappa\phi\sigma(1 - \beta\mu\chi) \tag{53}$$

Squaring both sides, substituting out for Ω using equation (51), and rearranging gives:

$$(\kappa\phi\sigma)^2\left((1 + \beta\mu\chi)^2 - (\beta(2 + \mu\chi) - 1)^2\right) > -4\kappa\phi\sigma(1 - \beta)(1 + \mu\chi) \tag{54}$$

Expanding brackets and cancelling terms, this becomes:

$$\beta(1 + \mu\chi)(1 + \kappa\phi\sigma) > 0 \tag{55}$$

This is true if $\mu\chi > -1$. With the parameter restrictions in Proposition 1, this equilibrium is therefore explosive.

We now proceed to show that the other equilibrium is stable under the same parameter restrictions. First, we show $h_z\chi < 1$. This is true if

$$\sqrt{\Omega} > 1 + \kappa\phi\sigma(1 - \beta\mu\chi) - 2(1 + \beta\kappa\phi\sigma) \tag{56}$$

Simplifying the right hand side yields:

$$\sqrt{\Omega} > -(1 + \kappa\phi\sigma(\beta(\mu\chi + 2)) - 1) \tag{57}$$

A sufficient condition for this to be true is that the right hand side is strictly negative.

The restriction $\mu\chi > -1$ gives an upper bound on that parameter combination

$$-(1+\kappa\phi\sigma(\beta(\mu\chi+2))-1) < -(1+\kappa\phi\sigma(\beta-1)) \tag{58}$$

In turn, that upper bound is strictly negative as long as $(1-\beta)\kappa\phi\sigma < 1$, as specified in Proposition 1. With these parameter restrictions, we therefore have $h_z\chi < 1$. The final step is then to show that, in addition, $h_z\chi > -1$.

Following the same steps above, equation (49) at the solution containing $-\sqrt{\Omega}$ implies $h_z\chi > -1$ if

$$\sqrt{\Omega} < 1+\kappa\phi\sigma(1-\beta\mu\chi)+2(1+\beta\kappa\phi\sigma) \tag{59}$$

Squaring both sides and simplifying we obtain

$$-\kappa\phi\sigma(2\beta+(1+\beta)(1-\mu\chi))-(\kappa\phi\sigma)^2\beta(1+\beta)(1-\mu\chi) < 8 \tag{60}$$

With the restriction in Proposition 1 that $\mu\chi < 1$, the left hand side of this inequality is always strictly negative, so the inequality holds. This equilibrium is therefore stable.

Step 5: properties of the unique stable equilibrium. Combining equations (20) with $k=b$ and (45), we obtain

$$\mathbb{E}_t^b\, y_{t+1} = h_z\chi y_t \tag{61}$$

Similarly, combining (20) with $k=z$ and (37):

$$\mathbb{E}_t^z\, y_{t+1} = h_z\chi y_t + h_z v_t^z \tag{62}$$

These combine to give equation (11) in Proposition 1, with $\mathcal{G} = h_z$. The properties in equation (12) are a direct consequence of equations (42) and (47).

Finally, we differentiate equation (38) with respect to λ, holding the predetermined z_{t-1} fixed, to obtain equation (13) with

$$\mathcal{H} = \frac{h_z\psi^z}{\Lambda} \tag{63}$$

Substituting out for ψ^z and Λ using equations (46) and (48), this becomes

$$\mathcal{H} = \frac{h_z(1-\beta)}{1+\kappa\phi\sigma - \chi h_z(1+\beta\kappa\phi\sigma)} \tag{64}$$

This shares the same sign as h_z (and so $\mathcal{G}$) if and only if the denominator of this fraction is positive. After some algebraic manipulation, we find that this condition holds if

$$h_z\chi < \frac{1+\kappa\phi\sigma}{1+\beta\kappa\phi\sigma} \tag{65}$$

Since $\beta \in [0,1]$, the right hand side of this is weakly greater than 1. This therefore reduces to the same condition as that ensuring the stability of the equilibrium, which we have shown to be satisfied above. This therefore confirms equation (14).

Proposition 2. In any DAG k, the conditional independence assumptions can be summarized by the Bayesian network factorization formula, defined as

$$\tilde{p}_k(x_{\mathcal{N}}) = \prod_{n\in\mathcal{N}} p(x_n|x_{\mathcal{L}^k(n)}) \tag{66}$$

where $x_{\mathcal{N}}$ denotes the set of all variables in the narrative; and $x_{\mathcal{L}^k(n)}$ denotes the subset of those variables which have a direct causal link to x_n in that narrative.

Two DAGs with the same $\tilde{p}_k(.)$ always imply the same conditional expectations, for any information set (Spiegler, 2016, 2020). We proceed to show that the three z narratives have such identical factorizations.

First, we show $\tilde{p}_c(\cdot) = \tilde{p}_b(\cdot)$. By the definitions of joint and conditional probabilities:

$$\begin{aligned}
\tilde{p}_c(r_s, r_{s+1}, y_s, y_{s+1}, z_s) &= p(r_s|y_s)p(r_{s+1}|y_{s+1})\frac{p(y_s, z_s)}{p(z_s)}p(y_{s+1}|r_s, y_s, z_s)p(z_s) \\
&= p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)p(z_s|y_s)p(y_{s+1}|r_s, y_s, z_s) \\
&= \tilde{p}_b(r_s, r_{s+1}, y_s, y_{s+1}, z_s)
\end{aligned}$$

Similarly, we can show $p_b(\cdot) = p_a(\cdot)$:

$$
\begin{aligned}
\tilde{p}_b(r_s, r_{s+1}, y_s, y_{s+1}, z_s) &= p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)\frac{p(y_{s+1}, z_s|r_s, y_s)}{p(z_s|r_s, y_s)}p(z_s|y_s) \\
&= p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)p(y_{s+1}|r_s, y_s)p(z_s|y_s, y_{s+1}) \\
&= \tilde{p}_a(r_s, r_{s+1}, y_s, y_{s+1}, z_s)
\end{aligned}
$$

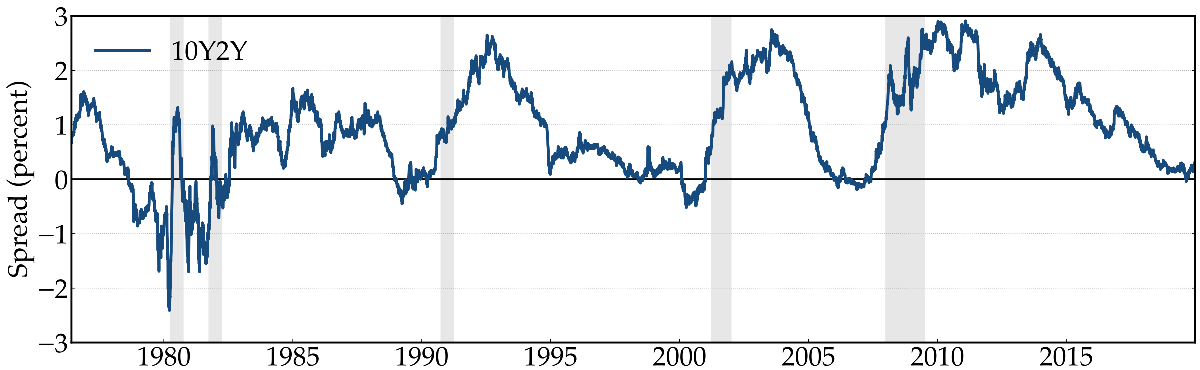
where the penultimate equality uses that $p(z_s|y_s, r_s) = p(z_s|y_s)$, as r_s is not directly causally related to z_s.

Lemma 3. Equivalent to equation (23) in the proof of Proposition 1.

Lemma 1. Follows directly from stacking equations (26)-(28) into vector form.

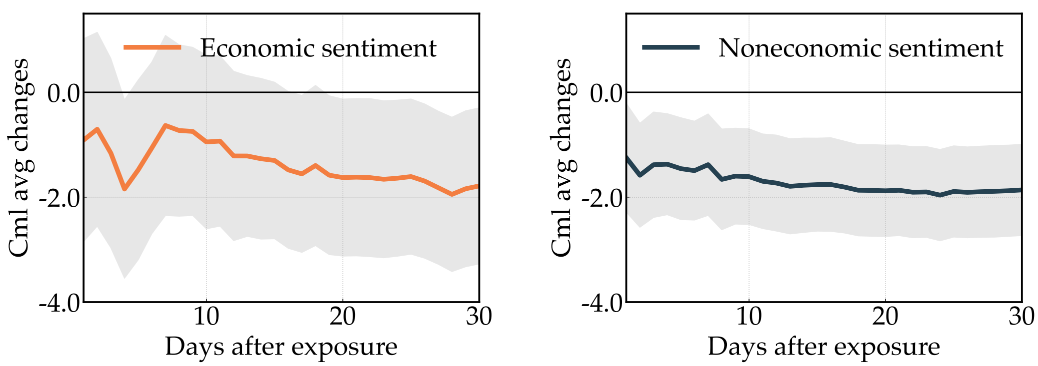
B. Additional Tables and Figures

Figure B.1: Yield curve inversion and recessions in the US



Notes: Yield curve and recessions in the US for 1976–2019. The blue solid line displays the spread between 10-year treasury yield and 2-year treasury yield ("10Y2Y"). Recession dates as classified by NBER are shaded in grey.

Figure B.2: Effects of recession narratives on economic and noneconomic sentiment



(a) Economic sentiment **(b)** Noneconomic sentiment

Notes: Panels (a) and (b) report β_h in basis points from estimating $\Delta_h y = \alpha + \beta_h \cdot \mathbb{1}(d, \text{recession}) + \varepsilon_{\text{idh}}$, where $y \in \{s_{id}^{\text{econ}}, s_{id}^{\text{non}}\}$ denotes the average sentiment change sentiment in tweets with and without economic discussion, respectively; and $\mathbb{1}(d, k)$ denotes an indicator variable of whether the loading of an article d on the recession narrative is above the cross-sectional mean. We estimate (18) separately for each horizon $h = 1, \cdots, 30$. Shaded areas represent 90% confidence intervals.

Table B.1: Top positive and negative scores: tweets on yield curve

Panel (a): Top negative tweets (most negative first)

	Tweet	Score	Sentiment
1	@USER @USER @USER Real recessions have real inverted yield curves. That really invert and stay there. Then the real Recession starts. Probably July, 2020 just in time for the election. Isn't that what the Deep State wants? But they'll blame it on "don't cry for me Argentina!"	0.211	negative
2	@USER: IT DIDN'T WORK: Despite the Fed, the yield curve is stuck in 'recession' mode, stocks are a mess, and manufacturing is ...	0.218	negative
3	@USER: Global mkts in bad mood after hawkish Fed cut. Stocks fell, yield curve flattened worryingly & dollar strengthened as ...	0.218	negative
4	@USER: It doesn't always mean a recession's coming, but you don't get a recession without an inverted yield curve. Therein lies the worr ...	0.225	negative
5	@USER: Economics can't be spun. An inverted yield curve is the sign of a sick economy. Period... Trump had tried to spin the ...	0.233	negative

Panel (b): Top positive tweets (most positive first)

	Tweet	Score	Sentiment
1	@USER: Nice article and agree 100%... the market is treating the "yield curve" inversion like the Ebola virus for stocks... REAL M...	0.677	positive
2	Japanese yen stands tall as US yield curve inversion stokes economic worries HTTPURL via @USER HTTPURL	0.668	positive
3	@USER: A simple graph does a better job of predicting recessions than the experts. @USER remind us why the yield curve matters ...	0.655	positive
4	@USER: U.S. yield curve flattens on supply, trade worries HTTPURL HTTPURL	0.651	positive
5	White House trade advisor Navarro: 'Technically we did not have a yield curve inversion' HTTPURL via @USER HTTPURL	0.634	positive

Notes: This table reports the top 5 positive and negative tweets about the yield curve classified by the naïve Bayes model described in Appendix Section D. User names and URLs have been anonymized to tokens "@USER" and "HTTPURL", respectively. Sentiment scores represent the probability of a tweet being positive and have a range of $[0, 1]$

Table B.2: Topics estimated with LDA: yield curve inversion

Topic 1 "Recession"		Topic 2 "Nonrecession"	
Term	**Probability**	**Term**	**Probability**
recession	0.016	year	0.052
rate	0.016	bond	0.048
yield	0.011	said	0.036
economy	0.011	bank	0.025
cut	0.010	yield	0.021
curve	0.010	market	0.016
year	0.009	minus	0.015
yield curve	0.009	investor	0.015
trump	0.008	note	0.014
inversion	0.008	five	0.013
growth	0.008	easing	0.013
say	0.008	monetary	0.012
economic	0.008	three	0.011
even	0.008	rate	0.011
would	0.008	bond market	0.010
bank	0.006	analyst	0.010
risk	0.006	longer dated	0.010
long	0.006	mortgage	0.010
aug	0.006	crisis	0.009
term	0.006	billion	0.009

Topic 3		Topic 4		Topic 5	
Term	**Probability**	**Term**	**Probability**	**Term**	**Probability**
yield	0.040	yield	0.024	year	0.025
curve	0.036	curve	0.021	yield	0.023
yield curve	0.026	year	0.016	curve	0.016
inversion	0.016	*recession*	0.014	china	0.015
inverted	0.016	inversion	0.013	*recession*	0.014
market	0.015	rate	0.013	treasury	0.012
year	0.013	treasury	0.009	bond	0.012
recession	0.012	market	0.008	economy	0.011
rate	0.010	time	0.008	trade	0.010
stock	0.010	yield curve	0.008	global	0.008
month	0.010	point	0.008	growth	0.008
economic	0.009	month	0.008	market	0.008
term	0.008	bond	0.007	even	0.008
investor	0.008	fed	0.007	inverted	0.007
bond	0.008	long	0.007	signal	0.007
energy	0.008	term	0.007	yield curve	0.007
u	0.007	inflation	0.006	time	0.007
longer	0.007	note	0.006	country	0.006
america	0.007	much	0.006	chinese	0.006
inverted yield	0.007	equity	0.006	cause	0.006

Notes: This table reports topics estimated with the LDA on articles of the yield curve with $K = 5$ and symmetric Dirichlet priors. For each topic, we report the distribution over vocabulary terms estimated with the LDA model.

Table B.3: Automated topic labelling with guided LDA

	(1)	(2)
	Tweet Sentiment Changes	
Recession narrative		
$\mathbb{1}(d,k)$	-0.44	
	(0.43)	
Nonrecession narrative		
$\mathbb{1}(d,k)$		0.44
		(0.43)
R^2	0.003	0.003
Observations	352	352

Notes: This table reports results from estimating $\Delta s_{id} = \alpha + \beta_k \cdot \mathbb{1}(d,k) + \varepsilon_{id}$, where topic $k \in \{\text{recession}, \text{nonrecession}\}$ is estimated with guided LDA as described in the main text. As in the baseline specification, Δs_{id} denotes changes in user i's tweet sentiment 24 hours around reading article d; and $\mathbb{1}(d,k)$ is an indicator variable for whether the loading of article d on narrative k is above the cross-sectional mean. Tweet sentiment is measured with naïve Bayes classifier and an article's loading on a narrative is measured with the LDA model, as described in the main text. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table B.4: Removing potential bots

	(1)	(2)	(3)	(4)	(5)	(6)
			Tweet Sentiment Changes			
Nonrecession narrative						
$\mathbb{1}(d,k)$	-1.45**		-1.40**			
	(0.72)		(0.70)			
$\theta(d,k)$		-1.96**		-1.86**		
		(0.92)		(0.90)		
Nonrecession narrative						
$\mathbb{1}(d,k)$	-0.13				0.14	
	(0.51)				(0.50)	
$\theta(d,k)$		-0.36				-0.03
		(0.72)				(0.70)
R^2	0.013	0.014	0.012	0.013	0.000	0.000
Observations	323	323	323	323	323	323
Exclude bots	yes	yes	yes	yes	yes	yes

Notes: This table reports results from estimating variants of the baseline specification in Table 3 while excluding users with top 5% average daily tweets. Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table B.5: Controlling for financial markets

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment Changes					
Recession narrative						
$\mathbb{1}(d,k)$	-1.13*		-1.26**			
	(0.65)		(0.63)			
$\theta(d,k)$		-1.63*		-1.62**		
		(0.83)		(0.80)		
Nonrecession narrative						
$\mathbb{1}(d,k)$	0.47				0.74	
	(0.60)				(0.58)	
$\theta(d,k)$		-0.01				0.32
		(0.67)				(0.65)
R^2	0.020	0.019	0.019	0.019	0.012	0.008
Observations	352	352	352	352	352	352
Financial controls	yes	yes	yes	yes	yes	yes

Notes: This table reports results from estimating variants of the baseline specification in (16) while controlling for macroeconomic and financial fluctuations. Column (1) reports β_r and β_{nr} from estimating the baseline specification

$$\Delta s_{id} = \alpha + \beta_r \cdot \mathbb{1}(d, \text{recession}) + \beta_{nr} \cdot \mathbb{1}(d, \text{nonrecession}) + \Gamma' Z_t + \varepsilon_{id},$$

where Δs_{id} denotes changes in user i's tweet sentiment 24 hours around reading article d; and $\mathbb{1}(d,k)$ for $k \in \{\text{recession}, \text{nonrecession}\}$ denotes an indicator variable for whether the loading of article d on narrative k is above the cross-sectional mean; Z_t is a vector of macro and financial controls including the S&P 500 and VIX indices. Tweet sentiment is measured with naïve Bayes classifier and an article's loading on a narrative is measured with the LDA model, as described in the main text. Column (2) reports β_r and β_{nr} from estimating $\Delta s_{id} = \alpha + \beta_r \cdot \theta(d, \text{recession}) + \beta_{nr} \cdot \theta(d, \text{nonrecession}) + \Gamma' Z_t + \varepsilon_{id}$, where $\theta(d,k)$ denotes the loading of article d on narrative k. Columns (3) through (6) report β from estimating univariate models $\Delta s_{id} = \alpha + \beta \cdot x_{dk} + \Gamma' Z_t + \varepsilon_{id}$, where x_{dk} is $\mathbb{1}(d, \text{recession})$, $\theta(d, \text{recession})$, $\mathbb{1}(d, \text{nonrecession})$, or $\theta(d, \text{nonrecession})$. Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table B.6: Controlling for days of the week

	(1) **Tweet**	(2) **Sentiment**	(3) **Changes**
Recession narrative			
$\mathbb{1}(d,k)$	-1.17*	-1.13*	-1.29**
	(0.65)	(0.67)	(0.65)
Nonrecession narrative			
$\mathbb{1}(d,k)$	0.23	-0.03	-0.05
	(0.55)	(0.48)	(0.51)
R^2	0.016	0.014	0.012
Observations	352	352	352
Day of the week control	Monday	Friday	Weekend

Notes: This table reports results from estimating the baseline specification in (16) while including an indicator variable that takes the value 1 if the quote retweet is posted on Monday, Friday, or weekend, respectively. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

C. Latent Dirichlet Allocation

Latent Dirichlet allocation (LDA) developed by Blei et al. (2003) is a generative probabilistic model that is aimed at reducing the dimensionality of text corpus. This section presents details of the model.

We represent each *word* from our vocabulary as a basis vector of length V with a single component equal to 1 and all other components equal to zero. For example, the vth word is denoted as $w = (0, \cdots, 0, 1, 0, \cdots, 0)$ where $w_v = 1$ and $w_u = 0$ if $u \neq v$. Then, an *article* is a vector consisting of N words, i.e., $w = (w_1, \cdots, w_N)$ where w_n is the nth word. Finally, A *corpus* is a collection of M articles, i.e., $D = \{w_1, \cdots, w_M\}$.

Consider a k-dimensional Dirichlet random variable θ with a parameter vector $\alpha = (\alpha_1, \cdots, \alpha_K)$, whose probability density over a $(k-1)$-simplex is given by

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \theta_1^{\alpha_1 - 1} \cdots \theta_k^{\alpha_k - 1} \tag{67}$$

where $\Gamma(x)$ is the Gamma function. Then, LDA assumes the following data generating process for each article d in our corpus D:

1. Draw $N \sim \text{Poisson}(\xi)$;
2. Draw $\theta \sim \text{Dirichlet}(\alpha)$;
3. Each word w_n is generated from a two-step process:
   (a) Draw a topic $z_n \sim \text{Multinomial}(\theta)$;
   (b) Draw a word w_n from $p(w_n|z_n, \beta)$, the multinomial probability conditioned on the topic;

where β denotes a k-by-V matrix with $\beta_{ji} = p(w_j = 1|z_i = 1)$ that represent word probabilities.

Given the parameters α, β, the distribution over a topic θ, a set of topics z, and a set of N words, the joint likelihood is given by

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^{N} p(z_n|\theta) p(w_n|z_n, \beta). \tag{68}$$

We can integrate over θ and sum over z to obtain the marginal distribution of an article as

$$p(w|\alpha, \beta) = \int p(\theta|\alpha) \left(\prod_{n=1}^{N} \sum_{z_n} p(z_n|\theta) p(w_n|z_n, \beta) \right), \tag{69}$$

and we can obtain the probability of a corpus by taking the product of all marginal probabilities of single documents

$$p(D|\alpha, \beta) = \prod_{d=1}^{M} \int p(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn}|\theta_d) p(w_{dn}|z_{dn}, \beta) \right) \tag{70}$$

The inference problem that we solve with the LDA is to compute the posterior distribution of the unobserved variables given a document:

$$p(\theta, z|w, \alpha, \beta) = \frac{p(\theta, z, w|\alpha, \beta)}{p(w|\alpha, \beta)} \tag{71}$$

where

$$p(w|\alpha, \beta) = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \gamma(\alpha_i)} \int \left(\prod_{i=1}^{k} \theta_i^{\alpha_i - 1} \right) \left(\prod_{n=1}^{N} \prod_{i=1}^{k} \prod_{j=1}^{V} (\theta_i \beta_{ij})_{w_n^j} \right) d\theta, \tag{72}$$

which we approximate using the online variational Bayes algorithm developed by Hoffman, Bach and Blei (2010).

Our text preprocessing is standard. We remove stop words such as "a" and "the", numbers, words with a single character, and capitalization. We reduce the dimensionality of the corpus by lemmatizing, grouping together words with different forms that express the same meaning into a single token (for example, "curve" and "curves" are both lemmatized to "curve").

D. Measuring Tweet Sentiment

Based on the tweets from users' timelines collected as described in the previous subsection, we estimate consumer sentiment using the naïve Bayes classifier developed by Rish et al. (2001). Using the Bayes law, the classifier represents the probability of the sentiment $y = \{0, 1\}$ of a tweet consisting of terms $(t_1, \cdots, t_n)$ as:

$$p(y|(t_1, \cdots, t_n) \propto p(y) \prod_{i=1}^{n} p(t_i|y) \tag{73}$$

As recognized by Buehlmaier and Whited (2018), naïve Bayes is one of the oldest tools in natural language processing and has better out-of-sample performance in text-based tasks than alternative models (Friedman et al., 2001). The special features in tweets require additional preprocessing. We convert all user mentions and links into single tokens (`@USER` and `HTTPURL`), remove special characters (`RT` and `FAV`), and fix common typos. For example, a raw tweet:

RT @UMich @UMichFootball: Victors valiant, champion of the west! https://umich.edu/

will be transformed to:

@USER @USER: victors valiant, champion of the west! HTTPURL

After pre-processing, we vectorize tweets using term-frequency inverse-document-frequency (tf-idf), which weighs a token by its importance to a document relative to the corpus (Ramos et al., 2003). The weighting is specified as:

$$\text{tf-idf}_{t,d} = \underbrace{\frac{w_{t,d}}{\sum_{\tau \in d} w_{\tau,d}}}_{\text{term frequency}} \cdot \underbrace{\log \frac{D}{|\{d \in D : t \in d\}|}}_{\text{inverse document frequency}} \tag{74}$$

where $w_{t,d}$ represent the frequency count of term t in document d, D represents the total number of documents, and $|\{d \in D : t \in d\}|$ is the number of documents term t appears. Tf-idf reduces the importance of words that appear with high frequency, such as "the" or "we."

Then we use the naïve Bayes algorithm to classify the sentiment of tweets. Specifically, we represent the probability that a tweet j conveys positive sentiment as a function of the tf-idf-weighted terms $t_1, \cdots, t_n$ of in the tweet:

$$\tilde{p}_j(\text{positive}) = f(t_1, \cdots, t_n) \tag{75}$$

where tildes indicate that the probability $\tilde{p}$ is predicted by the naïve Bayes classifier.

We pre-train the naïve Bayes classifier using 100,000 pre-classified tweets in Go, Bhayani and Huang (2009), who use emoticons to automatically classify the sentiment of tweets as positive and negative. For example, smiley faces `:)` indicate positive tweets, and sad faces `:(` indicate negative tweets.

Based on the predicted sentiment from the naïve Bayes classifier, we define the sentiment of consumer i in day t as:

$$s_{it} = \frac{1}{J}\sum_j \tilde{p}_j(\text{positive}) \quad \text{for } j \text{ posted in day } t \tag{76}$$

where s_{it} measures the average sentiment of tweets posted by the consumer in a day. Values of s_{it} lie between 0 and 1, with values greater than 0.5 corresponding to positive sentiment. The higher the values of s_{it}, the more optimistic a consumer is of the outlook.

E. Incomplete Passthrough from Media to Household Narratives

In our data, we cannot observe people's narratives directly. We only observe whether someone was exposed to a particular narrative in news media. We therefore examine what happens when households in the model experience this narrative exposure, to guide the interpretation of our empirical results.

Since our data concerns small windows around Twitter user interactions with yield curve narratives, we suppose for this exercise that over the window there are no changes in y_t. However, there is a change in z_t: the yield curve inverts. We normalize z_t such that the inversion implies it goes from 0 to 1 over the window.[26]

[26]Note some people in our sample engage with a yield curve narrative in the week before, or two weeks after, the event taking place. We are therefore assuming that the news article also makes them aware of the inversion, even though it hasn't quite happened yet, or happened in the recent past.

To explore how expected income changes with exposure to each narrative in this window, it will be useful to note that since the model is linear with Gaussian uncertainty, expectations formed with either narrative are linear combinations of realized variables:

***Lemma* 3** (narrative expectations)**.** *The conditional expectations formed using each narrative* $\{b, z\}$ *are such that*

$$E_t^b y_{t+1} = h_y^b y_t \tag{77}$$

$$E_t^z y_{t+1} = h_y^z y_t + h_z z_t \tag{78}$$

Proof. Appendix A. □

A household who uses the baseline narrative throughout the window does not therefore change their income expectations. In contrast, a household who uses the z narrative throughout sees their expectations change by $\partial\, \mathbb{E}_t^z\, y_{t+1}/\partial z_t = h_z$. Finally, households who switch narratives during the window change expectations according to

$$\Delta\, \mathbb{E}_t^{z,b}\, y_{t+1} = (h_y^b - h_y^z) y_t \tag{79}$$

$$\Delta\, \mathbb{E}_t^{b,z}\, y_{t+1} = (h_y^z - h_y^b) y_t + h_z \tag{80}$$

where $\Delta\, \mathbb{E}_t^{k,j}$ denotes the change in expectations for a household starting the window on narrative k and ending it on narrative j.

$$\Delta\, \mathbb{E}_t^{k,j}\, y_{t+1} \equiv \mathbb{E}_t^j\, y_{t+1} - \mathbb{E}_t^k\, y_{t+1} \tag{81}$$

Along with any direct effect of the yield curve inversion on expectations, these switchers may also therefore update expectations if the coefficients h_y^k differ across narratives k. Intuitively, when a household switches to a new narrative, they re-estimate those coefficients, using the specification implied by their new narrative. If the coefficients change as a result, then that household will change the way they extrapolate from their existing observations of y_t to y_{t+1}. These equations therefore pin down how expectations change over the window, conditional on a household's narratives.

To introduce exposure to particular narratives in news media, we proceed with two further assumptions.

1. (No-defiers). If a household begins a period with narrative k, and they are exposed to the same narrative k, they do not change their narrative. They end the period with narrative k with probability 1.

2. (Compliers). If a household begins a period with narrative k, and they are exposed to the alternative narrative j, they are 'infected' and switch to narrative j with probability $\phi \in (0, 1]$.

With these assumptions, the average expectation change among households exposed to the baseline narrative is given by

$$\bar{\Delta}\,\mathbb{E}_t(y_{t+1}|\text{exposed to } b) = \Pr(\text{start on } z|\text{exposed to } b)\left[(1-\phi)h_z + \phi\Delta\,\mathbb{E}_t^{z,b}\,y_{t+1}\right] \tag{82}$$

Within this group, only those who start the window using the z narrative change expectations, as all those starting with the baseline narrative keep it and therefore do not react to the yield curve inversion. A fraction ϕ of those households switch to the baseline narrative on exposure, while $1 - \phi$ do not switch, but update expectations because the news article alerts them to the inversion.

Similarly, the average expectation change among households exposed to the z narrative is given by

$$\begin{aligned}\bar{\Delta}\,\mathbb{E}_t(y_{t+1}|\text{exposed to } z) = \Pr(\text{start on } z|\text{exposed to } z)h_z \\ + \Pr(\text{start on } b|\text{exposed to } z)\phi\Delta\,\mathbb{E}_t^{b,z}\,y_{t+1}\end{aligned} \tag{83}$$

Here those who start on the z narrative all react to the news of the inversion, as do those who start on the baseline narrative and switch.

To make further progress on how narrative switching affects the average expectation changes in each group, we use equations (79) and (80) to note that:

$$\Delta\,\mathbb{E}_t^{b,z}\,y_{t+1} = h_z - \Delta\,\mathbb{E}_t^{z,b}\,y_{t+1} \tag{84}$$

This reflects the fact that the re-estimation of coefficients h_y^k when a household switches from the baseline narrative to the z narrative is the exact opposite of the re-estimation done by a household switching in the other direction.

Substituting equations (82) and (84) into equation (83), we obtain:

$$\bar{\Delta}\,\mathbb{E}_t(y_{t+1}|\text{exposed to } z) = h_z - \frac{\Pr(\text{start on } b|\text{exposed to } z)}{\Pr(\text{start on } z|\text{exposed to } b)} \cdot \bar{\Delta}\,\mathbb{E}_t(y_{t+1}|\text{exposed to } b) \quad (85)$$

Equations (82) and (85) therefore imply that the average expectation change among households exposed to a particular narrative is equal to the response of expectations to z_t under that narrative (0 and h_z for the baseline and z narratives respectively), plus a bias. That bias in each case comes from the fact that some households are exposed to narratives which they did not already hold at the start of the window: exposure is not a pure observation of a household's narrative. Importantly, the biases in the groups exposed to each narrative are proportional to one another.

In Table 3, we find that the average effect of exposure to the baseline narrative is close to zero, and not significant. Since the model predicts that observation is purely bias, this suggests the bias for that group is small. In addition, equation (85) then implies the bias is also small in the group exposed to the z narrative. We can therefore interpret the effect of exposure to the z narrative as capturing $\partial E_t^z y_{t+1}/\partial z_t$, as long as $\Pr(\text{start on } z|\text{exposed to } b)$ is not also close to zero. Although not directly testable, this is plausible given our data. The main thrust of many articles containing the baseline narrative, such as Peter Coy's Bloomberg article quoted in Section 4, is that recent developments are what has led to the yield curve losing its association with recessions. It is therefore likely that the recession narrative was prevalent at the start of the window, even among those exposed to the baseline narrative.