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

Alistair Macaulay
University of Oxford

Wenting Song
Bank of Canada

This version: December 22, 2022
First version: June 29, 2022

Abstract

This paper studies the role of narratives for macroeconomic fluctuations. Micro-founding narratives as directed acyclic graphs, we show how exposure to different narratives can affect expectations in an otherwise-standard macroeconomic framework. We capture such competing narratives in news media reports on the US yield curve inversion, using techniques in natural language processing. Linking media narratives to social-network 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 a model with frictions in financial intermediation, these effects of narrative-driven beliefs create a novel trade-off: extended periods of quantitative easing make narrative-driven waves of pessimism more frequent, but smaller in magnitude.

*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, John Leahy, Michael McMahon, Pablo Ottonello, Christopher Roth, Robert Shiller, Ran Spiegler, Leif Anders Thorsrud, and participants at various seminars and conferences for useful 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.

1. Introduction

Many decisions made by households and firms are influenced by expectations of future macroeconomic developments, but the factors determining these developments are often varied and nuanced. To help individuals make such decisions, popular narratives in the media provide simple causal stories of how variables affect one another. Shiller (2017) tracks the spread of certain "viral" narratives, which could cause or exacerbate existing macroeconomic fluctuations through their effect on expectations. To what extent do these narratives affect expectations once they have spread? And consequently, how substantial are the aggregate effects of shifting popular narratives about the macroeconomy?

In this paper, we answer those questions by linking traditional newspaper articles of macroeconomic events to engagement with that coverage on social media. Our theoretical framework specifies narratives as *directed acyclic graphs* (DAGs) (as in Eliaz and Spiegler, 2020; Andre, Haaland, Roth and Wohlfart, 2022b), which drive fluctuations in expectations. Motivated by the theory, we capture competing narratives in traditional news media using natural language processing and trace the influence of those narratives by comparing the sentiment of Twitter users before and after engaging with a particular narrative. Focusing on an episode of yield curve inversion in the US, we provide empirical evidence that exposure to a narrative associating the inversion with an imminent recession leads to a display of more pessimistic sentiment.

Our paper begins by developing a theoretical framework of how narratives affect expectations. We start with a textbook consumption-and-saving problem faced by households and specify narratives as *directed acyclic graphs* (DAGs), or network representations of the underlying models, which have natural interpretation as "causal" stories. We consider two competing narratives: a baseline narrative in which expectations of future income depend on the current income and interest rate, and an extraneous narrative in which expectations of future income also depend on an extraneous variable, such as a popular recession indicator. The key insight from this model is an equivalence result: While the extraneous variable can enter into a household's narrative about future incomes in a variety of ways—as a *shock* affecting future income or as a *signal* of other variables—the resulting DAGs have observationally equivalent effects on expectations. For this subset of narratives, we do not need to

distinguish between the direction of causality; it is sufficient to identify whether a link exists between the extraneous variable and remaining variables.

Motivated by the theoretical framework, we measure narratives as the media's competing interpretations of the same economic event. We collect news articles devoted to an economic event and use topic models from natural language processing to extract narratives surrounding the event. We obtain empirical estimates of both the prevailing narratives and each article's reliance on the narratives. Using these narratives, we provide empirical evidence on the importance of narratives for sentiment fluctuations. To isolate the effects of narratives, we focus on an episode of yield curve in version in 2019——a popular recession indicator in the US with a nebulous theoretical foundation. Two competing narratives emanate from major news outlets: a "recession" narrative that links the inverted yield curve to an imminent recession and a "nonrecession" narrative that makes no such connection.

Our main analysis studies the effects of narratives on the readers who are exposed. The most novel part of our data is the link from narratives in newspaper coverage to rich social network data from Twitter, which allows us to measure the spread of narratives. We use retweeting activities on Twitter to trace whether a Twitter user has engaged with news articles containing certain narratives. We find that after users are exposed to the recessionary narrative, their posted tweets display a more pessimistic sentiment, while exposure to the more neutral, nonrecessionary narrative has no such effect. The drop in sentiment following engagement with a recessionary narrative is persistent, remaining significant 30 days after the retweet.

To assess the potential for viral narratives to drive aggregate sentiment, we then turn to a quantitative model with narrative beliefs, informed by our empirical results. As in the data, households are split between two narratives. In the first, yield curve inversions are irrelevant for expectations of future real variables; the second links inversion events to future incomes. We find that these narratives generate a novel trade-off for monetary policy. Quantitative easing flattens the yield curve, which simultaneously increases the frequency with which shocks cause the yield curve to invert, and increases the prevalence of the baseline narrative in which inversions are irrelevant. Expansions of QE therefore increase the frequency of narrative-driven waves of pessimism, but decrease their magnitude.

Related literature Our paper relates to four strands of the literature. First, a growing literature pioneered by Shiller (2017) studies the role of narratives in economics.[1] Our contribution to this literature is twofold. Theoretically, we microfound narratives in a macroeconomic framework building on the Bayesian network literature (Spiegler, 2016, 2020a; Eliaz and Spiegler, 2020). 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.

Our empirical methodology complements the semantics-based approach (e.g., Ash, Gauthier and Widmer, 2021; Goetzmann, Kim and Shiller, 2022) that captures causal directions in narratives. We instead leverage the theoretical insight that DAGs with the same skeletons are observationally equivalent and use interpretable the topic model to capture narratives. Using similar methodology of topic models, Larsen and Thorsrud (2019) extract narratives as significant economic events based on newspaper articles; Flynn and Sastry (2022) measure narratives as varying sentiment on macroeconomic topics. We instead capture narratives as news media's competing interpretations of the *same* underlying economic event, building on models of competing narratives (Eliaz and Spiegler, 2020) and our empirical framework.

We provide empirical evidence on the importance of narratives by studying narratives around the yield curve inversion in 2019, complementing survey-based evidence from Andre et al. (2022b) and Kendall and Charles (2022), who conduct survey experiments to establish the causal effects of narratives on expectations, and Macaulay (2022) who presents evidence from UK household surveys on the importance of inflation narratives. Our empirical framework has the benefit of providing an ongoing measure of narratives outside of existing surveys, as illustrated with an application to inflation narratives by Macaulay and Song (2022), documenting the evolving prevalence and impact of competing inflation narratives.

Second, narratives provide a way for individuals to interpret economic news and translate that into expectations, and therefore also relate to studies of differences of opinion (Harris and Raviv, 1993; Patton and Timmermann, 2010; Xiong and Yan, 2010; Atmaz and Basak, 2018) and subjective models (Dräger, Lamla and Pfajfar, 2016; Andrade, Gaballo, Mengus and Mojon, 2019; Molavi, 2019; Andre, Pizzinelli, Roth and Wohlfart, 2022a).

This paper also relates to the broader literature of belief formation. Empirical evidence

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

documents the 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), learning from personal experiences (Malmendier and Nagel, 2016; Kozlowski, Veldkamp and Venkateswaran, 2020), 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.[2]

Third, we relate to the literature on sentiment and media. Our results highlight the role of economic narratives in shaping household sentiments, which are important sources of macroeconomic fluctuations (see, for example, Angeletos and La'O, 2013; Greenwood and Shleifer, 2014; Levchenko and Pandalai-Nayar, 2020; Maxted, 2019; Krishnamurthy and Li, 2020; Acharya, Benhabib and Huo, 2021). We contribute to the literature by showing that narratives constructed by the media provides a microfoundation for fluctuations in sentiment. We highlight, in particular, the role of media in curating news and constructing narratives, consistent with theories of news media as optimizing agents whose news reporting drives aggregate fluctuations (Nimark, 2014; Chahrour, Nimark and Pitschner, 2021).

Lastly, our unique data linking news coverage to its influence on social media allows us to measure the impact of narratives constructed by the media on household beliefs, which relates to the growing literature that uses unstructured data sources to study the economic effects of news (see, for example, Calomiris and Mamaysky, 2019; Bybee, Kelly, Manela and Xiu, 2020; Nyman, Kapadia and Tuckett, 2021), and 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).

Outline The rest of the paper proceeds as follows: in Section 2 we present our theoretical framework that connects narratives with expectations and derive conditions for observationally-equivalent narratives; in Section 3 we describe the episode of yield curve inversion in 2019; in Section 4 we describe our data and sample; in Section 5 we conduct our

[2]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).

main empirical analysis on the narratives surrounding the yield curve inversion by linking news articles and social media; in Section 6 we use these results in a quantitative model to explore the consequences of quantitative easing for narrative-driven fluctuations in sentiment; Section 7 concludes.

2. Model

In this section we develop a framework to analyse the role of narratives in shaping household expectations and actions, in an otherwise standard consumption-saving problem. A narrative is defined as a causal ordering of variables, represented by a DAG. Importantly, we show that certain groups of narratives are observationally equivalent, in that they always produce the same household expectations. This will guide our approach to distinguishing between relevant narratives in text data.

2.1. Households

Each household chooses consumption to maximise their life-time utility subject to the budget constraint under the expectation $\mathbb{E}_{it}$, taking the interest rate and income as given. Each household i has preferences over consumption given by

$$\sum_{s=0}^{\infty} \beta^s \, \mathbb{E}_{it} \, u(C_{it+s}) \tag{1}$$

where β is the discount factor; $\mathbb{E}_{it}$ is the subjective expectation of household i given the time-t information set; and the instantaneous utility function is CRRA, specified as

$$u(C_{it}) = \frac{C_{it}^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} \tag{2}$$

Each period, the household receives real income Y_t, and can purchase one-period bonds B_{it} with a real interest rate of R_t. Their budget constraint is therefore given by

$$C_{it} + B_{it} = R_{t-1} B_{it-1} + Y_t$$

For simplicity, we take income Y_t to be exogenous to household i's decisions. This is relaxed

in the quantitative model in Section 6.

The optimization leads to a standard consumption Euler equation. Log-linearizing the Euler equation and the budget constraint about a steady state in which $\beta R = 1$ and $B_i = 0$ gives the household's time-t consumption function as

$$c_{it} = (1-\beta)\sum_{s=0}^{\infty}\beta^s\,\mathbb{E}_{it}\,y_{t+s} - \sigma\beta\sum_{s=0}^{\infty}\beta^s\,\mathbb{E}_{it}\,r_{t+s} \tag{3}$$

where lower case c_{it}, y_t, r_t denote log-deviations of consumption, income, and real interest rates from their respective steady states.

Equation (3) shows that households' current consumption is driven by their expectations of future real income and real interest rates. Households observe the history of y, r up to the current period, but to form expectations of future realizations they must combine this with a belief about the evolution of both variables. We introduce narratives as the source of these beliefs relating observations to expectations.

2.2. Narratives

We follow Eliaz and Spiegler (2020) and Andre et al. (2022b) and define a narrative as a *directed acyclic graph* (DAG), that defines a series of causal relationships between variables. For a thorough review of this approach to modeling expectations, see Spiegler (2020a).

***Definition* 1** (narrative as a DAG)**.** *A narrative for income and interest rates 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.*

The set of nodes $\mathcal{N}$ contains current and future values of y and r, and potentially other additional variables. The links $\mathcal{L}$ are acyclic: they are such that the graph contains no directed path from a node back to itself.

The nodes of the DAG correspond to variables in the household environment. The links correspond to perceived causal relationships between those variables.

These links are crucial in the household's decision problem. To choose consumption, the household forms expectations of future income and interest rates, *conditional* on observed

current variables. To form that conditional expectation, they require a belief about the joint distribution of the variables involved. The narrative guides that belief, through the Bayesian factorization formula

$$\tilde{p}(x_{\mathcal{N}}) = \prod_{n \in \mathcal{N}} p(x_n | x_{\mathcal{L}(n)}) \tag{4}$$

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

A narrative therefore specifies which conditional distributions should be involved in forming their beliefs about the joint distribution of all variables $\tilde{p}(x_{\mathcal{N}})$, which may or may not equal the true joint distribution. The perceived causal links between variables imply a series of conditional independence assumptions, that will affect how the households interpret data on the variables in their environment. We follow Eliaz and Spiegler (2020) and assume that the households observe a long time series of such data on each variable, and so are able to accurately recover the true conditional distributions $p(x_n | x_{R(n)})$ involved in this factorization.

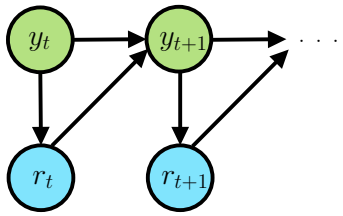
Expectations are then formed using the perceived joint distribution $\tilde{p}(x_{\mathcal{N}})$. This means that if the household's narrative correctly accounts for the true causal links between variables, Equation (4) yields the true joint distribution of the variables in their environment. Such a household will therefore have rational expectations.

However, if the narrative incorrectly specifies the true causal links between variables, the implied $\tilde{p}(x_{\mathcal{N}})$ may not coincide with the true joint distribution. A household with such a narrative interprets data through the lens of a misperceived causal model, which may cause them to use incorrect assumptions about the conditional (in)dependence of certain variables. That, in turn, may generate incorrect beliefs about the joint distribution of variables in their environment. In that case, the expectations of these households will not coincide with rational expectations.

Baseline Narrative The first narrative we consider is displayed in Figure 1. In this narrative, real income is persistent, so y_t has a causal effect on y_{t+1}. In addition, real incomes affect contemporaneous interest rates, for instance because the central bank reacts to demand conditions through a standard Taylor rule. Changes in real interest rates then in turn

affect real incomes with a lag, so r_t has a causal effect on y_{t+1}.[3] For now, we leave it unspecified whether this narrative represents a correct understanding of causal relationships in the equilibrium of the economy, or whether it is a misperception of true economic relationships.

Figure 1: DAG representation of the baseline narrative



We refer to this narrative as the "baseline narrative". The only variables that matter for expectations of y_t and r_t are lags of those variables themselves. Formally, the baseline narrative is defined as follows:

***Definition* 2** (baseline narrative)**.** *Let nRm denote a directed link from node n to node m. The baseline narrative is a DAG consisting of:*

1. *the set of nodes, $\mathcal{N} = \{y_s, r_s\}_{s=t}^{\infty}$; and*
2. *the set of links, $\mathcal{L} = \{y_s R y_{s+1}, y_s R r_s, r_s R y_{s+1}\}$.*

Extraneous Narratives We now introduce a competing group of narratives, which introduce an extraneous variable, z, into the causal ordering of variables.

These "extraneous narratives" could reflect true causal relationships in the economy, or the extraneous variable could be entirely spurious. Politicians or news media may have incentives to create such spurious narratives to influence expectations or household behavior (Gentzkow and Shapiro, 2008; Eliaz and Spiegler, 2020).

In particular, we consider a class of narratives in which z_s is perceived to be related to real income in periods s and $s+1$.

***Definition* 3** (extraneous narratives)**.** *The extraneous narratives are DAGs consisting of:*

1. *the set of nodes, $\mathcal{N} = \{y_s, r_s, z_s\}_{s=t}^{\infty}$; and*
2. *one of the sets of links $\mathcal{L}_a$, $\mathcal{L}_b$, or $\mathcal{L}_c$, where:*

[3] This lag is important to ensure that the graph remains acyclic, as required by Definition 1.

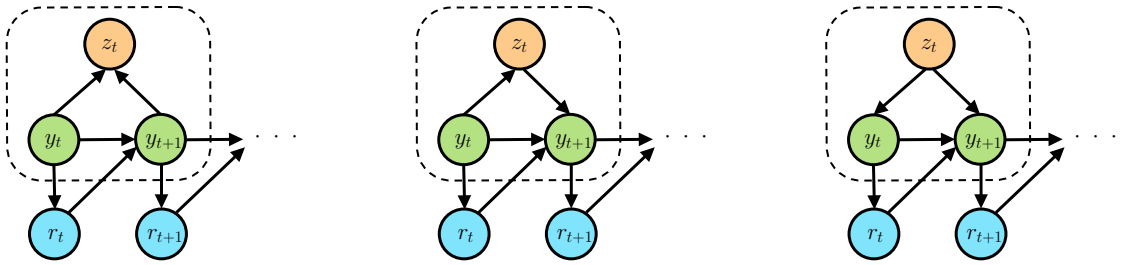
(a) $\mathcal{L}_a = \mathcal{L} \cup \{y_s R z_s, y_{s+1} R z_s\}$;

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

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

Importantly, even restricting extraneous narratives to this class, there are still three possible ways for z to enter the household's causal model. These are shown in Figure 2.

Figure 2: DAG representations of extraneous narratives



(a) "Signal" **(b)** "Channel" **(c)** "Shock"

In the narrative in Panel (a), z is caused by the income process and is a symptom of the underlying economic fundamentals. It therefore signals changes in income without being a cause of that change. In the narrative in Panel (b), z is a channel through which the current income affects the future income. In the narrative in Panel (c), z is an exogenous shock that affects income.

We now go on to derive the processes for expectations implied by these narratives.

2.3. Expectations

To find how narratives affect expectations, we first find the Bayesian factorization formulae for each of the narratives described above.

For the baseline narrative, we have

$$\tilde{p}(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)p(y_{s+1}|r_s, y_s)p(z_s) \tag{5}$$

The equivalent factorization formulae for the three extraneous narratives are

$$\tilde{p}_a(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)p(y_{s+1}|r_s, y_s)p(z_s|y_s, y_{s+1}) \tag{6}$$

$$\tilde{p}_b(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)p(y_{s+1}|r_s, y_s, z_s)p(z_s|y_s) \tag{7}$$

$$\tilde{p}_c(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s|z_s)p(y_{s+1}|r_s, y_s, z_s)p(z_s) \tag{8}$$

Systematically distinguishing between these different extraneous narratives in media would be challenging. However, despite the different structural interpretations of these three DAGs, Proposition 1 shows that their effects on households' beliefs are in fact observationally equivalent.

***Proposition* 1** (observational equivalence of extraneous narratives)**.** *The Bayesian factorization formulae for the three extraneous narratives are equivalent:*

$$\tilde{p}_a(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = \tilde{p}_b(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = \tilde{p}_c(r_s, r_{s+1}, y_s, y_{s+1}, z_s)$$

Proof. Appendix A. □

Intuitively, if one household believes that a rise in the variable z causes incomes to fall ("shock"), and another believes instead that falling incomes cause z to rise ("signal"), then both will revise their expected incomes down when they observe higher z. Formally, this property emerges because all of the extraneous narrative DAGs are "perfect": the direct causes of any downstream variable are all themselves directly linked together ("all parents are married"). All perfect DAGs with the same skeleton necessarily share the same Bayesian factorization formula (Verma and Pearl, 1990).[4]

Proposition 1 implies that we do not need to consider the three extraneous narratives in Figure 2 separately. From here, we therefore refer to *the* extraneous narrative to mean any narrative satisfying Definition 3.

The next Proposition shows that, in general, the baseline narrative and the extraneous narrative generate different Bayesian factorization formulae.

[4]See Spiegler (2020b) for a detailed discussion of the implications of perfection in DAGs used to represent the causal mental models of decision-makers in a variety of economic contexts.

***Proposition* 2** (nonequivalence of baseline and extraneous narratives)**.** *If z_s is correlated with y_s and/or y_{s+1}, then the following two Bayesian factorization formulae are nonequivalent:*

1. $\tilde{p}(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)p(y_{s+1}|r_s, y_s)p(z_s)$

2. $\tilde{p}_a(r_s, r_{s+1}, y_s, y_{s+1}, z_s) = p(r_s|y_s)p(r_{s+1}|y_{s+1})p(y_s)p(y_{s+1}|r_s, y_s)p(z_s|y_s, y_{s+1})$

Proof. Appendix A. □

Note that the narratives only imply different perceived joint distributions if z is correlated with current or future real incomes. However, this does not necessarily imply that the extraneous narrative describes the true causal relationships between the variables, as this reduced-form correlation may be present even if there is no true causal relationship between z and y. Some other variable not in any household's narrative, for example, could simultaneously cause both variables. In that case households believing the extraneous narrative are basing their expectations on a spurious correlation.

This model therefore predicts that a household's expectations depend on which narrative they are exposed to. Using the extraneous narrative, expectations of real income one period ahead are

$$\mathbb{E}^e_{it}(y_{t+1}|\mathcal{I}_t) = \int y_{t+1} p(y_{t+1}|r_t, y_t, z_t) dy_{t+1} \tag{9}$$

In contrast, using the baseline narrative, the same expectation is

$$\begin{aligned} \mathbb{E}^b_{it}(y_{t+1}|\mathcal{I}_t) &= \int y_{t+1} p(y_{t+1}|r_t, y_t) dy_{t+1} \\ &= \int \int y_{t+1} p(y_{t+1}|r_t, y_t, z_t) p(z_t|r_t, y_t) dz_t dy_{t+1} \end{aligned} \tag{10}$$

From Proposition 2, these are different whenever the extraneous variable z is correlated with real income. In particular, expectations react to realized z_t under the extraneous narrative, but do not react under the baseline narrative.

$$\frac{\partial \mathbb{E}^e_{it}(y_{t+1}|\mathcal{I}_t)}{\partial z_t} \neq 0, \quad \frac{\partial \mathbb{E}^b_{it}(y_{t+1}|\mathcal{I}_t)}{\partial z_t} = 0 \tag{11}$$

2.4. Production and Market Clearing

We now close the model to examine the effects of these competing narratives in equilibrium. The focus of this model is the role of expectations, so we keep the supply side simple. All income from production flows equally to all households, so real income is equal to real output. We consider an output process which nests situations in which both the baseline and extraneous narratives capture the true dynamics of output:

$$y_t = \rho y_{t-1} + \gamma c_{t-1} + \mu z_{t-1} + v_t^y \tag{12}$$

where c_{t-1} denotes aggregate consumption, and the shock $v_t^y \sim i.i.d.N(0, \sigma_y^2)$.

If $\gamma = \mu = 0$, output follows a simple AR(1) process and is unaffected by any other variables. This would capture an endowment economy, or an economy where output is solely determined by exogenous technology. In this case neither the baseline nor extraneous narratives correctly capture the y_t process, as both contain causal links from r_t to y_{t+1}.

If $\gamma > 0$, output is affected by past consumption demand. This could reflect, for example, an economy in which firms have some market power, but can only increase production in response to demand with a lag. In this case, different combinations of parameters, and different distributions of narratives among households, can render the model consistent with either the baseline or extraneous narrative.

For the goods market to clear, this output must equal aggregate consumption each period. Letting λ_t denote the proportion of households using the baseline narrative in period t, this market clearing condition is

$$c_t \equiv \lambda_t c_t^b + (1 - \lambda_t) c_t^e = y_t \tag{13}$$

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

Finally, we specify a process for the extraneous variable

$$z_t = \chi y_t + \nu_t^z, \quad \nu_t^z \sim i.i.d.N(0, \sigma_z^2) \tag{14}$$

This is consistent with the baseline narrative if $\chi = 0$, and with the extraneous narrative if

$\chi \neq 0$.

All that remains to close the model is now to specify how households arrive at the conditional probabilities used to form their expectations.

2.5. Narrative Equilibrium

Households form expectations by fitting their narratives to long histories, of data, which enables them to accurately estimate the conditional distributions that feature in their narrative (i.e. Equations (9) and (10)). Lemma 1 shows that, due to the fact the model is linear and all shocks are Gaussian, these conditional distributions are such that expectations under both baseline and extraneous narratives can be written as a linear combination of realized y_t, r_t, z_t.

***Lemma* 1** (narrative expectations)**.** *The conditional expectations formed using each narrative $k \in \{b, e\}$ are such that*

$$\mathbb{E}_t^k \boldsymbol{x}_{t+s} = \mathcal{H}_k^s \boldsymbol{x}_t \tag{15}$$

where $\boldsymbol{x}_t = (y_t, r_t, z_t)'$ and $\mathcal{H}_k$ is a 3×3 matrix of coefficients. In the baseline narrative $k = b$, the final row and column of $\mathcal{H}_b$ consist of all zeros.

Proof. Appendix A. □

With this result, we can define the full equilibrium of the model.

***Definition* 4** (narrative equilibrium)**.** *Given a distribution of households across narratives λ_t, and shocks v_t^y, v_t^z, a narrative equilibrium consists of $c_t^b, c_t^e, r_t, y_t, z_t$, narrative coefficients Ω^b, Ω^e, and expectations $\mathbb{E}_{it}^b(r_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_{it}^b(y_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_{it}^e(r_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_{it}^e(y_{t+s}|\mathcal{I}_t)$, such that:*

1. *Given prices and expectations, households maximize (1) subject to (2);*
2. *Output is determined according to (12);*
3. *The extraneous variable is determined according to (14);*
4. *The goods market clears according to (13);*

5. *Expectations are determined according to* (15)*, where the narrative coefficients* $\mathcal{H}_k$ *are functions of the distributions of* y_t, r_t, z_t*, defined in Appendix A.*

In general, solving for $\mathcal{H}_b, \mathcal{H}_e$ involves a system of nonlinear equations with no general analytic solution. There are, however, two cases in which there is a unique analytic solution, which provide insight into the nature of narratives in equilibrium.

***Proposition* 3** (equilibrium with $\mu = 0$)**.** *If* $\mu = 0$*, the narrative equilibrium in Definition 4 is such that*

$$\mathcal{H}_b = \begin{pmatrix} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{16}$$

$$\mathcal{H}_e = \begin{pmatrix} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & 0 \\ h_{31} & h_{32} & 0 \end{pmatrix} \tag{17}$$

where h_{11} *to* h_{32} *are combinations of parameters defined in Appendix A.*

Proof. Appendix A. □

In this case, with $\mu = 0$ in equation (12), there is no direct channel through which the extraneous variable z_t affects future output y_{t+1}. This does not immediately rule out all feedback from z_t to y_{t+1}, as z_t may still in principle affect aggregate consumption. However, in equilibrium, this indirect channel is also absent. That is, even households who believe the extraneous narrative converge on the view that z_t has no impact on y_{t+1}, and as such they do not condition their consumption decisions on z_t. Those who believe the baseline narrative also do not condition on z_t by construction, so z_t has no effect on y_{t+1}, which confirms the parameter estimates in the extraneous narrative. The baseline narrative therefore accurately describes aggregate dynamics, and all households converge on this. Equilibrium coincides with the rational expectations equilibrium, and narratives have no effect on sentiment: expectations are identical for both groups of households. The only difference in the narrative coefficients between groups comes from the perceived effect of y_t and r_t on z_{t+1},

but as no household believes z_{t+1} has any effect on any other variables of interest that makes no difference to expectations.

The reason expectations coincide across narratives is that the extraneous narrative nests the baseline narrative in our setup. If the baseline narrative is correct, then eventually households who believe the extraneous narrative will converge to it. This convergence may be slow, but we leave the study of such models in the transition to our narrative equilibrium for further study.

There are, however, cases in which narratives do affect sentiment.

***Proposition* 4** (equilibrium with $\chi = 0$)**.** *If $\chi = 0$, the narrative equilibrium in Definition 4 is such that*

$$\mathcal{H}_b = \begin{pmatrix} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{18}$$

$$\mathcal{H}_e = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 0 & 0 & 0 \end{pmatrix} \tag{19}$$

Proof. Appendix A. □

In this case z_t is a simple shock variable. It may affect y_{t+1}, but it is not affected by y_t. In this case, the equilibrium coefficients of the baseline and extraneous narratives are identical, *except* for the coefficient on z_{t-1} in the perceived processes for y_t, r_t.

Households who believe the extraneous narrative have rational expectations in this case,[5] but households using the baseline narrative do not. They miss the dependence of y_{t+1} on z_t, and so ignore information on z_t in forming expectations of y_{t+1} and other variables.

Narratives therefore matter for expectations, consumption, and output in this environment. In equilibrium, the effects of a shock to z_t differ across households with different

[5]Strictly, the extraneous narrative does not exactly coincide with the true model, as it does not contain any passthrough from z_t to r_t, which will be present in equilibrium. However, as $\mathbb{E}_t^e z_{t+1} = 0$ in all periods, this has no effect on expectations.

narratives, as in equation (11).

$$\frac{\partial \mathbb{E}^e_{it}(y_{t+1}|\mathcal{I}_t)}{\partial z_t} = \frac{\mu}{1-\beta\gamma(1-\lambda_{t-1})} \neq 0, \quad \frac{\partial \mathbb{E}^b_{it}(y_{t+1}|\mathcal{I}_t)}{\partial z_t} = 0 \tag{20}$$

Similarly, for any given realization of z_t, a change in the distribution of narratives across households will cause a change in aggregate consumption, and so in output.

$$\frac{\partial c_t}{\partial \lambda_t} = -\frac{\beta\mu}{(1-\beta\gamma(1-\lambda_t))^2} z_t, \quad \frac{\partial y_{t+1}}{\partial \lambda_t} = \gamma \frac{\partial c_t}{\partial \lambda_t} \tag{21}$$

This comes through two channels. First, the change in the distribution of narratives affects aggregate expectations, and so directly affects aggregate consumption. Second, there is an indirect effect through the equilibrium narratives. If more households start to hold the extraneous narrative (lower λ_t), then more households react to changes in z_t, meaning aggregate consumption and output are more responsive to z_t. In equilibrium, this will lead to households who believe the extraneous narrative estimating a stronger passthrough from z_t to y_{t+1}, and therefore reacting even more strongly to z_t shocks. This is why λ_{t-1} is present in equation (20): the distribution of narratives in the population affects not just the average response to information, but also the behavior of expectations at the individual level.

Overall, the lesson from these two tractable cases is that narratives only affect expectations in equilibrium when the baseline narrative is misspecified. Otherwise, the extraneous narrative converges on the baseline narrative, and all households hold identical rational expectations.

2.6. Link to Empirical Analysis

Observational equivalence of narratives In the following sections, we consider narratives in which the extraneous variable z_t is the slope of the yield curve on US Treasuries. Insights from Proposition 1 guide the design of our empirical methodology: what differentiates the extraneous narrative from the baseline narrative is whether an inverted yield curve is linked to future incomes. Within this class of narratives, there is no need to further identify the direction of causality within each narrative, as different cases of the extraneous narrative ("signal", "channel", and "shock") all imply the same expectations.

Connection between expectations and sentiment Matching these narratives in media coverage to data from Twitter, we test whether exposure to such an extraneous narrative implies a differential response of sentiment. Due to data availability, our empirical analysis focuses on sentiment rather than expectations. However, this measure of sentiment has tight connection to expectations, both in our theoretical framework and in the broader literature.

To make this connection from the expectations in equation (11) to sentiment, we notice that all expectations of interest can be expressed as scalings of a single variable: the household's expectations of one-period ahead income.

***Proposition* 5** (common factor of expectations)**.** *Household expectations under narrative beliefs can be written:*

$$\mathbb{E}_t^k \, x_{t+s} = \Gamma_{k,s} \, \mathbb{E}_t^k \, y_{t+1} \tag{22}$$

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

Proof. Appendix A. □

A single common factor, $\mathbb{E}_t^k \, y_{t+1}$, is therefore sufficient to summarize all expectations, at any horizon. This single-factor structure 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. The 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.

This notation of sentiment corresponds to several existing studies. Kamdar (2019) finds that the bulk of the variation in a household's expectations as measured in the Michigan Survey of Consumers can be explained by a single factor. Like us, she labels this overall driver of expectations as sentiment. This is consistent with the Michigan Survey's own sentiment index, which similarly constructs an average level of optimism or pessimism from many expectations (see Lagerborg, Pappa and Ravn (2022) for a recent application). In related work, Andre et al. (2022a) use hypothetical vignettes to study the response of expectations to various shocks, and find that those expectation responses are driven by whether the shock is viewed as "good" or "bad", echoing our single-factor structure. Other theoretical models

in which a single sentiment-like factor drives expectations can replicate a range of features of macroeconomic and financial data (Molavi, 2019; Molavi, Tahbaz-Salehi and Vedolin, 2021).

Importantly, this differs from the definition of sentiments in e.g. Angeletos and La'O (2013) or Acharya et al. (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.

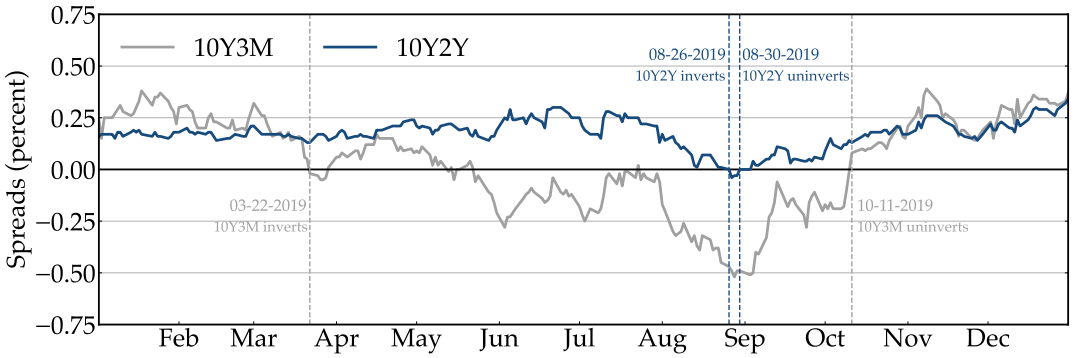
3. Yield Curve Inversion

The theoretical framework suggests that narratives may have an important effect on household sentiment, but measuring the effect is challenging because narratives are often ingrained with the underlying economic events. In this section, we focus on an episode of yield curve inversion in 2019 to isolate the effects of narratives, providing suggestive evidence on its importance.[6]

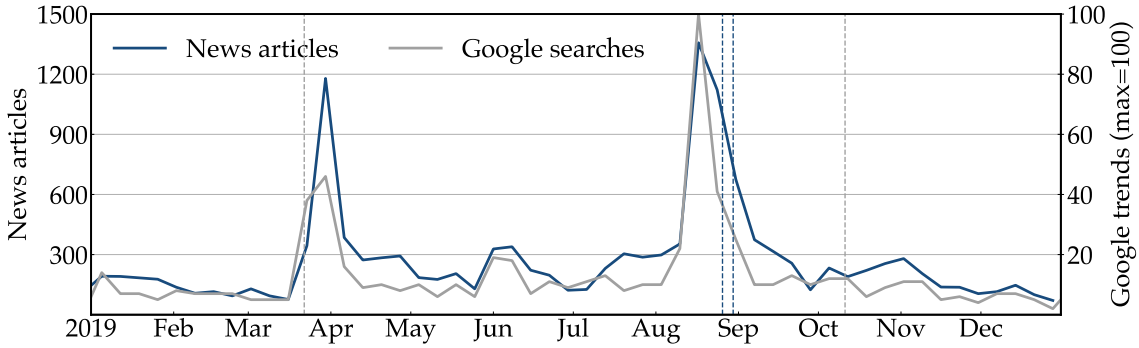
Yield curve inversions have been a closely-watched indicator of upcoming recessions in the U.S. since Harvey (1988) documented their predictive power for major recessions from the 1960s to the 1980s. Figure B.1 in the Appendix shows that the spread between the 10-year and 2-year Treasury bond yields has turned negative within 12 months before every recession in the US for the past 40 years. Despite the good track record of predicting recessions, it has also given false-positive signals (for example in 1966). The spread between long-term and short-term treasury yields is influenced by 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, Mertens et al., 2018). The implication of the yield curve inversion—specifically whether it predicts an imminent recession—is therefore not driven by the inversion itself, but rather the interpretations of it, providing a clear example in which heterogeneous competing narratives circulate simultaneously about the same economic event.

[6]While the yield curve inversion provides a laboratory to observe the effects of narratives, competing narratives are prevalent in the coverage of all economic news. Macaulay and Song (2022), for example, document competing narratives around the current elevated levels of inflation and find that shifting media narratives toward one that emphasize inflation's link to the real economy may have contributed to declining consumer sentiment in 2021.

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.

This episode of yield curve inversion, therefore, provides an ideal laboratory to isolate the effects of narratives.

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 widly-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[7] and Google searches for the term "yield curve" spiked before and during the inversions of both the 10Y2Y term spread and the 10-year-over-3-month (10Y3M) term spread, with a peak of interest right before the inversion

[7]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.

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[8]:

> 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 1: both of these narratives imply readers should update their expectations towards believing a recession is likely. It also underlines the importance of Proposition 1 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[9]:

> 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 on the yield curve inversion instead present a mix

[8] "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.

[9] "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.

between the two narratives. For example, Brian Chappatta's Bloomberg article[10] explains the nature of the yield curve and the historical significance of its inversion:

> What's a yield curve? [...] What are flat and inverted yield curves? [...] Why does it matter?

This defines an inverted yield curve, explains its history of proceeding recessions, but does not draw strong conclusions of what the inversion implies for the current economy.

We next study whether these narratives influence the outlook of their readers and estimate their respective importance.

4. Data

4.1. Newspaper articles

We capture narratives as media's different interpretations of the yield curve inversion (see Section 5.1 for details of our measurement approach). 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.[11] The 10 news outlets included in our sample in listed in Table 1.

During the event window of August 19 to September 13, 2019 (one week before the inversion and two weeks after the un-inversion, respectively)[12], we search for tweets by news 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.

[10] "The Yield Curve Is Inverted! Remind Me Why I Care" by Brian Chappatta. Link to the article on Bloomberg.

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

[12] 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.

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 5.1.

4.2. Twitter

Our Twitter data consists of three 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 collect base tweets using Twitter's Enterprise Search API, which contains the full archive of tweets since the start of Twitter in 2006.

Second, we use the rich network data available from Twitter to measure a user's exposure to narratives. Twitter provides four ways of interacting with posted tweets: quote retweet, retweet, reply and like. A "retweet" is when a user forwards a tweet without adding any comments, while a "quote retweet" requires that a user writes additional text when retweeting. The additional commentaries added by quote retweeters makes it more plausible that the users have digested the new information contained in the articles. Therefore, we use quote retweets as the main measure of exposure to narratives. Through Twitter's Standard API, we have information on the first 100 users who have quote retweeted each base tweet and the time of the quote retweet (timestamped to the second). 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, far below the API constraint of 100 users.

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) Retweeting users

	Mean	SD	5th Pctl	Median	95th Pctl	Obs
# tweets	3,863	14,948	6	637	15,368	404
# outlets	3.5	2.5	1	3	8	404

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.

Third, we measure changes in Twitter users' 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, again using the Enterprise API. Table 2b reports descriptive statistics of tweeting activity for the users in our sample, which shows that the median user is active and posts around 10 tweets per day. 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).[13] The sentiment score lies between 0 and 1, which is a uniform scale increasing with sentiment. 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, which demonstrates that the trained naïve Bayes classifier provides an accurate measure of tweet sentiment.

[13] 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).

5. Narrative-Driven Fluctuations in Sentiment

5.1. 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 an "extraneous 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. The coverage by Bloomberg's Brian Chappata can be empirically interpreted as a mix between the two narratives.

We extract these economic narratives from news articles using latent Dirichlet allocation (LDA), as developed by Blei, Ng and Jordan (2003) for natural language processing. Appendix C provides details on the LDA model.[14] 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, it replies on specialized vocabulary that are unique to each topic (for example, "risk" and "recession" versus "rainbow" and "pots of gold"). 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 1 shows that for the subset of *perfect* DAGs (in which "all parents are married") the direction of the causality within a DAG does not affect how a narrative influences a consumer's expectations. The important difference between narratives for fluctuations is whether phrases such as "yield

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

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 1. This greatly simplifies the measurement challenge and allows us to capture narratives with simple and interpretable LDA models.

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.

5.2. Yield-curve-inversion narratives

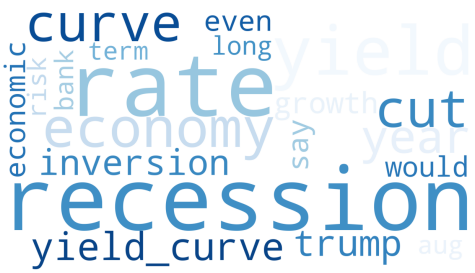
To estimate LDA outputs, we specify uniform Dirichlet priors, as in previous studies using LDA (e.g. Hansen et al., 2018).[15] The remaining parameter that we need to specify is the number of topic K. Our algorithm increments the number of topics from 2 until a topic emerges that does contain word "recession". 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 the highest probability of the word "recession" appearing as the *recession* narrative.

To ensure that our results are not sensitive to the labelling of the topics, in Appendix 5.1, 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. Appendix Table B.5 shows results that 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"

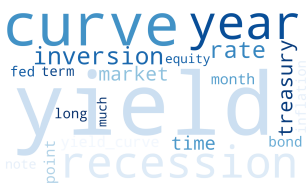
[15]The pre-processing of texts includes removing stop words and numbers, lemmatizing, and representing the documents with a bigram model.

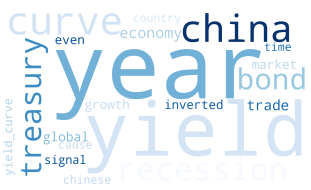
Figure 4: Economic narratives of the yield curve inversion: LDA outputs

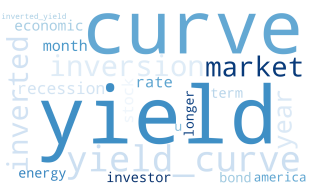


(a) "Recession" narrative

(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.

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 3 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 $\theta(\text{nonrecession}) = 0.05$. For the neutral coverage by Brian Chappata which introduces

the yield curve, the model produces more balanced loadings of $\theta(\text{recession}) = 0.67$ and $\theta(\text{nonrecession}) = 0.11$.

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 a narrative's salience in an article relative to other media coverage. 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.

5.3. Empirical importance of narratives

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 who has read news article d, the baseline model is:

$$\Delta s_{id} = \alpha + \beta_r \cdot \mathbb{1}(d, \text{recession}) + \beta_{nr} \cdot \mathbb{1}(d, \text{nonrecession}) + \varepsilon_{id}. \tag{23}$$

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. The exposure to a narrative is measured using quote retweeting activities on Twitter. We focus on the high-frequency changes in consumer sentiment 24 hours around the exposure to isolate the effect of the narrative. The timing is normalized so that the time when a consumer is exposed to a narrative is $t = 0$. Therefore, the time dimension of the baseline model in (23) is collapsed. The explanatory variables are narratives conveyed in an article. The binary variable $\mathbb{1}(d, k)$ measures whether the loading of an article d on narrative $k \in \{\text{recession}, \text{nonrecession}\}$ is above the cross-sectional mean. 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 consumer sentiment, respectively.

Table 3 contains our main results from estimating variants of (23). Column 1 reports our baseline estimates of β_{r} and β_{nr}, displayed in basis points. Exposure to the recession narrative

Table 3: Effects of narratives on consumer sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment					
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 (23). 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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

is associated with a significantly more pessimistic outlook. After a Twitter user is exposed to an article emphasizing the recession narrative, tweets posted by user display 1.3-basis-point more pessimistic sentiment. In contrast, the exposure to the nonrecession narrative leads to no significant changes in sentiment. This is not surprising, since the nonrecession narrative downplays the scenario of a potential recession and conveys that there is no change in economic fundamentals. These results are robust to different measures of narratives, as reported in Column 2.

Because newspaper subscription is not exogenously assigned, a substantial concern with our interpretation is that unobserved differences of Twitter users can drive both their sentiment changes and retweeting decisions. To ameliorate the concern, we show in Appendix

Figure B.2 that narratives do not have an effect on sentiment changes in the days before the exposure. 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 Reserver's open market operations do not control the exact timing of the yield curve inversion, which makes it plausible that the inversion is exogenous and unpredictable based on macroeconomic and financial information available prior to the event.

We conduct two additional robustness checks to verify the importance of narratives. First, we show the estimates are not sensitive to regression specifications. We use univariate models to estimate the effect of each narrative individually. Columns 3 and 4 in Table 3 confirm the baseline results that the exposure to the recession narrative leads to a more pessimistic outlook. Both the economic and statistical significance are similar to those from the baseline estimates. Columns 5 and 6 also confirm that the nonrecession narrative is not associated with significant changes in consumer sentiment.

Second, we show that results are not driven by business cycle fluctuations. Appendix Table B.3 considers potential confounding economic factors by controlling for market conditions and macroeconomic uncertainty, measured by the S&P 500 Index and the VIX Index respectively. Our estimates are little changed, which suggests that the impact on sentiment is not driven by current economic conditions or uncertainty, but rather by the media's interpretations of the yield curve inversion.

Focusing on susceptible consumers In Shiller (2017)'s epidemiological model of narratives, the economy consists of three types of agents: susceptibles, infectives, and recovereds. We now focus on measuring the effects of narratives on susceptible consumers, the households most likely to react to a new narrative. To do so, we limit our sample to users who retweet articles from a small number of news outlets only. The assumption here is that 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.

Appendix Table B.4 shows that, as in our main exercise, the recession narrative causes a decline in sentiment and the nonrecession narrative has no effect. However, the impact of

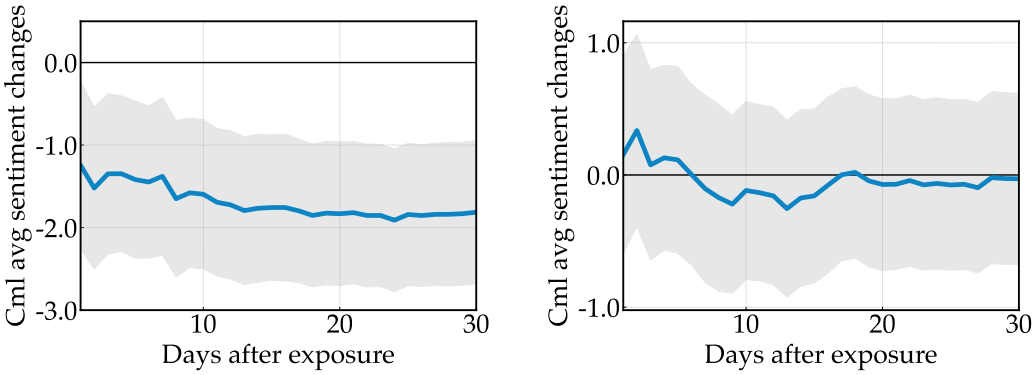
recession narrative is about 50% stronger on susceptible users than on the general population. We can alternatively interpret the results in Table B.4 as a robustness check, ensuring that the effects are not driven by users who selectively retweet many articles with a particular narrative to promote that agenda, rather than processing the information contained in a narrative.

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

$$\Delta_h s_{id} = \alpha + \beta_{kh} \cdot \mathbb{1}(d,k) + \varepsilon_{idh}, \tag{24}$$

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 5: 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 (24): $\Delta_h s_{id} = \alpha + \beta_{kh} \cdot \mathbb{1}(d,k) + \varepsilon_{idh}$ 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 $\mathbb{1}(d,k)$ denotes an indicator variable of whether the loading of an article d on a narrative k is above the cross-sectional mean. We estimate (24) separately for each horizon $h = 1, \cdots, 30$. Shaded areas represent 90% confidence intervals.

Figure 5 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.

Content of tweets In Appendix figure B.3, we study the source of pessimism of the recession narrative. We compare sentiment changes in tweets discussing economic conditions (containing keywords ∗economic∗ or ∗economy∗) and tweets discussing general, noneconomic conditions. While exposure to the recession narrative leads to a substantial decline in economic-specific sentiment, pessimism spreads to general sentiment and has a lasting effect. Consistent with findings by Kamdar (2019) and Molavi (2019) that household expectations appear to be driven by a single factor, recessionary narratives shape not only users' sentiment of the economic outlook but also their sentiment on everyday lives.

6. Unintended Consequences of QE

We now consider the impact of our empirical results in a quantitative model. As the results concern narrative-driven fluctuations in sentiment after a yield curve inversion, we study a model in which financial frictions generate a non-trivial yield curve, based on Gertler and Karadi (2013). Narratives differentiate whether financial frictions play a role in the transmission of monetary policy. The competition between baseline and extraneous narratives in this environment implies a novel trade-off in the effects of quantitative easing (QE). QE increases the frequency of narrative-driven waves of pessimism, but decreases their magnitude.

6.1. Model environment

We start with the set up in Gertler and Karadi (2013), featuring a New Keynesian model with frictional financial intermediation. In the interest of space, we highlight key equations and describe the remaining parts of the model verbally.[16] To relate the model to our empirical analysis, we allow QE to affect the slope of the yield curve. Depending on whether financial frictions exist, QE transmits differentially to the real economy, which gives rise to competing narratives regarding the role of yield curve inversion.

[16]For full details see Gertler and Karadi (2013).

Households are large families, consisting of a continuum of consumers and bankers. Consumers choose consumption, labor supply, and bank deposits to maximise expected discounted utility. The bank deposits flow to banks other than the ones controlled by their own household. Finally, consumers have a fixed probability of becoming a banker each period, and similarly the bankers have a fixed probability of transitioning back into being a consumer. Consumers provide a constant level of start-up equity for all new bankers in their household.

There is perfect consumption insurance for consumers within the household. This means that even when consumers have heterogeneous expectations due to heterogeneous narratives, this does not generate any heterogeneity in asset positions. To focus on the demand-side effects of narratives, we assume that consumers delegate the running of banks and firms to agents with rational expectations. The stochastic discount factors involved in their optimization problems are therefore formed using rational expectations.

Bankers intermediate funds from households to nonfinancial firms and the government. Their sources of funds are deposits from households and net worth, accumulated through retained earnings from returns on lendings. Using these funds, bankers provide debt financing to firms and the government subject to a financial constraint

$$V_t \geq \theta \cdot Q_t s_t + \Delta\theta \cdot q_t b_t, \tag{25}$$

where V_t denotes the bank value; $\{Q_t, s_t\}$ denote the price and quantity of banks' holdings of non-financial firm debt; and $\{q_t, b_t\}$ denote the price and quantity of banks' holdings of government debt. Parameters Δ and θ govern the severity of financial frictions in the economy. Banks can divert θ fraction of private debt and $\Delta\theta$ fraction of government debt. The financial constraint, when binding, introduces limits to arbitrage, which leads to a spread between the risk-free rate and the yield on private and government debt.

Firms form the New Keynesian block of the model. They produce using capital and labor. To fund investment, firms issue state-contingent securities to banks. A binding financial constraint in (25) restricts this debt issuance, which in turn restricts capital investment and output.

The central bank conducts QE through purchasing private loans or long-term government bonds at the prevailing market rates. When banks' financial constraint is binding, the

additional demand bids up prices of these assets and reduces financing costs of nonfinancial firms. To finance these purchases, the central bank issues riskless short-term debt.

We define slope of the yield curve as

$$\zeta_t = i_{b10,t} - i_t + \nu_{\zeta,t}, \tag{26}$$

where i_t denotes the yield on short-term riskless government bond; $i_{b10,t}$ denotes the nominal yield on the 10-year government bond;[17] and $\nu_{\zeta,t} \sim N(0, \sigma_\zeta^2)$ is a news shock which affects the long-term yield and therefore the slope of the yield curve.

6.2. Expectations

To stay as close as possible to standard models, we assume that all agents except for consumers have rational expectations. Consumers, however, form expectations using versions of the baseline and extraneous narratives explored in Section 2.

In each period, some consumers use a baseline narrative which does not relate the slope of the yield curve to any real variables. Others use an extraneous narrative where such links are included. Expectations formed using each narrative will in general differ from one another. However, as workers all belong to large families with perfect consumption insurance, this heterogeneity does not create heterogeneity in wealth across consumers.

Baseline Narrative When the financial constraint in (25) is not binding, the slope of the yield curve is irrelevant for all real variables. We therefore set up the baseline narrative such that expectations formed using this narrative coincide with rational expectations when the credit constraint does not bind.

It should be noted that the true model, whether the credit constraint binds or not, cannot be expressed as a DAG. With the general equilibrium, there are many pairs of variables (such as wages and labor supply) between which causal links run in both directions. The

[17] As Gertler and Karadi (2013), we consider the long-term government debt which pays unity for the first 10 years (40 periods) and a principal payment of $q_{ss}^n = 1/(R^n - 1)$ subsequently, where R^n denotes the steady-state nominal interest rate. The yield on the 10-year government bond, $i_{b10,t}$ is thus given by

$$P_t q_t = \sum_{\tau=1}^{40} \frac{1}{(1 + i_{b10,t})^\tau} + \frac{q_{ss}^n}{(1 + i_{b10,t})^{40}}. \tag{27}$$

true model is not therefore acyclic, as required in our definition of narratives. Even though there is no "structural" narrative to represent the true model, we can still specify a "reduced-form" narrative such that, when the credit constraint is not binding, consumers using that narrative form a correct perception of the joint distribution of all relevant variables as in Section 2.2.

Under log linearization, the true model without binding financial constraint can be represented in the state-space form

$$\boldsymbol{x}_t = \mathcal{H}\boldsymbol{x}_{t-1} + \mathcal{B}\boldsymbol{\varepsilon}_t, \tag{28}$$

where $\boldsymbol{x}_t$ is a vector of endogenous variables, $\boldsymbol{\varepsilon}_t$ is a vector of exogenous shocks, and $\mathcal{H}, \mathcal{B}$ are transition matrices. This gives rise to the rational expectations of y_{t+1} and r_{t+1} as

$$\mathbb{E}_t^* \, y_{t+1} = \mathcal{H}_y \boldsymbol{x}_t, \quad \mathbb{E}_t^* \, r_{t+1} = \mathcal{H}_r \boldsymbol{x}_t, \tag{29}$$

where $\mathcal{H}_y$ and $\mathcal{H}_r$ are the rows of $\mathcal{H}$ corresponding to y_t and r_t.

The baseline narrative induces expectations of the form

$$\mathbb{E}_t^b \, y_{t+1} = \mathcal{H}_y^b \boldsymbol{x}_t, \quad \mathbb{E}_t^b \, r_{t+1} = \mathcal{H}_r^b \boldsymbol{x}_t. \tag{30}$$

Narratives impose exclusion restrictions on $\mathcal{H}_y^b$ and $\mathcal{H}_y^b$ where there are perceived to be no causal relationships. We choose a baseline narrative such that $\mathcal{H}^b$ shares the same pattern of zero- and non-zero elements as its rational-expectation equivalent $\mathcal{H}$. Although the causal model behind $\mathcal{H}^b$ does not accurately capture all relationships within the economy, the resulting "reduced-form" law of motions for expectations is correctly specified. Fitting to many periods of data, expectations formed with this narrative will therefore coincide with rational expectations. The baseline narrative represents rational expectations if and only if the credit constraint (25) is not binding. Since this narrative excludes any financial variables, it will no longer coincide with rational expectations when those variables also affect real variables, i.e. when the credit constraint is binding.

Extraneous Narrative As in Section 2, we assume that the extraneous narrative is identical to the baseline narrative, except that it includes a perceived causal link from the sign of the spread between the interest rates on 10 year and short term debt. That is, consumers with the extraneous narrative form expectations of output and interest rate based on

$$\mathbb{E}_t^e\, y_{t+1} = \mathcal{H}_y^e \boldsymbol{x}_t + h_y^e \mathbb{1}_{\zeta_t<0}, \quad \mathbb{E}_t^e\, r_{t+1} = \mathcal{H}_r^e \boldsymbol{x}_t + h_r^e \mathbb{1}_{\zeta_t<0} \tag{31}$$

where h_y^e, h_r^e are parameters and $\mathbb{1}_{\zeta_t<0}$ is an indicator variable for periods when the yield curve is inverted.

There are two effects of including this extra variable in the narrative. First, when the financial constraint binds, the coefficients on other variables in the narrative may change. In this case, the baseline narrative is mis-specified, so when fitting the extraneous narrative to a long history of data, including a financial variable will affect the estimated parameters on all variables.[18] Second, a yield curve inversion event will cause a discrete change in the expectations of consumers holding this narrative, as in our empirical results.

6.3. The Distribution of Narratives

In Section 2, the fraction of agents using the baseline narrative λ_t was assumed to be fixed. However, it is plausible that the prevalence of different narratives depends on aspects of the economic environment (see, for example, Larsen and Thorsrud, 2019; Eliaz and Spiegler, 2020; Macaulay and Song, 2022). In our context, many of the articles in our news media sample with strong loadings on the "nonrecession" narrative explicitly cite QE as a force leading to a flatter yield curve. For example, Emily Barrett and Katherine Greifeld of Bloomberg write in "Treasuries Buying Wave Triggers First Curve Inversion since 2007":[19]

> That said, many downplay the curve's predictive powers. Some argue that technical factors have distorted the curve's shape and signaling capacity, particularly as crisis-era policy has tethered yields for the past decade.

This suggests that central bank asset purchases make the baseline narrative more popular by flattening the yield curve. Without QE, the yield curve only inverts if there is a large

[18] When the constraint never binds, the baseline narrative is correctly specified, and so when fit to many periods of data the extraneous narrative converges to the baseline narrative, as in Proposition 3.

[19] Link to the article on Bloomberg.

shock, which implies consumers ought to pay attention to such an event. However, because of large QE and the resulting flatter yield curve, inversions may be triggered by much smaller shocks, making it easier for a "nonrecession" narrative to prevail. This is related to the idea in Bordalo, Gennaioli, Ma and Shleifer (2020) that replacements for rational expectations should still contain a "kernel of truth", as they evolve to reflect changes in the equilibrium processes of the model.[20]

We incorporate this channel in the model by specifying a simple reduced-form process for the fraction of consumers who subscribe to the baseline narrative

$$\lambda_t = \lambda + \gamma \cdot \zeta_t, \tag{32}$$

where the parameter λ denotes a time-invariant proportion of consumers using the baseline narrative, and ζ_t is the slope of the yield curve. The parameter γ controls the relationship between the yield curve and the distribution of narratives. We assume that $\gamma < 0$, so that as greater asset purchases by the central bank flatten the yield curve, more consumers start to believe the baseline narrative.

6.4. Calibration

Table 4: Calibration

Parameter	Description	Value
h_y^e	effect of extraneous narrative	-1.3×10^{-4}
h_r^e	effect of extraneous narrative	0
λ	narrative process	0.94
γ	narrative process	-0.2
σ_ζ	news shock process	0.15

Our calibration consists of two parts. For standard parameters and monetary-policy parameters, we follow the calibration by Gertler and Karadi (2013). Table 4 contains parameters specific to our model. We calibrate h_y^e, the discrete shifts in expectations in response to yield curve inversion by households with extraneous narratives, to match our empirical

[20] While Eliaz and Spiegler (2020) also have a notion of empirical consistency in their model of competing narratives, they differ from us in assuming agents select "hopeful" narratives promising the highest anticipatory utility.

estimates from Table 3 Column 1. To focus on the effects of narratives on output sentiment, we set $h_y^e = 0$ and shut down the effects of narratives on interest-rate expectations. We calibrate parameters for the process of the baseline narrative in (32) so that the largest QE injection considered in the quantitative analysis implies that all consumers use the baseline narrative ($\lambda_t = 1$). Finally, we calibrate the volatility for news shocks using the estimate from Khan and Tsoukalas (2012).

6.5. Frequency vs. Magnitude of Narrative-Driven Fluctuations in Sentiment

The role of narratives in consumer expectation formation implies that yield curve inversions cause a decline in output. This occurs because consumers using the extraneous narrative lower their expectations of future income, consistent with the declines in sentiment observed empirically in Section 5. An expansion in QE has two effects on these narrative-driven fluctuations, both arising from a flatter yield curve.

Figure 6 illustrates these two effects. Panel (a) shows that as the yield curve becomes flatter with larger doses of QE, the probability of inversion in a given period rises. This implies that as inversion becomes more frequent, so does narrative-driven declines in sentiment. Panel (b) shows a counteracting effect. Greater QE increases the prevalence of the baseline narrative, implying that fewer consumers subscribe to the extraneous narrative that incorporates yield curve inversions. A given inversion episode, therefore, has a smaller impact on average sentiment, and in turn on output.

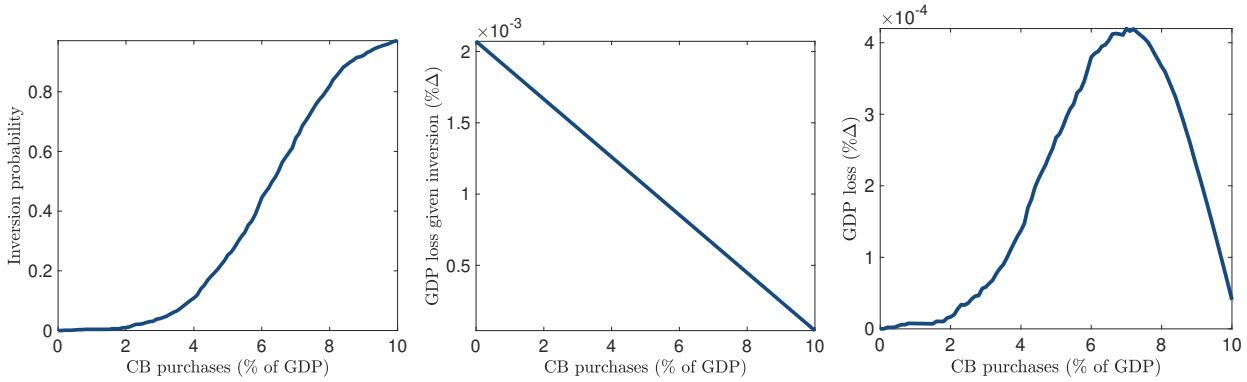
Larger doses of QE, therefore, increase the frequency, but decrease the magnitude, of narrative-driven fluctuations. To illustrate the novel trade-off, we decompose the expected changes in output y_t in response to an arbitrary shock ε_t using the law of iterated expectations. The first part is a change in output if the shock causes the yield curve to invert and the second part is an equivalent change if there is no inversion. These two effects differ because inversions affect the expectations of all households using the extraneous narrative.

$$\mathbb{E}(y_t|\varepsilon_t) = \mathbb{E}(y_t|\varepsilon_t, \mathbb{1}_{\zeta_t<0}) \Pr(\mathbb{1}_{\zeta_t<0}|\varepsilon_t) + \mathbb{E}(y_t|\varepsilon_t, \mathbb{1}_{\zeta_t\geq 0}) \Pr(\mathbb{1}_{\zeta_t\geq 0}|\varepsilon_t) \tag{33}$$

$$= \underbrace{\mathbb{E}(y_t|\varepsilon_t, \mathbb{1}_{\zeta_t\geq 0})}_{\text{no inversion}} + \underbrace{\Pr(\mathbb{1}_{\zeta_t<0}|\varepsilon_t)}_{\text{frequency}} \Big[\underbrace{\mathbb{E}(y_t|\varepsilon_t, \mathbb{1}_{\zeta_t<0}) - \mathbb{E}(y_t|\varepsilon_t, \mathbb{1}_{\zeta_t\geq 0})}_{\text{magnitude}} \Big] \tag{34}$$

The effects of a given shock in expectation are therefore given by the effect without

Figure 6: Effects of quantitative easing on narrative-driven fluctuations



(a) Frequency **(b)** Magnitude **(c)** Total effects

Notes: This figure reports the responses to various levels of central bank (CB) purchases. Panel (a) reports the probability of yield curve inversion based on 1000 simulations of $\nu_{\zeta,t}$ in (26). Panel (b) reports the loss of GDP conditional on yield curve inversion, $\mathbb{E}(y_t|\nu_{\zeta,t}, \mathbb{1}_{\zeta_t<0}) - \mathbb{E}(y_t|\nu_{\zeta,t}, \mathbb{1}_{\zeta_t\geq 0})$, in percent deviation from the steady state. Panel (c) reports the unconditional loss of GDP.

a yield curve inversion, and then the product of two additional terms. The first gives the probability that a given shock leads to a yield curve inversion. The second is the difference between output responses with and without such an inversion, which reflects the strength of narrative-driven sentiment fluctuations induced by an inversion event. QE flattens the yield curve, which leads to an increase in the frequency of inversion, but a decline in the magnitude of its effect. The overall effects of QE on the expected output sensitivity to such shocks are plotted in Figure 6c.

At very low levels of QE, the yield curve is steep, so inversion events are rare. In expectation, narrative-driven waves of pessimism are therefore small. At very high levels of QE, the majority of consumers believe the baseline narrative, and so yield curve inversions do not affect sentiment. At intermediate levels of QE, however, yield curve inversions happen with non-negligible probability and affect the expectations of a substantial fraction of consumers, causing output losses.

7. Conclusion

Narratives are increasingly seen as an important factor in how economic agents form their expectations, by both scholars (Shiller, 2017, 2020) and policymakers (Schnabel, 2020).[21] We provide evidence that exposure to particular narratives in the media does indeed have significant effects on sentiment.

Formalizing narratives as directed acyclic graphs, we show that certain groups of narratives will in fact have exactly the same effect on 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 link connecting the inverted yield curve with an upcoming recession.

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).

We confirm this aggregate implication in a quantitative model informed by our empirical results. Yield curve inversions cause declines in expected incomes among households holding narratives in which such events are linked to recessions. This implies that extended periods of quantitative easing generate two novel offsetting effects: by flattening the yield curve, they make such narrative-driven fluctuations in sentiment more frequent. However, they also reduce the prevalence of the "recession" narrative, reducing the magnitude of those fluctuations.

Our approach using tools from natural language processing to extract relevant groups

[21] See, for example, the speech by Isabel Schnabel, Member of the Executive Board of the ECB, at the Karlsruhe Law Studies Society entitled "Narratives about the ECB's Monetary Policy – Reality or Fiction?" (Schnabel, 2020).

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 and by firm managers in earnings reports. 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.

Andrade, Philippe, Gaetano Gaballo, Eric Mengus, and Benoît Mojon, "Forward Guidance and Heterogeneous Beliefs," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 1–29.

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

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

Atmaz, Adem and Suleyman Basak, "Belief Dispersion in the Stock Market," *Journal of Finance*, 2018, *73* (3), 1225–1279.

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.

_ , Carolin Pflueger, and Adi Sunderam, "Perceptions about Monetary Policy," Technical Report, National Bureau of Economic Research 2022.

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

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.

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.

_ , _ , Yueran Ma, and Andrei Shleifer, "Overreaction in Macroeconomic Expectations," *American Economic Review*, 2020, *110* (9), 2748–2782.

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

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.

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.

Dräger, Lena, Michael J. Lamla, and Damjan Pfajfar, "Are Survey Expectations Theory-Consistent? The Role of Central Bank Communication and News," *European Economic Review*, 2016, *85*, 84–111.

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.

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.

Gentzkow, Matthew and Jesse M Shapiro, "Competition and Truth in the Market for News," *Journal of Economic perspectives*, 2008, *22* (2), 133–154.

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

Gertler, Mark and Peter Karadi, "QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool," *International Journal of Central Banking*, 2013, *29.*

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.

Greenwood, Robin and Andrei Shleifer, "Expectations of Returns and Expected Returns," *Review of Financial Studies*, 2014, *27* (3), 714–746.

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.

Harris, Milton and Artur Raviv, "Differences of Opinion Make a Horse Race," *Review of Financial Studies*, 1993, *6* (3), 473–506.

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

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

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," Technical Report, National Bureau of Economic Research 2022.

Khan, Hashmat and John Tsoukalas, "The quantitative importance of news shocks in estimated DSGE models," *Journal of Money, Credit and Banking*, 2012, *44* (8), 1535–1561.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran, "The tail that wags the economy: Beliefs and persistent stagnation," *Journal of Political Economy*, 2020, *128* (8), 2839–2879.

Krishnamurthy, Arvind and Wenhao Li, "Dissecting Mechanisms of Financial Crises: Intermediation and Sentiment," Technical Report, National Bureau of Economic Research 2020.

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, "Shock Transmission and the Sources of Heterogeneous Expectations," Technical Report, University of Oxford, UK 2022.

_ **and Wenting Song**, "News Media, Inflation, and Sentiment," *Manuscript*, 2022.

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.

Maxted, Peter, "A Macro-Finance Model with Sentiment," *Manuscript*, 2019.

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.

Patton, Andrew J. and Allan Timmermann, "Why do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion," *Journal of Monetary Economics*, 2010, *57* (7), 803–820.

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.

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.

_ , "Can Agents with Causal Misperceptions be Systematically Fooled?," *Journal of the European Economic Association*, 2020, *18*, 583–617.

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

Xiong, Wei and Hongjun Yan, "Heterogeneous Expectations and Bond Markets," *Review of Financial Studies*, 2010, *23* (4), 1433–1466.

Appendices

A. Proofs for Section 2

Proposition 1. We begin by showing $p_c(\cdot) = p_b(\cdot)$. 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.

Proposition 2. Since z_s is correlated with y_s and/or y_{s+1}, they are not conditionally independent. As a result, $p(z_s) \neq p(z_s|y_s, y_{s+1})$, which implies $\tilde{p}(r_s, r_{s+1}, y_s, y_{s+1}, z_s) \neq \tilde{p}_a(r_s, r_{s+1}, y_s, y_{s+1}, z_s)$.

Lemma 1. 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 Definitions 2 and 3 can be written as if they reflect linear

perceived laws of motion for each variable

$$y_t = A^k y_{t-1} + B^k r_{t-1} + Z^k z_{t-1} + v_t^y \tag{35}$$

$$r_t = C^k y_t + v_t^r \tag{36}$$

$$z_t = D^k y_t + v_t^z \tag{37}$$

for $k \in \{b, e\}$, where $D^b, Z^b = 0$ by assumption, and v_t^y, v_t^r, v_t^z are all mean-zero shocks. Substituting out for y_t in the perceived laws of motion for r_t, z_t and stacking the resulting equations gives

$$x_t = \begin{pmatrix} A^k & B^k & Z^k \\ A^k C^k & B^k C^k & Z^k C^k \\ A^k D^k & B^k D^k & Z^k D^k \end{pmatrix} x_{t-1} + v_t \tag{38}$$

where v_t is a 3×1 vector of shocks, each element of which is a linear (mean-zero) combination of v_t^y, v_t^r, v_t^z. Since v_t has zero mean, taking expectations of this implies equation (15).

Equations for finding equilibrium and narrative coefficients. First, we find equilibrium under fixed narrative coefficients $\mathcal{H}_k$, then solve for those coefficients in narrative equilibrium. In this, it will be convenient to work with the parameters $A^k - Z^k$ defined in the proof of Lemma 1, rather than the combinations of these that form the coefficients of $\mathcal{H}_k$. It is also useful to note that

***Lemma* 2** (rewriting expectations)**.** *With the linear narratives defined in Lemma 1, expectations are given by*

$$\mathbb{E}_t^k y_{t+s} = (A^k + B^k C^k + Z^k D^k)^{s-1} \mathbb{E}_t^k y_{t+1} \tag{39}$$

$$\mathbb{E}_t^k r_{t+s} = C^k (A^k + B^k C^k + Z^k D^k)^{s-1} \mathbb{E}_t^k y_{t+1} \tag{40}$$

$$\mathbb{E}_t^k z_{t+s} = D^k (A^k + B^k C^k + Z^k D^k)^{s-1} \mathbb{E}_t^k y_{t+1} \tag{41}$$

Proof. From equations (35)-(37), we have

$$\begin{aligned}
\mathbb{E}_t^k\, y_{t+s} &= A^k\, \mathbb{E}_t^k\, y_{t+s-1} + B^k\, \mathbb{E}_t^k\, r_{t+s-1} + Z^k\, \mathbb{E}_t^k\, z_{t+s-1} \\
&= (A^k + B^k C^k + Z^k D^k)\, \mathbb{E}_t^k\, y_{t+s-1} \\
&= (A^k + B^k C^k + Z^k D^k)^{s-1}\, \mathbb{E}_t^k\, y_{t+1} \qquad (42) \\
\mathbb{E}_t^k\, r_{t+s} &= C^k\, \mathbb{E}_t^k\, y_{t+s} \\
&= C^k (A^k + B^k C^k + Z^k D^k)^{s-1}\, \mathbb{E}_t^k\, y_{t+1} \qquad (43) \\
\mathbb{E}_t^k\, z_{t+s} &= D^k\, \mathbb{E}_t^k\, y_{t+s} \\
&= D^k (A^k + B^k C^k + Z^k D^k)^{s-1}\, \mathbb{E}_t^k\, y_{t+1} \qquad (44)
\end{aligned}$$

□

As these expectations are determined entirely by observed r_t, y_t, and (for the extraneous narrative) z_t, the consumption of each group of households is given by

$$c_t^b = \Theta_y^b y_t + \Theta_r^b r_t \qquad (45)$$

$$c_t^e = \Theta_y^e y_t + \Theta_r^e r_t + \Theta_z^e z_t \qquad (46)$$

where

$$\Theta_y^k = 1 - \beta + A^k \psi^k \qquad (47)$$

$$\Theta_r^k = -\beta\sigma + B^k \psi^k \qquad (48)$$

$$\Theta_z^e = Z^e \psi^k \qquad (49)$$

and

$$\psi^k = \frac{\beta(1 - \beta - \beta\sigma C^k)}{1 - \beta(A^k + B^k C^k + Z^k D^k)} \qquad (50)$$

is the elasticity of consumption to $\mathbb{E}_t^b\, y_{t+1}$.

Substituting (45) and (46) into (12) gives

$$y_t = (\rho + \gamma\lambda_{t-1}\Theta_y^b + \gamma(1-\lambda_{t-1})\Theta_y^e)y_{t-1} + (\gamma\lambda_{t-1}\Theta_r^b + \gamma(1-\lambda_{t-1})\Theta_r^e)r_{t-1} + (\mu + \gamma(1-\lambda_{t-1})\Theta_z^e)z_{t-1} + v_t^y \quad (51)$$

Finally, substituting (45) and (46) into (13) gives an expression for equilibrium r_t

$$r_t = \frac{1-\lambda_t\Theta_y^b - (1-\lambda_t)\Theta_y^e}{\lambda_t\Theta_r^b + (1-\lambda_t)\Theta_r^e}y_t - \frac{(1-\lambda_t)\Theta_z^e}{\lambda_t\Theta_r^b + (1-\lambda_t)\Theta_r^e}z_t \quad (52)$$

Equations (14), (51), and (52) therefore give the true equilibrium DGPs of z_t, y_t, r_t under a given narrative. We now find the narrative coefficients.

Begin with the extraneous narrative. The first row of the narrative equation (38) has the same functional form as the true DGP for output (51). Matching coefficients yields

$$A^e = \rho + \gamma\lambda_{t-1}\Theta_y^b + \gamma(1-\lambda_{t-1})\Theta_y^e \quad (53)$$

$$B^e = \gamma\lambda_{t-1}\Theta_r^b + \gamma(1-\lambda_{t-1})\Theta_r^e \quad (54)$$

$$Z^e = \mu + \gamma(1-\lambda_{t-1})\Theta_z^e \quad (55)$$

Similarly, from (14) and the final row of (38)

$$D^e = \chi \quad (56)$$

Finally, substitute equation (14) into equation (52) to obtain

$$r_t = \frac{1-\lambda_t\Theta_y^b - (1-\lambda_t)(\Theta_y^e + \chi\Theta_z^e)}{\lambda_t\Theta_r^b + (1-\lambda_t)\Theta_r^e}y_t - \frac{(1-\lambda_t)\Theta_z^e}{\lambda_t\Theta_r^b + (1-\lambda_t)\Theta_r^e}v_t^z \quad (57)$$

Matching coefficients with row two of (38) gives

$$C^e = \frac{1-\lambda_t\Theta_y^b - (1-\lambda_t)(\Theta_y^e + \chi\Theta_z^e)}{\lambda_t\Theta_r^b + (1-\lambda_t)\Theta_r^e} \quad (58)$$

There is no role for v_t^z in fitting this aspect of the narrative as v_t^z can only affect output in period $t+1$, and is therefore independent of y_t. All other elements of Ω^e equal zero.

We now move on to construct similar equations for the baseline narrative parameters.

Using equation (14) to substitute for z_{t-1} in equation (51) we obtain

$$y_t = (\rho + \chi\mu + \gamma\lambda_{t-1}\Theta_y^b + \gamma(1-\lambda_{t-1})(\Theta_y^e + \chi\Theta_z^e))y_{t-1} + (\gamma\lambda_{t-1}\Theta_r^b + \gamma(1-\lambda_{t-1})\Theta_r^e)r_{t-1} + (\mu + \gamma(1-\lambda_{t-1})\Theta_z^e)v_{t-1}^z + v_t^y \tag{59}$$

Matching coefficients with row one of (38)

$$A^b = \rho + \chi\mu + \gamma\lambda_{t-1}\Theta_y^b + \gamma(1-\lambda_{t-1})(\Theta_y^e + \chi\Theta_z^e) \tag{60}$$

$$B^b = \gamma\lambda_{t-1}\Theta_r^b + \gamma(1-\lambda_{t-1})\Theta_r^e \tag{61}$$

where again we use that v_{t-1}^z is independent of r_{t-1}, y_{t-1}, as it does not affect those variables until the following period. Finally, as the baseline narrative gives the same causal underpinnings to r_t as the extraneous narrative, we also have:

$$C^b = \frac{1 - \lambda_t\Theta_y^b - (1-\lambda_t)(\Theta_y^e + \chi\Theta_z^e)}{\lambda_t\Theta_r^b + (1-\lambda_t)\Theta_r^e} \tag{62}$$

from matching coefficients between (57) and row two of (38).

Notice that $B^b = B^e$ and $C^b = C^e$. This is because both baseline and extraneous narratives share the same causal links between r_t and all other variables. However, $A^b \neq A^e$, because a household fitting the baseline narrative assigns some of the variability in output due to z_{t-1} to variation in y_{t-1}. We therefore have a system of 6 equations ((53), (54), (55), (56), (58), (60)) in 6 unknowns ($A^e, B^e, C^e, D^e, Z^e, A^b$). These equations are nonlinear in the unknown parameters, because of the nonlinear combination terms $\Theta_y^b, \Theta_r^b, \Theta_y^e, \Theta_r^e, \Theta_z^e$. Propositions 3 and 4 follow from solving two special cases of this system.

Proposition 3. With $\mu = 0$, equation (56) has a unique solution at $Z^e = \Theta_z^e = 0$. Through equations (53) and (60), we then have that $A^b = A^e$. This reduces the system to four equations, which when solved and substituted into equation (38) give equations (16)

and (17), with Ω^*_{ij} defined as

$$\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \frac{1}{1-\beta\gamma} \begin{pmatrix} \rho + \gamma(1-\beta) & -\beta\gamma\sigma \\ (\rho + \gamma(1-\beta))(\gamma+\rho-1)\sigma^{-1} & -\beta\gamma(\gamma+\rho-1) \end{pmatrix} \tag{63}$$

Furthermore,

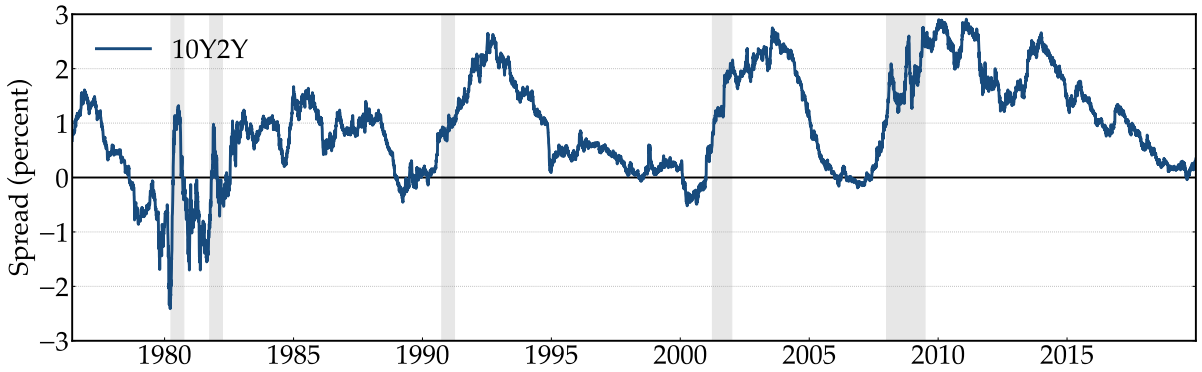
$$h_{31} = \frac{\chi(\rho + \gamma(1-\beta))}{1-\beta\gamma}, \text{ and } h_{32} = -\frac{\chi\beta\gamma\sigma}{1-\beta\gamma}. \tag{64}$$

Proposition 4. With $\chi = 0$, equations (53) and (60) coincide, so $A^e = A^b$. Furthermore, through equation (56), $D^e = 0$, which in turn implies that $\Omega^b_y = \Omega^e_y$ and $\Omega^b_r = \Omega^e_r$. Solving the reduced system of equations for A^e, B^e, C^e, Z^e and substituting into equation (38) gives equations (18) and (19). Furthermore, $h_{13} = \frac{\mu}{1-\beta\gamma(1-\lambda_{t-1})}$ and $h_{23} = \frac{\mu(\gamma+\rho-1)}{\sigma(1-\beta\gamma(1-\lambda_{t-1}))}$.

Proposition 5. Follows directly from stacking the equations in Lemma 2 into vector form.

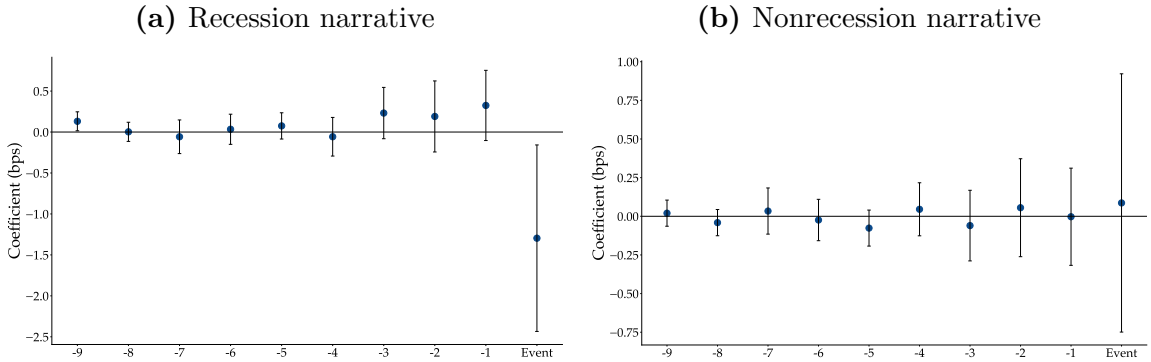
B. Additional Tables and Figures

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: 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.

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.

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: Controlling for macroeconomic conditions

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment					
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
Macro controls	yes	yes	yes	yes	yes	yes

Notes: This table reports results from estimating variants of the baseline specification in (23) 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.4: Limiting the number of outlets in user timelines

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment					
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 (23), 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

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

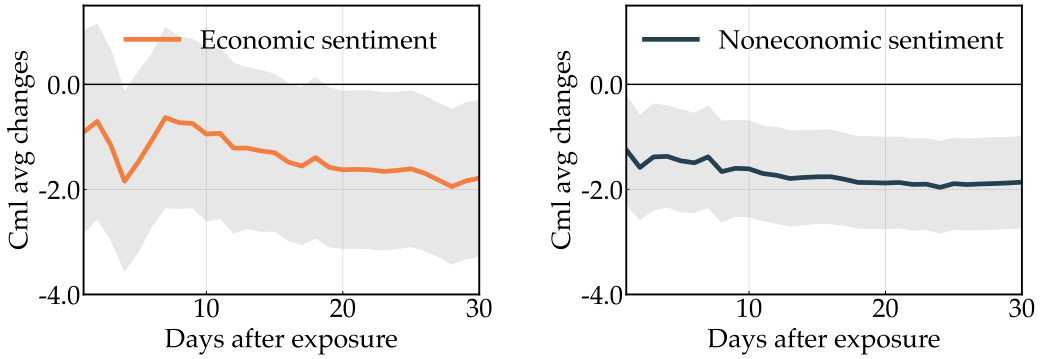
where Δs_{id} denotes changes in user i's tweet sentiment 24 hours around reading article d; and $\mathbb{1}(d,k)$ for $k \in \{\text{recession}, \text{nonrecession}\}$ denotes an indicator variable for whether the loading of article d on narrative k is above the cross-sectional mean. Tweet sentiment 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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table B.5: Automated topic labelling with guided LDA

	(1)	(2)
	Tweet Sentiment	
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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Figure B.3: 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 (24) separately for each horizon $h = 1, \cdots, 30$. Shaded areas represent 90% confidence intervals.

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

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \theta_1^{\alpha_1 - 1} \cdots \theta_k^{\alpha_k - 1} \tag{65}$$

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

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^{N} p(z_n|\theta) p(w_n|z_n, \beta). \tag{66}$$

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

$$p(w|\alpha,\beta) = \int p(\theta|\alpha)\left(\prod_{n=1}^{N}\sum_{z_n} p(z_n|\theta)p(w_n|z_n,\beta)\right), \tag{67}$$

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

$$p(D|\alpha,\beta) = \prod_{d=1}^{M}\int p(\theta_d|\alpha)\left(\prod_{n=1}^{N_d}\sum_{z_{dn}} p(z_{dn}|\theta_d)p(w_{dn}|z_{dn},\beta)\right) \tag{68}$$

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

$$p(\theta,z|w,\alpha,\beta) = \frac{p(\theta,z,w|\alpha,\beta)}{p(w|\alpha,\beta)} \tag{69}$$

where

$$p(w|\alpha,\beta) = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \gamma(\alpha_i)}\int\left(\prod_{i=1}^{k}\theta_i^{\alpha_i-1}\right)\left(\prod_{n=1}^{N}\prod_{i=1}^{k}\prod_{j=1}^{V}(\theta_i\beta_{ij})_{w_n^j}\right)d\theta, \tag{70}$$

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:

$$p(y|(t_1, \cdots, t_n) \propto p(y) \prod_{i=1}^{n} p(t_i|y) \tag{71}$$

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:

$$\text{tf-idf}_{t,d} = \underbrace{\frac{w_{t,d}}{\sum_{\tau \in d} w_{\tau,d}}}_{\text{term frequency}} \cdot \underbrace{\log \frac{D}{|\{d \in D : t \in d\}|}}_{\text{inverse document frequency}} \tag{72}$$

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:

$$\tilde{p}_j(\text{positive}) = f(t_1, \cdots, t_n) \tag{73}$$

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:

$$s_{it} = \frac{1}{J} \sum_j \tilde{p}_j(\text{positive}) \quad \text{for } j \text{ posted in day } t \tag{74}$$

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.