

The Macroeconomics of Narratives

Joel P. Flynn*

MIT

Karthik A. Sastry†

MIT

June 19, 2022

Abstract

We study the macroeconomic implications of viral, belief-altering narratives. Empirically, we use natural-language-processing methods to extract narratives from the text of all U.S. public firms' end-of-year business and financial reports (Forms 10-K). We find that: (i) firms' hiring decisions respond strongly to discussion of narratives and (ii) firms' discussion of narratives is both viral and responsive to macroeconomic conditions. To understand the implications of narratives for macroeconomic fluctuations, we embed a viral optimistic narrative in a Neoclassical business-cycle model. We show that viral optimism can lead the economy's essentially unique equilibrium to feature: (i) multiple, self-fulfilling steady states, (ii) discontinuous and hump-shaped impulse responses, and (iii) endogenous regime switching between times of high and low optimism. When calibrated to match our empirical results, our model implies that declines in aggregate optimism account for approximately 32% of the output reduction during the Early 2000s Recession and 18% during the Great Recession.

*MIT Department of Economics, 50 Memorial Drive, Cambridge, MA 02142. Email: jpflynn@mit.edu.

†MIT Department of Economics, 50 Memorial Drive, Cambridge, MA 02142. Email: ksastry@mit.edu.

We are grateful to Daron Acemoglu, George-Marios Angeletos, Ricardo Caballero, Oğuzhan Çelebi, Hui Chen, Roberto Corrao, Luca Gagliardone, Adam Guren, Basil Halperin, Tarek Hassan, Leonid Kogan, Antoine Levy, Stephen Morris, Georgios Nikolakoudis, Jonathan Parker, Christina Patterson, Charlie Rafkin, Antoinette Schoar, David Thesmar, Adrien Verdelhan, Emil Verner, Jiang Wang, Christian Wolf, and participants in the MIT Macro Lunch, MIT Theory Lunch, and MIT Sloan Finance Lunch for helpful comments.

1 Introduction

In *The General Theory of Employment, Interest, and Money*, Keynes (1936) argues that *animal spirits* lead economic decisions to differ from what is justified purely by prospects’ true expected values. But what drives animal spirits? A prominent candidate is provided by the *Narrative Economics* of Shiller (2017, 2020), wherein viral stories and worldviews are hypothesized to be a primary driver of economic fluctuations. However, the importance of narratives for business cycles, and *vice versa*, remains unclear.

In this paper, we study joint fluctuations in narratives and the macroeconomy. We develop a framework in which narratives describe subjective models of the macroeconomy and form building blocks of macroeconomic beliefs. Based upon the narratives in which they believe, agents take actions under both fundamental and strategic uncertainty. The spread of narratives depends both on their prevalence in the population and how well they describe economic reality. Thus, narratives in our conceptualization are *relevant for decisions*, like Keynes’ animal spirits, and *viral*, like Shiller’s narratives.¹

We operationalize this framework to measure narratives and estimate their importance in the data. We do this by applying natural-language-processing methods to the universe of 10-K regulatory filings in which all U.S. public firms discuss “perspectives on [their] business results and what is driving them” (SEC, 2011). We find that measured narratives matter strongly for firms’ hiring decisions and that their spread is both viral and responsive to aggregate conditions.

To understand the macroeconomic implications of these findings, we embed a viral optimistic narrative in a microfounded business-cycle model. We show that viral optimism can lead the economy to feature hysteresis and boom-bust cycles. We calibrate this model to match our empirical evidence and find that measured variation in aggregate optimism accounts for approximately 32% of the decline in output over Early 2000s Recession and 18% over the Great Recession.

Measuring Narratives. Our first goal is to empirically evaluate the two premises of our model of narrative macroeconomics: narratives’ *decision-relevance* and *virality*. To this end, we construct a dataset of narrative adoption among U.S. public firms, whose perspectives on economic events are publicly available in compulsory regulatory filings (forms 10-K) and whose input choice and economic performance are easily measured.

We apply three techniques to extract narratives from the corpus of 10-K documents. The first technique is to measure the intensity of *positive and negative sentiment* using the 10-K-specific dictionary introduced by Loughran and McDonald (2011). We interpret this

¹See Online Appendix F for a description of the relationship between our study and Shiller (2017, 2020).

measure as capturing optimism. The second technique is to measure the frequency of words that best characterize the nine *Perennial Economic Narratives* introduced by Shiller (2020). Our method uses simple tools from natural language processing, applied also by Hassan, Holander, Van Lent, and Tahoun (2019) and Flynn and Sastry (2021), to select words that are common in Shiller’s description of narratives but relatively uncommon among 10-Ks. These “narratively-identified narratives,” by construction, are motivated by the historical evidence of relevance and virality provided by Shiller (2020). The third technique is to estimate a Latent Dirichlet Allocation (LDA) model (Blei, Ng, and Jordan, 2003), which extracts an underlying set of *topics*, probability distributions over words, based on the frequency with which certain words co-occur within documents. This measure is the most flexible, as it incorporates no information *a priori* about the topics’ relevance or virality.

Empirical Results. We first provide descriptive evidence about our estimated narratives. Across the three methods, almost all of our estimated narratives are persistent and cyclical. Very little of the narratives’ *total* variance, however, is at the time-series level – only 1.1% for optimism, and 0.2% and 3.5%, respectively, for the median Shiller or topic narrative.

We next leverage the cross-sectional variation to examine the *relevance* of our measured narratives for economic decisions. Our first main result is that optimistic firms hire 3.6 percentage points more than pessimistic firms in a given year, net of firm and sector-time means. We show that this effect is robust to accounting for firm-level productivity and financial conditions. Moreover, we find that optimism is neither correlated with future productivity growth nor future stock returns, and that optimism predicts positive forecast errors in managerial guidance about sales. We therefore interpret the association of optimism with hiring as arising from the impact of non-fundamental, narrative optimism on firms’ hiring decisions. To underscore this interpretation, we show that changes in optimism driven by plausibly exogenous changes in CEOs (*i.e.*, those caused by death, illness, personal issues, or voluntary retirement of an incumbent CEO, as coded by Gentry, Harrison, Quigley, and Boivie, 2021) lead to quantitatively similar effects on hiring. Finally, we study the relevance of the Perennial Economic Narratives and our estimated topics. In light of the high-dimensionality of this set of narratives, we use the rigorous LASSO method of Belloni, Chen, Chernozhukov, and Hansen (2012) for estimation. We find that two of the nine Perennial Economic Narratives and eleven of the one hundred topics are relevant for hiring.

We next examine how our measured narratives spread across firms over time. Our second main result is that greater aggregate optimism and higher aggregate real GDP growth are associated with a greater probability that a firm is optimistic in the following year. We find similar effects at the industry-level when we control non-parametrically for aggregate conditions. Moreover, both the aggregate and industry-level patterns are robust to controlling

for future idiosyncratic and aggregate economic conditions. This rules out the explanation that aggregate optimism drives future optimism through its correlation with omitted positive news about economic conditions. Thus, we interpret our estimated spillover of past optimism as evidence of virality and our estimated positive responsiveness of optimism to economic conditions as evidence of associativeness.² We perform similar analysis for the other decision-relevant narratives. We find that almost all of them are viral and many are associative.

Model. Having provided evidence of the premises of narrative macroeconomics, we embed them in a macroeconomic model to understand their implications. The consumption, production, and labor supply side of the model is a (purely real) variant of the standard Neoclassical synthesis model of [Woodford \(2003\)](#) and [Galí \(2008\)](#). In particular, our model features aggregate demand externalities ([Blanchard and Kiyotaki, 1987](#)), which generate a motive among firms to co-ordinate the levels of their production. Narratives affect firms’ beliefs about the state of aggregate productivity. In our main analysis, we specialize to a case with two narratives: *optimism* and *pessimism*. The evolution of narratives is governed by the probabilities that optimists and pessimists become optimistic as a function of aggregate output (associativeness) and the fraction of optimists in the population (virality).

Theoretical Results. We first establish that there is an essentially unique equilibrium in which aggregate output is log-linear in aggregate productivity and a non-linear function of a sufficient statistic for narratives, the fraction of optimists in the population.³ In the limiting case of unanimous optimism, the contribution of optimism equals the partial-equilibrium effect of optimism on one firm’s hiring, exactly as we measured empirically, times a general-equilibrium multiplier. This exemplifies that optimism matters both directly for firms and indirectly through aggregate demand externalities, or a “Keynesian-cross” logic.

We next describe the model’s dynamics. We first show that, for fixed fundamentals, there always exists a steady-state level of optimism. Second, we characterize when extreme optimism and pessimism are stable steady states. Third, we provide the precise conditions under which extreme optimism and pessimism are both simultaneously stable steady states. In this case, depending on the initial fraction of optimists, the economy is (almost everywhere) globally attracted to one of these extreme steady states. Thus, small differences in

²This analysis relies on aggregate variation to draw inference about cross-agent spillovers, and is subject to classic critiques of other studies of “peer effects” (see, *e.g.*, [Angrist, 2014](#)). To ameliorate these concerns, we use the same plausibly exogenous CEO changes as instruments for industry-level optimism and find qualitatively consistent effects.

³The sense in which this equilibrium is *essentially* unique is that: (i) it is unique when fundamentals are bounded and (ii) when fundamentals are not bounded, it is a ε -equilibrium for any $\varepsilon > 0$ to the unique equilibrium with bounded fundamentals for some sufficiently large bound on fundamentals.

the economy’s initial optimism can make the difference between the economy settling in an optimistic, high-output state and a pessimistic, low-output state – a form of hysteresis.

Owing to these forces, the economy can feature hump-shaped and discontinuous impulse responses. Moreover, the economy can regularly oscillate between extreme optimism and pessimism. We provide analytical upper bounds on the expected period of these oscillations.

Quantification. In the last part of the paper, we identify the parameters of our model using our empirical estimates, macroeconomic time-series evidence, and external calibration of standard preference and production parameters. We find aggregate output would be 7% higher in an economy in which all firms are optimistic than one in which all firms are pessimistic. Our measured aggregate peak-to-trough movement in optimism accounts for 32% of output loss during the Early 2000s Recession and 18% during the Great Recession. More systematically, we find that optimism accounts for 17% of the short-run (one-year) and 66% of the medium-run (two-year) autocovariance in output. We therefore argue that time-varying, endogenous optimistic narratives are a significant driver of macroeconomic fluctuations.⁴

Related Literature. Our analysis relates to a large literature that studies business-cycle and financial fluctuations through time-variation in agents’ beliefs. First, our modelling of narratives and their spread relates to the work of [Carroll \(2001\)](#), [Burnside, Eichenbaum, and Rebelo \(2016\)](#), [Schaal and Taschereau-Dumouchel \(2020\)](#), and [Shiller \(2017\)](#), in which beliefs spread virally between agents and explain inflation dynamics, boom-bust cycles, and discussion of narratives, respectively. At the same time, our theoretical approach differs in both modelling narratives as forming a common set of building blocks of agents’ beliefs and studying how endogenous macroeconomic outcomes shape the spread of narratives.⁵

Second, our work is related to papers by [Beaudry and Portier \(2006\)](#), [Christiano, Ilut, Motto, and Rostagno \(2008\)](#), [Lorenzoni \(2009\)](#), [Angeletos and La’O \(2013\)](#), [Benhabib, Wang, and Wen \(2015\)](#), [Benhima \(2019\)](#), and [Bhandari, Borovička, and Ho \(2019\)](#) which postulate that the economy undergoes exogenous shocks to demand via news, noise, or sentiment. Our work can be understood in this context as providing a micro-foundation for such shocks via the endogenous evolution of narratives and corresponding degree of optimism.⁶ Our work contrasts with that of [Maxted \(2020\)](#), where agents extrapolate from recent changes in funda-

⁴Normatively, we find that viral optimism is welfare-improving and welfare-equivalent to a 1.3% *ad valorem* production subsidy.

⁵Thus, our model also differs from recent theoretical work in which models correspond to likelihoods ([Schwartzstein and Sunderam, 2021](#)) or directed acyclic graphs ([Spiegler, 2016](#); [Eliaz and Spiegler, 2020](#)).

⁶In an asset pricing context, [Maenhout, Vedolin, and Xing \(2021\)](#) study model updating within the robust control approach of [Hansen and Sargent \(2001\)](#) and assess how optimism evolves and impacts prices.

mentals, in that both equilibrium outcomes and social dynamics determine agents’ models.⁷ We thereby also provide a micro-foundation for the exogenous state-variation in optimism in Caballero and Simsek (2020), in which optimism drives asset pricing and consumption dynamics.⁸ Moreover, our normative results contribute to the literature understanding the social value of information (Morris and Shin, 2002; Angeletos and Pavan, 2007) by demonstrating the potentially welfare-enhancing effects of misspecified beliefs.

Finally, in studying the dynamics of misspecified models, we relate to a large macroeconomics and theory literature on model misspecification and learning (Bray, 1982; Bray and Savin, 1986; Marcet and Sargent, 1989a,b; Esponda and Pouzo, 2016; Acemoglu, Chernozhukov, and Yildiz, 2016; Adam, Marcet, and Beutel, 2017; Molavi, 2019; Frick, Iijima, and Ishii, 2020; Bohren and Hauser, 2021; Fudenberg, Lanzani, and Strack, 2021). This literature primarily characterizes the limit points of agents’ models. Instead, we study short-run fluctuations and time-variation in the models held by agents. This approach is similar to that of Kozlowski, Veldkamp, and Venkateswaran (2020), but we differ in our non-Bayesian and analytical, rather than computational, approach.

By studying narratives empirically, we relate to a literature assessing narratives following Carroll (2001) and Shiller (2017).⁹ Of most relevance are the papers by Andre, Haaland, Roth, and Wohlfart (2021), who use surveys to understand narratives underlying inflation, and Bybee, Kelly, Manela, and Xiu (2021), who apply LDA to the full text of *Wall Street Journal* articles to extract narrative time series. Our approach differs in its use of text data about the cross-section of firms to extract narratives, uncover their effects on decision making, and study their spread. Our empirical analysis therefore relates to a growing literature in economics that uses text data (for a review, see Gentzkow, Kelly, and Taddy, 2019).

Outline. The rest of the paper proceeds as follows. In Section 2, we develop the general framework for studying narratives. In Section 3, we describe our data and measurement. In Section 4, we detail our empirical strategy and results. In Section 5, we build our macroeconomic model with viral narratives. In Section 6, we provide theoretical results on macroeconomic dynamics. In Section 7, we calibrate our model and quantify the role of narratives. Section 8 concludes.

⁷Bordalo, Gennaioli, Kwon, and Shleifer (2021) show how a similar mechanism leads to bubbles.

⁸This relates our work to the large literature on optimism, overconfidence and economic activity (see, e.g., Harrison and Kreps, 1978; Scheinkman and Xiong, 2003; Barberis, Greenwood, Jin, and Shleifer, 2018).

⁹By studying how narratives shape agents beliefs, we indirectly relate to the large empirical literature on macroeconomic expectations formation (for a review, see Angeletos, Huo, and Sastry, 2021).

2 Narratives: A Conceptual Framework

We first describe the conceptual framework underlying our analysis. This framework is designed to parsimoniously represent the two premises of the macroeconomics of narratives. First, narratives form a set of common building blocks of agents’ beliefs. Second, narratives spread between agents virally and as a function of how well they describe economic reality. We embed these two premises in an abstract game where agents care about their own actions, fundamentals, and aggregates of other agents’ actions. This allows us to derive two regression equations, comprising the mapping from narratives to actions and the ways in which agents update their narratives, which we will bring to the data. Our main macroeconomic model in Section 5 is a specialization of this framework that generates precisely these regressions.

2.1 Premise I: Narratives Are the Building Blocks of Beliefs

The first premise is that narratives form the building blocks of agents’ beliefs.¹⁰ To model this, suppose that there is a continuum of agents indexed by i , of unit measure, and uniformly distributed over $[0, 1]$. We think of these agents as the firms or households that comprise the economy. There is an underlying payoff-relevant state space of aggregate fundamentals $\theta \in \Theta$. For example, these fundamentals might represent aggregate productivity or the strength of demand.

An individual narrative is a model of fundamentals. We describe each narrative, indexed by $k \in K = \{1, \dots, K\}$, as a probability distribution $N_k \in \Delta(\Theta)$ within the set of narratives $\mathcal{N} = \{N_k\}_{k \in K}$. For example, if the fundamental θ describes the strength of productivity, then a pessimistic narrative N_P might correspond to “productivity in the economy is low,” while an optimistic narrative N_O may represent the view that “productivity in the economy is high.” In this example, we might capture the relationship between these narratives mathematically through the distribution of productivity under the optimistic narrative being greater than the distribution of productivity under the pessimistic narrative in the sense of first-order stochastic dominance (FOSD).

Agents combine narratives to form priors about the fundamental by placing a vector of weights $\lambda = \{\lambda_k\}_{k \in K} \in \Delta(K)$ on each narrative. An agent with narrative weights λ has an induced prior distribution over fundamentals given by the following linear combination of distributions in \mathcal{N} :

$$\pi(\theta) = \sum_{k \in K} \lambda_k N_k(\theta) \tag{1}$$

¹⁰This approach of viewing narratives as forming a common set of stories that underpin beliefs features prominently in the management and organizational literature (see, *e.g.*, Isabella, 1990; Maitlis, 2005; Loewenstein, Ocasio, and Jones, 2012; Vaara, Sonenshein, and Boje, 2016).

Continuing the example, an agent who is fully pessimistic might place weight $\lambda_P = 1$ on the pessimistic narrative and complementary weight $\lambda_O = 0$ on the optimistic narrative, so their subjective probabilities for each state θ are $\pi(\theta) = N_P(\theta)$. An agent who has been convinced by neither narrative might take a middle ground and consider both equally likely, which we would represent with $(\lambda_P, \lambda_O) = (\frac{1}{2}, \frac{1}{2})$ or beliefs $\pi(\theta) = \frac{1}{2}N_O(\theta) + \frac{1}{2}N_P(\theta)$.

The extent of narrative penetration in the minds of agents is summarized by the cross-sectional distribution of narratives in the population $Q \in \Delta(\Lambda)$. This represents the distribution of agents' distributions of narrative weights. For example, in an economy populated by only optimists $\lambda^O = (0, 1)$, pessimists $\lambda^P = (1, 0)$ and moderates $\lambda^M = (\frac{1}{2}, \frac{1}{2})$, we would have that $Q = (Q^O, Q^P, Q^M)$ corresponds to the fraction of the population with each combination of weights over optimism and pessimism.

2.2 Premise II: Narrative Spread Is Viral and State-Dependent

We now model the evolution of the distribution of narratives in the population as a function of current narrative penetration and endogenous aggregate outcomes. Time is discrete and infinite, indexed by $t \in \mathbb{N}$. We index aggregate endogenous outcomes by $Y \in \mathcal{Y}$. For example, these aggregate outcomes might capture aggregate production or stock prices. In the next subsection, we describe how these outcomes are determined. For now, we simply study their influence on narrative evolution.

This evolution is described by an updating rule $P : \Lambda \times \mathcal{Y} \times \Delta(\Lambda) \rightarrow \Delta(\Lambda)$, which returns the probabilities $f^P(\lambda, Y, Q)$ that an agent with narrative weights λ changes their weights to λ^θ when the endogenous state is Y and the distribution of narratives in the population is Q .¹¹ Hence, conditional on a distribution of narratives at time t given by Q_t , if the realized endogenous outcomes are Y_t then:

$$Q_{t+1} = \sum_{\lambda} Q_t \cdot P(\lambda, Y_t, Q_t) \quad (2)$$

At this level of generality, the updating function can capture classical Bayesian updating by agents given some latent information structure. However, we can also model behavioral phenomena such as associative learning where agents associate certain states of the economy with certain models (*e.g.*, “aggregate output is high, therefore productivity is high”), and virality, wherein the distribution of narratives itself affects updating.¹²

¹¹In Appendix B.3, we extend this setting to allow for idiosyncratic fundamentals and updating that depends on their realizations.

¹²Associativeness in updating has recently been shown to be important in explaining forecasts by both households and experts (Andre, Pizzinelli, Roth, and Wohlfart, 2022).

2.3 Closing the Framework: From Beliefs to Actions to Outcomes

To close the framework, we need to describe the mapping from beliefs to agents' actions and model how agents' actions are mapped into endogenous outcomes. Agents care about their own actions $x_{it} \in X$, aggregate outcomes $Y_t \in Y$, the fundamental state $\theta_t \in \Theta$, and an idiosyncratic preference shifter $\omega_i \in \Omega$. They have utility functions $u : X \times Y \times \Theta \times \Omega \rightarrow \mathbb{R}$ and information sets I_{it} . Agents' place weights $\lambda_{it} \in Q_t$ on the various narratives with resulting priors π_{it} (see Equation 1). Given a conjecture about the mapping from fundamental states to aggregates $\hat{Y}_t : \Theta \rightarrow Y$, the agents chooses their action to maximize their expected utility:

$$\max_{x_{it} \in X} E_{\lambda_{it}} \left[u(x_{it}, \hat{Y}_t(\theta_t), \theta_t, \omega_i) \mid I_{it} \right] \quad (3)$$

where we note that narratives have the potential to shift agents' best replies by altering their prior beliefs about fundamentals and resulting aggregate outcomes.

Finally, cross-sectional distributions of agents' actions $G \in \Delta(X)$ are aggregated according to some aggregation functional $F : \Delta(X) \times \Theta \rightarrow Y$. Hence, given a distribution of play by the agents G and state θ , the endogenous aggregate outcome is $Y = F(G, \theta)$. As fundamentals enter not only agents' preferences and production technologies but also aggregation, they can capture differences in how agents interact and how their decisions aggregate.

An equilibrium in this setting can then be compactly described by an aggregate outcome function that is consistent with aggregating the distribution of agents' actions and agents' actions being derived from utility maximization given this aggregate outcome function:

Definition 1 (Equilibrium). *An equilibrium (Y_t, G_t) comprises an aggregate outcome function $Y_t : \Theta \rightarrow Y$ and an action distribution functional $G_t : \Theta \rightarrow \Delta(X)$ such that:*

1. *Aggregate outcomes are consistent with aggregation: $Y_t(\theta) = F(G_t(\theta), \theta)$ for all $\theta \in \Theta$*
2. *Distributions are consistent with utility maximization: $G_t(x; \theta) = \int_{[0,1]} P[x_{it} = x \mid \theta] di$ for all $\theta \in \Theta$ and $x \in X$ where x_{it} is a solution to Equation 3 given Y_t for all $i \in [0, 1]$.*

In an equilibrium, agents have “rational expectations” in the sense that they correctly understand the mapping from true fundamentals to aggregate outcomes. But, insofar as they entertain different narratives, they agree to disagree on the probability that different fundamental states occur, and hence the probability that different outcomes are realized. In this way, narratives shape both fundamental and strategic uncertainty.

Moreover, aggregate outcomes Y_t depend on the distribution of narratives Q_t . Thus, narratives regarding the joint evolution of fundamentals and outcomes depend on not only the narrative considered by an agent, but also those considered by others.

The combination of an equilibrium aggregate outcome function Y_t and the structure of narrative evolution (Equation 2) induces a first-order, non-linear, and stochastic difference equation for the path of the narrative distribution in the population. Our theoretical analysis will study the properties of this system and its implications for macroeconomic dynamics.

2.4 From the Conceptual Framework to an Empirical Framework

To test these premises, we need to identify narratives, extract agents' loadings onto these narratives and then assess: (i) how agents' narratives affect their actions and (ii) how agents update their narratives. We now describe how we operationalize this framework to accomplish these two steps. First, to study how agents' narratives affect their actions in general equilibrium, we linearize their best reply conditions to obtain:

$$x_{it} = \gamma_i + \chi_t + \sum_{k=1}^K \delta_k \lambda_{k;it} + \varepsilon_{it} \quad (4)$$

See Appendix A.1 for the formal derivation of this equation (Proposition 8). Intuitively, γ_i captures time-invariant factors such as preference shifters that affect the choice variable, χ_t captures time-varying general equilibrium factors such as the distribution of narratives in the population and fundamentals, δ_k correspond to the appropriately normalized expectation of both fundamentals and endogenous aggregate outcomes under narrative k , and ε_{it} corresponds to noise around these expectations caused by differences in the information sets across agents. In Appendix A.1, we also provide assumptions sufficient to guarantee a quadratic misspecification bound and show that ε_{it} is mean zero and independent from γ_i, χ_t , and λ_{it} . This latter point implies that, modulo issues of misspecification, the δ_k can be estimated consistently via a simple OLS regression. In our structural model in Section 5, this equation will hold in equilibrium without approximation.

Second, to study how agents update their narratives in equilibrium, we linearize the narrative updating equations to obtain the following system of linear probability models:

$$P[\lambda_{it} = \lambda | \lambda_{i,t-1}, Y_{t-1}, Q_{t-1}] = \zeta + \sum_{\theta=2} u_{\theta} |[\lambda_{i,t-1} = \lambda^{\theta}] + r^{\theta} Y_{t-1} + s^{\theta} Q_{t-1} \quad (5)$$

See Appendix A.1 for the formal derivation of this equation (Proposition 9). Intuitively, u captures the agents' stubbornness in updating, r captures the associativeness in updating, and s captures virality in updating.

3 Data, Measurement, and Descriptive Statistics

To take this framework to the data, we seek to recover the structure of narratives from agents’ language. To this end, we use the universe of public firm communication from SEC Form 10-K as our corpus of language. We then use both supervised and unsupervised methods to extract both the structure of narratives and firms’ loadings onto these narratives. We combine these measures of narratives with data on firm fundamentals and choices. Finally, we provide descriptive facts regarding the time-series and cross-sectional properties of narratives.

3.1 Data

Narratives. To extract the narratives that firms express in their language, we use the universe of SEC Forms 10-K. Each publicly traded firm in the U.S. submits a 10-K each year to the SEC. These forms provide “a detailed picture of a company’s business, the risks it faces, and the operating and financial results of the fiscal year.” Moreover, “company management also discusses its perspective on the business results and what is driving them” (SEC, 2011). This description is consistent with our notion that agents’ narratives constitute a view of the world and its rationalization via some model.

We download the universe of SEC forms 10-K from the SEC Edgar database from 1995 to 2019. This yields a corpus of 182,259 html files comprising the underlying text of the 10-K, various formatting information, and tables. We describe our exact method for processing the text data in Online Appendix C.1. The three key steps are pre-processing the raw text data to isolate English-language words, associating words with their common roots via lemmatization, and fitting a bigram model that groups together co-occurring two-word phrases. We then count the occurrences of all words, including bigrams, in all documents to obtain the bag-of-words representation (*i.e.*, a vector of word counts) for each document. Our final sample consists of 100,936 firm-by-year observations from 1995 to 2018.

As an alternative source of text data, we use public firms’ sales and earnings conference calls. Our initial sample consists of 158,810 documents from 2002 to 2014. We apply the same natural-language-processing techniques that we employ for the 10-Ks to this corpus. We average variables over the periods between successive 10-Ks to obtain a firm-by-fiscal-year dataset. Our final sample consists of 25,589 firm-by-year observations. We describe more details in Online Appendix C.2.

Firm Fundamentals and Choices. We compile our dataset of firm fundamentals and choices using Compustat Annual Fundamentals from 1995 to 2018. This dataset includes information from firms’ financial statements on employment, sales, input expenses, and

capital. We apply standard selection criteria to screen out firms that are very small, report incomplete information, or were likely involved in an acquisition. We also, for our main analysis of employment effects, ignore firms in the financial and utilities sectors due to their markedly different production and/or market structure. Details about our sample selection can be found in Online Appendix D.1.

We organize firms into 44 industries, which are defined at the NAICS 2-digit level, but for Manufacturing (31-33) and Information (51), which we split into the 3-digit level. To study narrative transmission at a finer level, we also define peer sets for the subset of firms traded on the New York Stock Exchange using the method of Kaustia and Rantala (2021). These authors exploit common equity analyst coverage to define peers for each firm.¹³

To measure total factor productivity, we estimate a constant-returns-to-scale, Cobb-Douglas, two-factor production function in materials and capital, for each industry. We estimate the output elasticities using the ratio of materials expenditures to total sales and an assumed revenue returns to scale of 0.75. More details are provided in Online Appendix D.2. We denote our estimated log-TFP variable as $\log \hat{\theta}_{it}$.

Manager and Analyst Beliefs. We collect data from IBES (the International Brokers' Estimate System) on quantitative sales forecasts by companies and their equity analysts. Specifically, we use the IBES Guidance dataset which records, for specific variables, both (i) a numerical management expectation recorded from press releases or transcripts of corporate events and (ii) a contemporaneous consensus value from equity analysts. We restrict to the first recorded forecast per fiscal year of that year's sales. When managers' guidance is reported as a range, we code a point-estimate forecast as the range's midpoint. We construct two variables from these data at the level of firms i and fiscal years t . The first, $\text{GuidanceOptExAnte}_{it}$, is an indicator of managers' guidance exceeding the analyst consensus. The second, $\text{GuidanceOptExPost}_{it}$, is an indicator of managers' guidance minus the realization (both in log units) exceeding the sample median.¹⁴

CEO Changes. To obtain plausibly exogenous variation in narratives held at the firm level, we will examine the year-to-year change in firm-level narratives stemming from plausibly exogenous CEO changes. To do this, we use the dataset of categorized CEO exits compiled by Gentry, Harrison, Quigley, and Boivie (2021). These data comprise 9,390 CEO

¹³Firm j is a peer of firm i at time t if they have more than C common analysts, where C is chosen so that the probability of having C or more common analysts by chance is less than 1% when analysts following firm i randomly choose the firms they follow among all firms with analysts in period t .

¹⁴This method corrects for the fact that, in more than half of our observations, guidance is lower than the realized value, presumably due to asymmetric incentives.

turnover events categorized by the reason for the CEO exit.¹⁵ We restrict attention to CEO exits caused by death, illness, personal issues, and voluntary retirements. Importantly, we exclude all CEO exits caused by inadequate job performance, quits, and forced retirement.

3.2 Measurement: Recovering Narratives from Language

We identify sentiment in firm language and a more specific set of narratives using both a supervised method comparing the text of filings to the description of the various narratives identified by Shiller (2020) and Latent Dirichlet Allocation (LDA), an unsupervised machine-learning method for topic analysis.

Sentiment Narratives. To measure firm sentiment, we first categorize individual words as either positive or negative using the dictionaries constructed by Loughran and McDonald (2011). The goal of these dictionaries is to correct standard tools for English-language sentiment analysis to more precisely score financial communications, in which certain words (*e.g.*, the leading example “liability”) have specific definitions.¹⁶ We first define \mathcal{W}_P as the set of positive words and \mathcal{W}_N as the set of negative words. For reference, we print the 20 most common words in each set in Appendix Table A1. We calculate positive and negative sentiment as:

$$\text{pos}_{it} = \sum_{w \in \mathcal{W}_P} \text{tf}(w)_{it} \quad \text{neg}_{it} = \sum_{w \in \mathcal{W}_N} \text{tf}(w)_{it} \quad (6)$$

where $\text{tf}(w)_{it}$ is the term frequency of word w in the time- t 10-K of firm i . We then construct a one-dimensional measure of net sentiment, sentiment_{it} , by computing the across-sample z -scores of both positive and negative sentiment and taking their difference. Finally, we define a firm i as being optimistic at time t if its sentiment is above the entire-sample median:

$$\text{opt}_{it} = \mathbb{1}[\text{sentiment}_{it} > \text{med}(\text{sentiment}_{it})] \quad (7)$$

Narrative Identification of Narratives. To measure more specific narratives entertained by firms, we first consider a supervised strategy by narratively identifying a set of narratives using the text of Shiller’s *Narrative Economics*. Shiller identifies a set of nine *Perennial Economic Narratives*: Panic versus Confidence; Frugality versus Conspicuous Consumption; The Gold Standard versus Bimetallism; Labor-Saving Machines Replace Many Jobs; Automation and Artificial Intelligence Replace Almost All Jobs; Real Estate

¹⁵The categorization was performed using primary sources (*e.g.*, press releases, newspaper articles, and regulatory filings) by undergraduate students in a computer lab, supervised by graduate students, with the final dataset checked by both a data outsourcing company and an additional student.

¹⁶Loughran and McDonald (2011) generate the dictionaries based on human inspection of the most common words in the 10-Ks and their usage in context. We describe more details of our method in Section C.3.

Booms and Busts; Stock Market Bubbles; Boycotts, Profiteers, and Evil Businesses; and The Wage-Price Spiral and Evil Labor Unions. Each of these narratives and its history is described in its own chapter in *Narrative Economics*. To measure the extent to which each narrative is adopted by the firms in our sample, we compute the textual similarity between each 10-K filing and the relevant chapter of the book.

Formally, we use a method related to prior work by Hassan, Hollander, Van Lent, and Tahoun (2019) and our own implementation in Flynn and Sastry (2021). For each narrative k , we first compute the term-frequency-inverse-document-frequency (tf-idf) score to obtain a set of words most indicative of that narrative:

$$\text{tf-idf}(w)_k = \text{tf}(w)_k \log \left(\frac{1}{\text{df}(w)} \right) \quad (8)$$

where $\text{tf}(w)_k$ is the number of times that word w appears in the chapter corresponding to narrative k in *Narrative Economics* and $\text{df}(w)$ is the fraction of 10-K documents containing the word. Intuitively, if a word has a higher tf-idf score, it is both common in Shiller’s description of a narrative but relatively uncommon in 10-K filings. We define the set of 100 words with the highest tf-idf score for narrative k as \mathcal{W}_k . For reference, we print the twenty most common words in each set \mathcal{W}_k in Appendix Table A2.

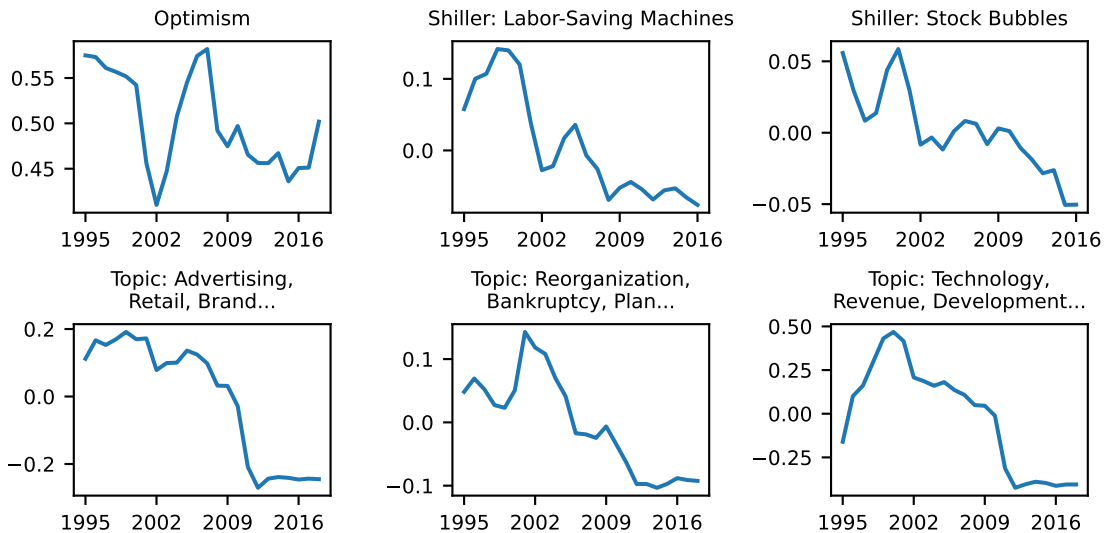
Finally, we score document (i, t) for narrative k by the total frequency of narrative words:

$$\text{Shiller}_{it}^k = \sum_{w \in \mathcal{W}_k} \text{tf}(w)_{it} \quad (9)$$

We then compute the document’s loading on each narrative, Shiller_{it}^k , by taking the z -score.

Unsupervised Recovery of Narratives. The final technique we use to recover the structure of narratives is Latent Dirichlet Allocation (LDA), which is a hierarchical Bayesian model of the corpus of text in which documents are constructed by combining a latent set of topics (Blei, Ng, and Jordan, 2003). Formally, given our corpus of 10-Ks with M documents, we postulate that there are $K = 100$ topics. First, the number of words in each document is drawn from a Poisson distribution with parameter ξ . Second, the distribution of topics in each document is given by $\vartheta = (\vartheta_1, \dots, \vartheta_M)$, over which we impose a Dirichlet prior with parameter $\alpha = \bar{f}\alpha_k g_{k2K}$, where α_k represents the prior weight that topic k is in any document. Third, the distribution of words across topics is given by $\phi = (\phi_1, \dots, \phi_K)$, over which we impose a Dirichlet prior with parameter $\beta = \bar{f}\beta_{jk} g_{k2K}$, where β_{jk} is the prior weight that word j is in topic k . Finally, we assume that individual words in each document d are generated by first drawing a topic z from a multinomial distribution with parameter ϑ , and then selecting a word from that topic by drawing a word from a multinomial dis-

Figure 1: Aggregate Time Series for Six Selected Narratives



tribution with parameter ϕ_z . Intuitively, in an LDA, the set of documents is formed of a low-dimensional space of narratives of co-occurring words.

To estimate the LDA, we use the Gensim implementation of the variational Bayes algorithm of Hoffman, Bach, and Blei (2010), which makes estimation of LDA on our large dataset feasible where standard Markov Chain Monte Carlo methods are very slow.¹⁷ Given the estimated LDA, we construct the document-level narrative score as the posterior probability of that topic in the estimated document-specific topic distribution \hat{p} :

$$\text{topic}_{it}^k = \hat{p}(k|d_{it}) \quad (10)$$

For each of the eleven topics that our subsequent analysis identifies as relevant for hiring (see Section 4.1), we print the ten highest-weight bigrams and their weights in Appendix Table A3. These topics are qualitatively different from the word sets used by our sentiment scoring (Appendix Table A1) and Shiller narratives (Appendix Table A2).

3.3 Descriptive Analysis of Narratives

Before our main empirical analysis, we first describe the time-series and cross-sectional structure of our measured narratives.

¹⁷For computational reasons, we use all available documents from a randomly sampled 10,000 of our 37,684 unique possible firms.

Time-Series Properties. In Figure 1, we show the time path of six selected measured narratives: optimism, “Labor-Saving Machines” and “Stock Bubbles” from Shiller’s perennial narratives, and three topics whose three most common terms are “Advertising, Retail, Brand”; “Reorganization, Bankruptcy, Plan”; and “Technology, Revenue, Development.” Our choices among the Shiller chapters and unsupervised topics are among the set that our later analysis suggests is particularly important for explaining hiring in the cross-section. At a glance, all of these narratives are highly persistent and feature business-cycle fluctuations; some also have notable trends.¹⁸ In Appendix Table A4, we report summary statistics for all narratives’ autocorrelation and correlation with unemployment. Almost all measured narratives are persistent, and several among the Shiller and topic sets are pro- or counter-cyclical. This observation is consistent with existing evidence in the literature on the cyclical nature of aggregate text-based measures of narratives (*e.g.*, Shiller, 2020) and news coverage (*e.g.*, Baker, Bloom, and Davis, 2016; Bybee, Kelly, Manela, and Xiu, 2021).

However, it is challenging to interpret these basic time-series facts for two reasons. First, as noted by Shiller (2017), it is hard to disentangle the dual roles of narratives in driving behavior versus describing fundamentals. In the next section, we will use cross-sectional variation in narratives to isolate the impact on behavior. Second, without an understanding of how narratives affect decisions and how decisions aggregate, it is difficult to understand the macroeconomic implications of even desirable time-series variation in optimism. We will later combine our macroeconomic model with our microeconomic evidence to evaluate the business-cycle impact of aggregate variation in narratives quantitatively.

Cross-Sectional Properties. Our firm-level panel allows us to explore variation that is more fine-grained than the time-series variation of Figure 1. We perform a variance decomposition of each narrative variable by comparing the total variance of each variable, N_{it} , with the variance after removing means at the time, industry, industry-by-time, and firm levels. In Appendix Table A5, we present the results in units of the fraction of variance explained by each level of fixed effects, relative to the total. Time-series variation constitutes a very small percentage of the total variation in our variables – only 1.1% for optimism, 0.2% for the median Shiller narrative, and 3.5% for the median topic narrative. Adding industry-specific trends increases these percentages, respectively, to only 6.7%, 8.7%, and 9.9%. The vast majority of variation is therefore at the firm level.

¹⁸In the Appendix, we report the time-series plot for Positive Sentiment, Negative Sentiment, and their difference (Figure A1); all nine Shiller (2020) Perennial Economic Narratives (Figure A2); and all eleven LDA topics that our later analysis determines are relevant for hiring (Figure A3).

4 Empirical Results

To move beyond descriptive evidence, we leverage our dataset of firm-level outcomes and narrative loadings to assess the effect of narratives on firm decision-making and to understand the diffusion of narratives across firms over time.

4.1 Narratives Matter for Decisions

Empirical Strategy. From the conceptual framework in Section 2 (see Equation 4 and Proposition 8 in Appendix A.1), we have that firm hiring $\Delta \log L_{it}$ can be described to first order by the following regression equation:

$$\Delta \log L_{it} = \sum_{k=1}^K \delta_k \lambda_{k;it} + \gamma_i + \chi_t + \varepsilon_{it} \quad (11)$$

where the $\lambda_{k;it}$ are firm-specific loadings on narratives indexed by k , γ_i is a fixed effect spanning static firm characteristics, χ_t is a fixed effect spanning aggregate conditions (including both fundamentals and the distribution of narratives), and ε_{it} is a residual term arising from idiosyncratic noise in individuals' signals.

We first operationalize this by estimating the following regression equation relating hiring to our optimism variable constructed in the 10-Ks:¹⁹

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i);t} + \tau^\theta X_{it} + \varepsilon_{it} \quad (12)$$

Hiring and optimism are constructed as described in Section 3, at the level of firms and fiscal years. We augment our theoretically implied specification (Equation 11) with controls, including industry-by-time fixed effects and a suite of firm-level time-varying controls X_{it} (current and past TFP, lagged labor, and financial variables). We later estimate analogues of Equation 12 with other estimated narratives as independent variables.

Main Result: Optimism Drives Decisions. We present our estimates of Equation 12 in Table 1. We first estimate the model with no additional controls beyond fixed effects and find a point estimate of $\hat{\delta}^{OP} = 0.0355$ with a standard error of 0.0030 (column 1). In column 2, we add controls for current and lagged TFP and lagged labor ($\log \hat{\theta}_{it}$, $\log \hat{\theta}_{i:t-1}$, $\log L_{i:t-1}$). These controls proxy both for time-varying firm fundamentals and, to first order, the presence of

¹⁹Throughout the analysis, when we estimate regressions with continuous outcome measures, we drop the top 1% and bottom 1% of the distribution. This is to guard against extremely large outliers likely due to measurement issues or unmodeled tail events.

Table 1: The Effect of Optimism on Hiring

	(1)	(2)	(3)	(4)	(5)
	Outcome is				
	$\Delta \log L_{it}$				$\Delta \log L_{i;t+1}$
opt_{it}	0.0355 (0.0030)	0.0305 (0.0030)	0.0249 (0.0033)	0.0322 (0.0028)	0.0216 (0.0037)
Industry-by-time FE	×	×	×	×	×
Firm FE	×	×	×		×
Lag labor		×	×	×	×
Current and lag TFP		×	×	×	×
Log Book to Market			×		
Stock Return			×		
Leverage			×		
N	71,161	39,298	31,513	40,580	38,402
R^2	0.259	0.401	0.416	0.142	0.398

Notes: Standard errors are two-way clustered by firm ID and industry-year. In column 5, control variables are dated $t + 1$.

adjustment costs in labor.²⁰ Our point estimate $\hat{\delta}^{OP} = 0.0305$ (SE: 0.0030) is quantitatively comparable to our uncontrolled estimate. In column 3, we add measures of firms' financial characteristics, the (log) book-to-market ratio, last fiscal year's stock return, and leverage (total debt over total assets). These controls proxy for both Tobin's q and firm-level financial frictions. The point estimate remains positive and quantitatively similar. In column 4, we estimate a specification with the controls from column 2 but no firm fixed effects to guard against small-sample bias (Nickell, 1981) and find similar results.

We finally estimate a specification in which the outcome and control variables are time-shifted one year in advance:

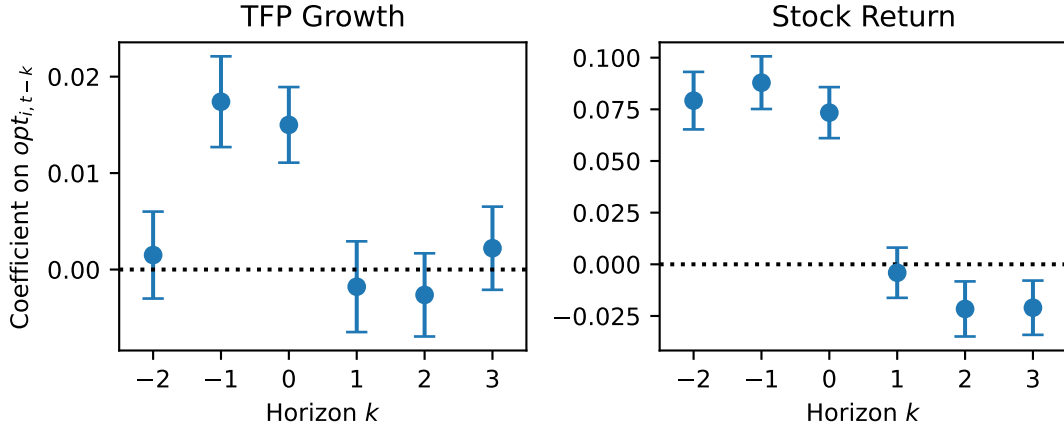
$$\Delta \log L_{i;t+1} = \delta_1^{OP} \text{opt}_{it} + \tau^\ell X_{i;t+1} + \gamma_i + \chi_{j(t);t+1} + \varepsilon_{i;t+1} \quad (13)$$

where δ_1^{OP} is the effect of lagged optimism on hiring and the (time-shifted) control variables $X_{i;t+1}$ are those studied in column 2. This specification ensures that hiring takes place in fiscal year $t + 1$ after the filing of the 10-K at the end of fiscal year t . Our point estimate in column 5 is similar in magnitude to our comparable baseline estimate (column 2).

In the Appendix, we report several additional results. We summarize them briefly here. Figure A4 shows estimates of a variant of our baseline regression interacting optimism with

²⁰For a forward-looking firm that observes current productivity, the analysis in Online Appendix B.5 makes precise that these controls capture the impact of adjustment costs to first order.

Figure 2: Dynamic Relationship of Optimism with Firm Fundamentals



Notes: Each coefficient is estimated from a separate projection regression. Error bars are 95% confidence intervals, based on standard errors clustered at the firm and industry-year level.

quartiles of firm characteristics. We find that the effect of optimism is decreasing in capital intensity, essentially flat in market capitalization, and U-shaped in the book-to-market ratio (*i.e.*, high for both growth firms and value firms). Table A6 repeats the analysis of Table 1 with our conference-call-based optimism measure, and finds similar results. Table A7 repeats our main analysis for different measured inputs – employment (the baseline), total variable input expenditure, and investment – and demonstrates a positive and comparably-sized effect of optimism on all three. Thus, optimism expands operations uniformly across inputs.

Inspecting the Mechanism: Optimism is Non-Fundamental. The coefficient of interest, δ^{OP} , measures the impact of optimism on hiring if $E[opt_{it}\varepsilon_{it} | \gamma_i, \chi_{j(\cdot):t}, X_{it}] = 0$. This is likely to be satisfied if optimism is unrelated to firm-level fundamentals that affect hiring. We have already demonstrated that controlling for firm-level productivity, current labor employed, and financial variables has minimal impact on the estimated value of δ^{OP} . Thus, any correlation between measured optimism and measured contemporaneous and lagged fundamentals does not generate quantitatively significant omitted variables bias. But we have not yet systematically investigated those correlations, or more formally explored whether measured optimism captures news about future fundamentals.

To investigate these issues, we estimate projection regressions of firm-fundamentals Z_{it} , either TFP growth $\Delta \log \hat{\theta}_{it}$ or stock returns R_{it} , on optimism at leads and lags k :

$$Z_{it} = \beta_k opt_{i,t-k} + \gamma_i + \chi_{j(\cdot):t} + \varepsilon_{it} \quad (14)$$

Table 2: Textual Optimism and Optimistic Forecasts

	(1)	(2)	(3)	(4)
	Outcome is			
	GuidanceOptExPost _{<i>i,t+1</i>}	GuidanceOptExAnte _{<i>i,t+1</i>}	GuidanceOptExPost _{<i>i,t+1</i>}	GuidanceOptExAnte _{<i>i,t+1</i>}
opt _{<i>it</i>}	0.0354 (0.0184)	0.0561 (0.0257)	0.0267 (0.0231)	-0.000272 (0.0353)
Ind.-by-time FE	×	×	×	×
Lag labor		×		×
Current and lag TFP		×		×
<i>N</i>	3,817	2,159	3,044	1,718
<i>R</i> ²	0.173	0.193	0.161	0.192

Notes: Standard errors are two-way clustered by firm ID and industry-year.

For negative k , β_k measures the relationship of current fundamentals with future optimism. For positive k , β_k measures the relationship of current fundamentals with past optimism.

We show our findings graphically in Figure 2, in which each point is a coefficient from a separate estimation of Equation 14 and the error bars are 95% confidence intervals.²¹ For $k < 0$, and both outcome variables, we find evidence of $\beta_k > 0$ – that is, a firm doing well today in terms of TFP growth or stock-market returns is likely to become optimistic tomorrow. However, for $k > 0$, and both outcome variables, we find no positive association – that is, a firm doing well today was not on average optimistic yesterday, or a firm that is optimistic today does not on average do better tomorrow. This is consistent with our required exclusion restriction that our narrative measure of optimism is non-fundamental, and inconsistent with a story that optimism is driven by news about fundamentals.²²

Inspecting the Mechanism: Optimistic Language Predicts Optimistic Beliefs. In our theoretical framework, optimistic narratives increase hiring by increasing firms’ expectations about fundamentals. To test this mechanism, we investigate the relationship between narrative optimism and the extent to which firms make more optimistic forecasts. As described in Section 3.1, we define variables GuidanceOptExPost_{*i,t+1*} and GuidanceOptExAnte_{*i,t+1*} to indicate firms’ optimism at the beginning of fiscal year $t + 1$ relative to realized sales and contemporaneous sales forecasts of equity analysts, respectively. For each variable

²¹Figure A5 replicates this analysis with conference-call-based optimism and finds similar results.

²²To further investigate the effects on stock prices, we also estimate the correlation of optimism with stock returns on and around the filing date of the 10-K. We present our results in Appendix Table A8. We find essentially no evidence of stock response on or before the filing day, and weak evidence of positive returns on the order of 15-25 basis points in the four days after.

GuidanceOpt $_{i,t+1}$, we estimate the following regression model:

$$\text{GuidanceOpt}_{i,t+1} = \beta \text{opt}_{it} + \tau^{\theta} X_{it} + \chi_{j(\cdot);t} + \varepsilon_{it} \quad (15)$$

The control variables X_{it} are current and lagged TFP and lagged labor, all in log units. As we have guidance data for only a small subset of firms, we do not include firm fixed effects.

Our findings are reported in Table 2. For optimism relative to realizations, we find a positive correlation that increases when we add the aforementioned control variables (columns 1 and 2). This is consistent with the notion that firms producing an optimistic 10-K truly hold optimistic views about firm performance. For optimism relative to analysts, we find an imprecise positive effect in an uncontrolled model and a zero effect in the controlled model. These findings are consistent with a story in which optimism is shared between management and investors, potentially due to persuasion in communications.

Alternative Strategy: CEO Change Event Studies. To further isolate variation in the narratives held by firms that is unrelated to fundamentals, we study the effects on hiring of changes in narratives induced by plausibly exogenous managerial turnover. As described in Section 3.1, for all firms i and years t such that i 's CEO leaves because of death, illness, personal issues or voluntary retirements, we estimate the regression equation

$$\Delta \log L_{it} = \delta^{CEO} \text{opt}_{it} + \delta_1^{CEO} \text{opt}_{i,t-1} + \tau^{\theta} X_{it} + \chi_{j(\cdot);t} + \varepsilon_{it} \quad (16)$$

This differs from our baseline Equation 12 by including parametric controls for lagged values of the narrative loadings, but removing a persistent firm fixed effect.²³ If the studied CEO changes are truly exogenous, as we have suggested, then the narrative loadings of the new CEO are, conditional on the narrative loadings of the previous CEO, solely due to the differences in worldview across these two senior executives. Of course, CEO exits may be disruptive and reduce firm activity. Any time- and industry-varying effects of CEO exits via disruption are controlled for by the intercept of the regression $\chi_{j(\cdot);t}$, since the equation is estimated only on the exit events. Moreover, any within-industry, time-varying, and idiosyncratic disruption is captured through our maintained productivity control. Under this interpretation, the coefficient of interest δ^{CEO} isolates the effect of optimism on hiring purely via the channel of changing managements' narratives.

We present our results in Table 3. We obtain estimates of δ^{CEO} that are quantitatively similar to our estimates of δ^{OP} in Table 1 (columns 1, 2, and 3). In column 4, we estimate a regression equation on the full sample that measures the direct effect of CEO changes and

²³With a firm fixed effect, the regression coefficients of interest would be identified only from firms with multiple plausibly exogenous CEO exits.

Table 3: The Effect of Optimism on Hiring, CEO Change Strategy

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0253 (0.0131)	0.0404 (0.0131)	0.0358 (0.0133)	0.0253 (0.0029)
$\text{opt}_{it} \text{ ChangeCEO}_{it}$				0.0220 (0.0099)
ChangeCEO_{it}				-0.0232 (0.0088)
Industry-by-time FE	×	×	×	×
Lag optimism	×	×	×	×
Lag labor		×	×	×
Current and lag TFP		×	×	×
Log Book to Market			×	
Stock Return			×	
Leverage			×	
N	1,725	982	872	36,953
R^2	0.243	0.375	0.379	0.134

Notes: Standard errors are two-way clustered by firm ID and industry-year.

its interaction with the new management’s optimism. Specifically, we estimate

$$\Delta \log L_{it} = \delta^{\text{NoChange}} \text{opt}_{it} + \delta^{\text{Change}} (\text{opt}_{it} \text{ ChangeCEO}_{it}) + \alpha^{\text{Change}} \text{ChangeCEO}_{it} + \delta_1^{\text{CEO}} \text{opt}_{i,t-1} + \tau^\theta X_{it} + \chi_{j(i);t} + \varepsilon_{it} \quad (17)$$

where ChangeCEO_{it} is an indicator for our plausibly exogenous CEO change events. We find that CEO changes in isolation reduce hiring ($\alpha^{\text{Change}} < 0$) but also that the effect of optimism is magnified when it accompanies a CEO change ($\delta^{\text{Change}} > 0$). This is further inconsistent with a story under which omitted fundamentals lead us to overestimate the effect of optimism on hiring.

The Narratives that Matter for Decisions. We now study the decision-relevance of the measured Shiller (2020) and topic narratives. Specifically, for each of the two sets of K narratives, we estimate the regression equation implied by our theoretical framework:

$$\Delta \log L_{it} = \sum_{k=1}^K \delta_k \hat{\lambda}_{k;it} + \gamma_i + \chi_{j(i);t} + \tau^\theta X_{it} + \varepsilon_{it} \quad (18)$$

Table 4: Narratives Selected as Relevant by LASSO

Shiller (2020) Chapters	Topics
1. Labor-Saving Machines	1. Lease, Tenant, Landlord
2. Stock Bubbles	2. Business, Public, Combination
	3. Value, Fair, Loss
	4. Advertising, Retail, Brand
	5. Financial, Control, Internal
	6. Stock, Compensation, Tax
	7. Gaming, Service, Network
	8. Debt, Credit, Facility
	9. Reorganization, Bankruptcy, Plan
	10. Court, Settlement, District
	11. Technology, Revenue, Development

We use our baseline controls, current and lagged TFP and lagged labor. Because we have many candidate narratives (9 and 100, respectively), and we expect only a few to matter for decisions, we apply the rigorous square-root lasso method of [Belloni, Chen, Chernozhukov, and Hansen \(2012\)](#) to estimate the subset of hiring-relevant narratives. In [Table 4](#), we list the selected [Shiller \(2020\)](#) and topic narratives. In the first two columns of [Appendix Table A9](#), we report the post-LASSO OLS estimates of [Equation 18](#) with the selected variables.

Among the [Shiller \(2020\)](#) Perennial Economic Narratives, the LASSO methodology selects two of nine as quantitatively relevant for hiring: “Labor-Saving Machines” and “Stock Bubbles.” Among the unsupervised topics, the LASSO methodology selects eleven variables out of 100. In [Table A9](#), we present these topics in the (essentially random) order they come out of our LDA exercise and identify them by their three highest-weight bigrams (in all cases, single words).²⁴ *Ex post*, based on their word combinations, we identify two as relating to demand conditions (Topics 4 and 7); two related to legal proceedings (Topics 9 and 10); one related to technology development (Topic 11); one related to real estate (Topic 1); and the remaining five related to financial conditions.²⁵

We finally quantify the extent to which loadings on these narratives are a mechanism for the effects of optimism on hiring. In particular, we estimate the following system of equations in which we treat optimism as an endogenous variable, and the (LASSO-selected)

²⁴[Appendix Table A3](#) prints the top ten words per topic and their numerical weights.

²⁵We can also verify in some cases that these more specific narratives, identified purely from textual data, affect related choices. For instance, we validate our interpretation of the “Advertising, Retail, Brand” narrative by measuring its relationship with the growth of SG&A expenditures. Specifically, we estimate an analogue of [Equation 12](#) with SG&A growth as the outcome and the “Advertising, Retail, Brand” topic score as the regressor, and we find a positive and significant estimate of $\delta = 0.0076$ (SE: 0.0022). We similarly test the relevance of the topic “Technology, Revenue, Development” for the growth of R&D spending and find a positive and significant estimate of $\delta = 0.0319$ (SE: 0.0036).

Shiller and topic narratives (indexed respectively up to K_S and K_T) as excluded instruments:

$$\begin{aligned} \Delta \log L_{it} &= \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i);t} + \tau^\theta X_{it} + \varepsilon_{it} \\ \text{opt}_{it} &= \sum_{k=1}^{K_S} \delta_{Sk} \text{Shiller}_{it}^k + \sum_{k=1}^{K_T} \delta_{Tk} \text{topic}_{it}^k + \tilde{\gamma}_i + \tilde{\chi}_{j(i);t} + \tilde{\tau}^\theta \tilde{X}_{it} + \tilde{\varepsilon}_{it} \end{aligned} \quad (19)$$

where X_{it} are, again, our baseline controls. We provide coefficient estimates for Equation 19 in column 4 of Appendix Table A9. The Shiller and topic narratives strongly predict optimism ($F = 189$), and our IV estimate of a 0.0597 log-point effect of optimism on hiring is larger than our baseline estimate of 0.0305.

4.2 Narrative Spread is Viral and Associative

Empirical Strategy. From the conceptual framework in Section 2 (see Equation 5 and Proposition 9 in Appendix A.1), we have that narrative updating is described by a system of linear probability models that depend on agents' fixed effects, agents' previous narrative weights, the previous narrative weights of the population, and economic outcomes.

To operationalize this idea in the context of our measured binary optimism, we first estimate the following model:

$$\text{opt}_{it} = u \text{opt}_{i;t-1} + s \overline{\text{opt}}_{t-1} + r \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it} \quad (20)$$

where $\overline{\text{opt}}_{t-1}$ is average optimism in the previous period, $\Delta \log Y_{t-1}$ is U.S. real GDP growth, and γ_i is an individual fixed effect. Following our earlier interpretation, u_{opt} measures stubbornness, s_{opt} measures virality, and r_{opt} measures associativeness.

Main Result: Optimism Spreads Virally and Associatively. In column 1 of Table 5, we show our estimates. We find strong evidence of $u > 0$, $s > 0$, and $r > 0$ – that is, firms are significantly more likely to be optimistic in year t if, in the previous year, they were optimistic, if other firms were optimistic, and if the economy grew.

Our estimates of Equation 20, and especially separate identification of virality (s) and associativeness (r), levers only the time-series variation over our studied 23-year period. We therefore also study a model that allows for virality and associativeness at the finer levels of our 44 industries and our firm-specific peer groups. Specifically, we estimate the equation:

$$\text{opt}_{it} = u_{\text{ind}} \text{opt}_{i;t-1} + s_{\text{ind}} \overline{\text{opt}}_{j(i);t-1} + s_{\text{peer}} \overline{\text{opt}}_{p(i);t-1} + r_{\text{ind}} \Delta \log Y_{j(i);t-1} + \gamma_i + \chi_t + \varepsilon_{it} \quad (21)$$

where $\overline{\text{opt}}_{j(i);t-1}$ and $\overline{\text{opt}}_{p(i);t-1}$ are (leave-one-out) means of optimism among a firm's in-

Table 5: The Virality and Associativeness of Optimism

	(1)	(2)	(3)
	Outcome is opt_{it}		
Own lag, $\text{opt}_{j:t-1}$	0.209 (0.0071)	0.214 (0.0080)	0.135 (0.0166)
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290 (0.0578)		
Real GDP growth, $\Delta \log Y_{t-1}$	0.804 (0.2204)		
Industry lag, $\overline{\text{opt}}_{j(i):t-1}$		0.276 (0.0396)	0.207 (0.0733)
Industry output growth, $\Delta \log Y_{j(i):t-1}$		0.0560 (0.0309)	0.0549 (0.0632)
Peer lag, $\overline{\text{opt}}_{p(i):t-1}$			0.0356 (0.0225)
Firm FE?	×	×	×
Time FE?		×	×
N	64,948	52,258	8,514
R^2	0.481	0.501	0.501

Notes: Standard errors are two-way clustered by firm ID and industry-year.

dustry and peer set, respectively, and $\Delta \log Y_{j(i):t-1}$ is the growth of sectoral value added measured by linking Bureau of Economic Analysis (BEA) sector-level data to our NAICS-based classification.²⁶ The time fixed effect χ_t absorbs aggregate virality and associativeness.

We show the results in columns 2 and 3 of Table 5. First, using just the industry-level data, we find strong evidence for virality and weaker evidence for associativeness within industries. Second, including the peer set optimism and restricting to the much smaller number of NYSE-listed firms, we find both a quantitatively similar industry-level effect and an independent peer-set effect. Moreover, the sum of coefficients $s_{\text{ind}} + s_{\text{peer}}$, the marginal effect of optimism in both the industry and peer set, is positive and strongly significant (estimate 0.243, standard error 0.075).

Inspecting the Mechanism: Spillovers are Not Driven by Common Fundamental Shocks. The coefficients of interest, u , r , and s identify stubbornness, virality, and associativeness, as defined in the model, exactly when $E[(\text{opt}_{it}, \overline{\text{opt}}_{t-1}, \Delta \log Y_{t-1}) \otimes \varepsilon_{it} \gamma_i] = 0$. An important case in which this would fail is if spillovers from optimism are correlated with the error term because they reflect staggered positive shocks to the economy. In this case,

²⁶These data are available only from 1997.

Table 6: The Virality of Optimism, Over-Controlling for Past and Future Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Outcome is opt_{it}						
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290 (0.0578)	0.339 (0.0763)	0.235 (0.1278)	0.222 (0.2044)			
Ind. lag, $\overline{\text{opt}}_{j(i):t-1}$					0.276 (0.0396)	0.241 (0.0434)	0.262 (0.0705)
Time FE					×	×	×
Firm FE	×	×	×	×	×	×	×
Own lag, $\text{opt}_{i:t-1}$	×	×	×	×	×	×	×
$(\Delta \log Y_{t+k})_{k=2}^2$		×	×	×			
$(\Delta \log Y_{j(i):t+k})_{k=2}^2$			×	×		×	×
$(\Delta \log \hat{\theta}_{i:t+k})_{k=2}^2$				×			×
N	64,948	49,631	38,132	13,272	52,258	38,132	13,272
R^2	0.481	0.484	0.497	0.543	0.501	0.498	0.545

Notes: Standard errors are two-way clustered by firm ID and industry-year. Columns 1 and 5 are “baseline estimates” corresponding, respectively, with columns 1 and 3 of Table 5.

our estimate of s would be biased upward, in particular.

To test for this important possibility, we augment our previous regressions to include controls for past and future fundamentals in the form of two leads and lags of real value added growth at the aggregate and sectoral levels as well as firm-level TFP growth. Specifically, for our aggregate specification Equation 20, we estimate

$$\begin{aligned} \text{opt}_{it} = & u \text{opt}_{i:t-1} + s \overline{\text{opt}}_{t-1} + \gamma_i + \\ & + \sum_{k=2}^2 (\eta_k^{\text{agg}} \Delta \log Y_{t+k} + \eta_k^{\text{ind}} \Delta \log Y_{j(i):t+k} + \eta_k^{\text{rm}} \Delta \log \theta_{i:t+k}) + \varepsilon_{it} \end{aligned} \quad (22)$$

We estimate an analogous specification at the industry level, but with the aggregate leads and lags absorbed. If staggered common positive shocks to the economy and sectors were driving some or all of the estimated spillovers, we would expect to find a severely attenuated estimate of the virality coefficient s . Even under our interpretation, future output growth could be a “bad control” that is caused by optimism and absorbs some of its effect.

We report our estimates of the virality coefficients in Table 6, adding the “bad controls” one at a time. In column 2 we find that instead of attenuating \hat{s} , controlling for past and future aggregate fundamentals in fact slightly increases the original point estimate reported in column 1 (within one standard error of the original value). In columns 3 and 4, when we additionally control for sectoral level value added growth and firm level TFP growth, the

point estimates drop slightly while standard errors increase significantly. Similarly, for our industry-level estimates, we find no statistically significant evidence of coefficient attenuation as additional controls are added (columns 5 to 7). Taken together, these estimates build confidence that our baseline virality results are not driven by staggered aggregate shocks.

Alternative Strategy: CEO Change Spillovers. To further build confidence that our spillover estimates capture virality, we leverage changes in within-sector and peer-set optimism induced by plausibly exogenous CEO changes as instruments for the level of optimism within these groups. Concretely, we construct an instrument equal to the contribution toward optimism from firms whose CEOs changed for a plausibly exogenous reason, or

$$\overline{\text{opt}}_{j(i);t}^{\text{ceo}} = \frac{1}{jM_{j(i);t}} \sum_{k \in M_{j(i);t}^c} \text{opt}_{k;t} \quad (23)$$

where $M_{j(i);t}$ is the set of firms in industry $j(i)$ at time t , and $M_{j(i);t}^c$ is the subset that had plausibly exogenous CEO changes. We construct the peer-set instrument $\overline{\text{opt}}_{\rho(i);t}^{\text{ceo}}$ analogously. We use $(\overline{\text{opt}}_{j(i);t}^{\text{ceo}}, \overline{\text{opt}}_{\rho(i);t}^{\text{ceo}})$ as instruments for $(\overline{\text{opt}}_{j(i);t}, \overline{\text{opt}}_{\rho(i);t})$ in the estimation of Equation 21. We present the corresponding estimates in Table A10. We find similar point estimates under IV and OLS, although the IV estimates are significantly noisier.

The Spread of Hiring-Relevant Narratives. We repeat the estimation of our equation measuring aggregate associativeness and virality, Equation 20, for the other thirteen narratives that are selected by our LASSO procedure as relevant for hiring. To allow for the greatest comparability with our estimates for optimism, we transform these narrative loadings into binary indicators for being above the sample median. We present our estimates of u , r , and s in the three panels of Appendix Figure A6. The circles are point estimates and the bars are 95% confidence intervals. We find significant evidence of stubbornness, or $u > 0$, in each case and significant evidence of virality, or $s > 0$, in all but two cases. We find some evidence of associativeness ($r \neq 0$) for certain narratives, with “Lease, Tenant, Landlord” (relating to real estate), “Debt, Credit, Facility” (relating to financial conditions and leverage), and “Reorganization, Bankruptcy, Plan” (relating to firm restructuring) having $r < 0$, and “Court, Settlement, District” (relating to legal proceedings), “Business, Public, Combination” (relating to firm origination), and “Technology, Revenue, Development” (relating to R&D) having $r > 0$. In Appendix Table A11, we instrument for optimism with the other 13 hiring-relevant narratives in the estimation of Equations 20 and 21 and find similar point estimates to our baseline analysis that are suggestive of increased virality.

5 A Narrative Business-Cycle Model

To study the implications of narratives for macroeconomic dynamics, we now specialize our abstract framework and develop a microfounded business-cycle model that embeds the two premises of the macroeconomics of narratives.

5.1 Technology and Preferences

The consumption, production, and labor supply side of the model is intentionally standard and is a purely real variant of the models expositied in Woodford (2003) and Galí (2008). Time is discrete and infinite, indexed by $t \geq \mathbb{N}$. There is a continuum of monopolistically competitive intermediate goods firms of unit measure, indexed by i , and uniformly distributed on the interval $[0, 1]$. Intermediate goods firms have idiosyncratic (Hicks-neutral) productivity θ_{it} . They hire labor L_{it} monopsonistically at wage w_{it} to produce a differentiated variety in quantity x_{it} that they sell at price p_{it} according to the production function:

$$x_{it} = \theta_{it} L_{it} \quad (24)$$

where $\alpha \geq (0, 1]$ describes returns to scale in production.

A final goods firm competitively produces aggregate output Y_t by using a constant elasticity of substitution (CES) production function

$$Y_t = \left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \quad (25)$$

where $\epsilon > 1$ is the elasticity of substitution between varieties.

A representative household consumes final goods C_t and supplies labor $fL_{it}g_{i2[0,1]}$ to the intermediate goods firms with isoelastic, separable, expected discounted utility preferences:

$$U(fC_t, fL_{it}g_{i2[0,1]}g_{t2\mathbb{N}}) = E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^1}{1-\gamma} - \int_{[0,1]} \frac{L_{it}^{1+\psi}}{1+\psi} di \right) \right] \quad (26)$$

where $\gamma \geq \mathbb{R}_+$ indexes the household's risk aversion and $\psi \geq \mathbb{R}_+$ indexes their labor supply elasticity to each firm.

Finally, we define the composite parameter:

$$\omega = \frac{1-\gamma}{1+\psi} + 1 \quad (27)$$

which indexes the strength of aggregate demand externalities (Blanchard and Kiyotaki, 1987) in generating strategic complementarities. So that complementarity is positive but not so extreme that the model features multiple equilibria, we assume that $\omega \geq [0, 1)$.

5.2 Narratives and Beliefs

Firm productivity θ_{it} is comprised of a common, aggregate component θ_t , an idiosyncratic time-invariant component γ_i , and an idiosyncratic time-varying component $\tilde{\theta}_{it}$:

$$\theta_{it} = \tilde{\theta}_{it}\gamma_i\theta_t \quad (28)$$

Firms know that $\log \gamma_i \sim N(\mu_\gamma, \sigma_\gamma^2)$, know their own value of γ_i , and believe that $\log \tilde{\theta}_{it} \sim N(0, \sigma_\theta^2)$ and independently and identically distributed (IID) across firms and time. Firms receive idiosyncratic Gaussian signals about $\log \theta_t$

$$s_{it} = \log \theta_t + \varepsilon_{it} \quad (29)$$

with $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ and IID across firms and time.

As in the conceptual framework from Section 2, narratives form a common factor structure of agents' prior beliefs about the aggregate component of productivity θ_t . To best fit our main empirical analysis, we suppose that there are two competing narratives: an optimistic narrative and a pessimistic narrative. According to each narrative, the aggregate component of productivity follows:

$$\log \theta_t \sim N(\mu, \sigma^2) \quad (30)$$

where $\mu = \mu_P$ under the pessimistic narrative and $\mu = \mu_O > \mu_P$ under the optimistic narrative. Both of these narratives are potentially misspecified, and the true distribution for fundamentals is given by H .

Firms either believe the optimistic narrative or the pessimistic narrative. Hence, each firm's prior belief regarding the fundamental can be described as:

$$\pi_{it}(\lambda_{it}) = N(\lambda_{it}\mu_O + (1 - \lambda_{it})\mu_P, \sigma^2) \quad (31)$$

where $\lambda_{it} \in [0, 1]$, $\lambda_{it} = 1$ corresponds to a firm believing in the optimistic narrative, and $\lambda_{it} = 0$ corresponds to a firm believing in the pessimistic narrative. We let $Q_t = \int_{[0,1]} \lambda_{it} di$ correspond to the fraction of optimists in the population.

5.3 Narrative Evolution

To describe the evolution of narratives, we need to describe two probabilities: the probability that optimists remain optimistic, P_O , and the probability that pessimists become optimistic, P_P . We specify that both probabilities depend on aggregate output Y_t and the fraction of optimists in the population Q_t . Hence, if aggregate output is Y_t and there are Q_t optimists in the population, the fraction of optimists in the following period is given by:

$$Q_{t+1} = Q_t P_O(\log Y_t, Q_t) + (1 - Q_t) P_P(\log Y_t, Q_t) \quad (32)$$

In view of our evidence that firms update associatively and virally, we assume that P_O and P_P are both increasing functions. As firms are stubborn, we assume that $P_O > P_P$. We assume that P_O and P_P are continuous and almost everywhere differentiable.

To illustrate our results, obtain closed-form expressions, and exactly match our regression evidence, we will often study the following updating probabilities:

Example 1 (Linear-Associative-Viral Updating Probabilities). The linear-associative-viral (LAV) specification for updating probabilities sets:

$$P_O(\log Y, Q) = [u + r \log Y + sQ]_0^1 \quad \text{and} \quad P_P(\log Y, Q) = [u + r \log Y + sQ]_0^1 \quad (33)$$

where $[z]_0^1 = \max\{\min\{z, 1\}, 0\}$ ensures that probabilities lie between zero and one, $u > 0$ indexes stubbornness, $r > 0$ indexes associativeness, and $s > 0$ indexes virality. 4

In Section 6.5, we compare the predictions of our model to two alternatives: a model in which agents are Bayesian and one in which agents are contrarian.

5.4 Equilibrium

An equilibrium is a path for all endogenous and exogenous variables:

$$E = \left\{ Y_t, C_t, Q_t, \theta_t, fL_{it}, x_{it}, p_{it}, w_{it}, \lambda_{it}, s_{it}, \tilde{\theta}_{it} g_{i2[0;1]} \right\}_{t \in \mathbb{N}} \quad (34)$$

such that (i) narrative weights λ_{it} follow the Markov process consistent with Equation 32 given Q_t and Y_t , (ii) x_{it} maximizes intermediate goods firms' expected profits given their narrative weights λ_{it} , signal s_{it} , and knowledge of E , (iii) L_{it} is consistent with production technology (Equation 24) given x_{it} and $\tilde{\theta}_{it}$, (iv) prices p_{it} are consistent with profit maximization by the final goods firm, (v) wages w_{it} clear the labor market for each firm (vi) Y_t aggregates intermediate good production according to Equation 25, (vii) C_t satisfies goods market clearing, $C_t = Y_t$, and (viii) Q_t evolves according to Equation 32.

6 Theoretical Results

We first characterize equilibrium dynamics in the model and establish that the model is identified conditional on both calibrating standard macroeconomic preference and production parameters and our regression estimates from Section 4. We then show three main results. First, despite the essential uniqueness of equilibrium, the presence of narratives can lead to multiple self-fulfilling steady states and hysteresis. Second, impulse responses to both fundamentals and narratives can feature both humps and discontinuities. Third, the economy endogenously cycles between regimes of high and low production, corresponding to times of aggregate optimism and pessimism. We show that none of these dynamic features are present under a Bayesian benchmark.

6.1 Characterizing Equilibrium Dynamics

We begin by characterizing the dynamics of the model as a single first-order nonlinear stochastic difference equation. The subsequent analysis will then study this equation to establish properties of business-cycle and narrative dynamics.

To solve for equilibrium production, it will suffice to solve for production by the intermediates goods firms. These firms face the problem of maximizing expected profits, as priced by the representative household:

$$\Pi = E_{it}[C_t (p_{it}x_{it} - w_{it}L_{it})] \quad (35)$$

where C_t is the (unnormalized) stochastic discount factor that converts the profits of the firm into their marginal value to the household. The intermediate goods firm acts as a monopolist in the product market and a monopsonist in the labor market.

We first solve for the demand curve faced by the intermediates goods firms. The final goods firm maximizes profits taking as given the prices set by intermediates goods firms. This implies the following constant-price-elasticity demand curve:

$$p_{it} = Y_t^{\frac{1}{\epsilon}} x_{it}^{\frac{1}{\epsilon}} \quad (36)$$

Increases in aggregate output shift out this demand curve via aggregate demand externalities.

Second, we solve for the wage schedule faced by the intermediate goods firm. When facing a wage w_{it} , the intratemporal Euler equation of the representative household implies that labor supply is given by:

$$L_{it} = w_{it}C_t \quad (37)$$

Third, given the production technology of the firm, when it commits to producing x_{it} , its implied labor input is given by:

$$L_{it} = \theta_{it}^{\frac{1}{\alpha}} x_{it}^{\frac{1}{\alpha}} \quad (38)$$

Finally, by imposing goods market clearing $C_t = Y_t$ and substituting Equations 36, 37, and 38 into Equation 35, we obtain that the intermediates goods firms solve the following profit maximization problems:

$$\max_{x_{it}} \mathbb{E}_{it} \left[Y_t \left(Y_t^{\frac{1}{\epsilon}} x_{it}^{1 - \frac{1}{\epsilon}} \quad Y_t \theta_{it}^{\frac{1+\psi}{\alpha}} x_{it}^{\frac{1+\psi}{\alpha}} \right) \right] \quad (39)$$

By the first-order condition of this program, we have that optimal production solves:

$$\left(1 \quad \frac{1}{\epsilon} \right) \mathbb{E}_{it} \left[Y_t^{\frac{1}{\epsilon}} \right] x_{it}^{\frac{1}{\epsilon}} = \frac{1+\psi}{\alpha} \mathbb{E}_{it} \left[\theta_{it}^{\frac{1+\psi}{\alpha}} \right] x_{it}^{\frac{1+\psi}{\alpha} - \frac{1}{\epsilon}} \quad (40)$$

where the left-hand side is the marginal expected revenue of the firm from expanding production and the right-hand side is the marginal expected cost of this expansion. We now take logarithms of all variables, and substitute this best reply into the production function of the final goods firm. From this, we obtain that the static equilibrium of the model is characterized by the solution to the following fixed-point equation:

$$\log Y_t = \frac{\epsilon}{\epsilon - 1} \log \mathbb{E}_t \left[\exp \left\{ \frac{-1}{1+} + 1 \left(\log \left(\frac{1}{1+} \right) \right. \right. \right. \quad (41)$$

$$\left. \left. \left. \log \mathbb{E}_{it} \left[\exp \left\{ \frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] + \log \mathbb{E}_{it} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) \log Y_t \right\} \right] \right) \right\} \right]$$

where the outer expectation operator integrates over the realizations of productivity shocks $(\tilde{\theta}_{it}, \gamma_t)$, narrative loadings λ_{it} , and signals s_{it} .

By employing a functional guess-and-verify argument, we obtain that the model has a unique *quasi-linear* equilibrium in which log output depends linearly on log aggregate productivity and non-linearly, but separably, on the fraction of optimists in the population:

Proposition 1 (Equilibrium Characterization). *There exists a unique equilibrium such that:*

$$\log Y(\log \theta_t, Q_t) = a_0 + a_1 \log \theta_t + f(Q_t) \quad (42)$$

for some coefficients a_0 and $a_1 > 0$, and a strictly increasing function f .

Proof. See Appendix A.2 □

Remark 1. This result claims uniqueness only within the quasi-linear class. As best replies and aggregation are non-linear and the spaces of actions and fundamentals are not compact, one cannot use classical arguments to ensure that the fixed point operator implicit in Equation 41 is a contraction. Nevertheless, in Appendix A.2, we show that there is a unique equilibrium when fundamentals are restricted to lie in a compact set (Lemma 2). Moreover, the claimed quasi-linear equilibrium is an ε equilibrium for any $\varepsilon > 0$ for some sufficiently large support for fundamentals (Lemma 3). Hence, in a formal sense, the quasi-linear equilibrium we study is the limit of the unique equilibrium with bounded fundamentals, justifying our restriction in analyzing this class of equilibrium. 4

The coefficient a_1 and function f describe how fundamentals and optimism drive aggregate output. In the proof of Proposition 1, we derive these objects as functions of the macroeconomic parameters $(\epsilon, \psi, \gamma, \alpha)$, the signal-to-noise ratio of agents' signals about productivity κ , and the extent of mean differences in the priors of optimists and pessimists $\mu_O - \mu_P$. The effect on output from going from full pessimism to full optimism is given by

$$f(1) = \frac{\alpha \delta^{OP}}{1 - \omega} \quad (43)$$

where δ^{OP} is the average partial equilibrium effect of a firm being optimistic on hiring, the returns-to-scale parameter α converts this into the effect on production, and $\frac{1}{1-\omega}$ is the general equilibrium multiplier of this effect.

The role of optimism in equilibrium has two subtle properties. First, the effect of optimism on output, $f(Q)$, is non-linear. The non-linearity arises from the fact that firms' heterogeneous priors induce heterogeneity in production conditional on productivity and hence also misallocation. Second, there is an equilibrium multiplier for optimism due to demand externalities. In particular, even a pessimistic firm will produce more if a large fraction of *other* firms is optimistic, as this optimism increases aggregate demand.

We can also derive the equilibrium equation describing firms' hiring decisions:

$$\log L_{it} = c_{0,i} + c_1 \log \theta_t + c_2 f(Q_t) + c_3 \log \theta_{it} + \delta^{OP} \lambda_{it} + \zeta_{it} \quad (44)$$

where ζ_{it} is an IID normal random variable with zero mean. This clarifies the exact interpretation of our regression model for hiring, Equation 12, in the model. The general-equilibrium effect of optimism on hiring, $c_2 f(Q_t)$, was absorbed in the regression equation as a fixed-effect. But our model structure allows us to construct these effects given knowledge of the partial equilibrium effect of optimism on hiring, δ^{OP} , and other structural parameters that govern general-equilibrium feedback. We formalize this below:

Corollary 1 (Identification of Model Parameters). *Conditional on $(\alpha, \epsilon, \gamma, \psi)$, δ^{OP} uniquely identifies f , the equilibrium effect of optimism on aggregate output.*

Proof. See Appendix A.4. □

Finally, we can now express the dynamics of the economy in terms of a first-order nonlinear stochastic difference equation for the fraction of optimists in the population:

Corollary 2 (Characterization of Dynamics). *Optimism evolves according to the following stochastic, nonlinear first-order difference equation $Q_{t+1} = T(Q_t, \log \theta_t)$, where*

$$T(Q_t, \log \theta_t) = Q_t P_O(a_0 + a_1 \log \theta_t + f(Q_t), Q_t) + (1 - Q_t) P_P(a_0 + a_1 \log \theta_t + f(Q_t), Q_t) \quad (45)$$

Proof. See Appendix A.5 □

6.2 Steady-State Multiplicity and Hysteresis

We first characterize the steady states of optimism and their stability. Let T be the equilibrium transition map from Corollary 2 and T_θ be the map for a fixed value of aggregate productivity. A level of optimism Q is a deterministic steady state for level of productivity θ if it is a fixed point of the corresponding map, $T_\theta(Q) = Q$. The following result establishes that a deterministic steady state always exists and provides necessary and sufficient conditions for extreme optimism and pessimism to be steady states.

Proposition 2 (Steady State Existence, Multiplicity, and Stability). *The following statements are true:*

1. *There exists a deterministic steady state level of optimism for every $\theta \geq \Theta$*
2. *There exist thresholds θ_P and θ_O such that: $Q = 0$ is a deterministic steady state for θ if and only if $\theta \leq \theta_P$ and $Q = 1$ is a deterministic steady state for θ if and only if $\theta \geq \theta_O$. Moreover, these thresholds are given by:*

$$\theta_P = \exp \left\{ \frac{P_P^{-1}(0; 0) - a_0}{a_1} \right\} \quad \text{and} \quad \theta_O = \exp \left\{ \frac{P_O^{-1}(1; 1) - a_0 - f(1)}{a_1} \right\} \quad (46)$$

where $P_P^{-1}(x; Q) = \sup \{ Y : P_P(Y, Q) = x \}$ and $P_O^{-1}(x; Q) = \inf \{ Y : P_O(Y, Q) = x \}$.²⁷

3. *Extreme pessimism is stable if $\theta < \theta_P$ and $P_O(P_P^{-1}(0; 0), 0) < 1$ and extreme optimism is stable if $\theta > \theta_O$ and $P_P(P_O^{-1}(1; 1), 1) > 0$.*

Proof. See Appendix A.6. □

²⁷With the convention that the infimum of an empty set is $+1$ and the supremum of an empty set is -1 .

This result establishes two important properties of narrative dynamics. First, it establishes conditions under which extreme optimism and pessimism can be stable steady states. These conditions can be checked with only a few parameters: the responsiveness of output to productivity a_1 , its baseline level a_0 , the impact of all agents being optimistic on output $f(1)$, the highest level of output such that all pessimists remain pessimistic when everyone is a pessimist $P_P^{-1}(0; 0)$, and the lowest level of output such that all optimists remain optimistic when all other agents are optimists $P_O^{-1}(1; 1)$.

Second, the result demonstrates the possibility for steady states of optimism and pessimism to be entirely self-fulfilling, in the sense that they can co-exist as stable steady states at a fixed level of fundamental productivity. Thus, differing initial conditions of narratives in the population can lead to differing steady state levels of narrative penetration and therefore output. The following corollary characterizes exactly when this can happen:

Corollary 3 (Characterization of Extremal Multiplicity). *Extreme optimism and pessimism are simultaneously deterministic steady states for θ if and only if $\theta \geq [\theta_O, \theta_P]$, which is non-empty if and only if*

$$P_O^{-1}(1; 1) > P_P^{-1}(0; 0) > f(1) \quad (47)$$

Proof. See Appendix A.7 □

To gain intuition for these results, we can compute these conditions parametrically in our running LAV example:

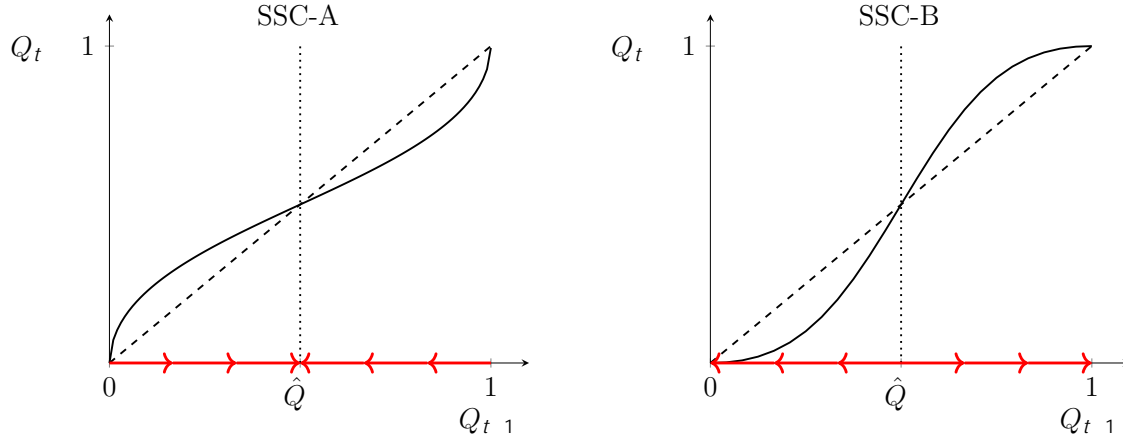
Example 1 (continuing from p. 29). In the LAV special case, we can compute the sufficient statistics for narrative updating analytically. In particular, we have that extreme optimism and pessimism can coexist if and only if:

$$1 > 2u + s + rf(1) \quad (48)$$

which is to say that stubbornness, associativeness, virality, and the equilibrium impact of optimism on output are sufficiently large. 4

To say more, we restrict attention to two important subclasses of updating rules that satisfy a natural single-crossing condition. We say that T is *strictly single-crossing from above* (SSC-A) if for all $\theta \geq \Theta$ there exists $\hat{Q} \in [0, 1]$ such that $T(Q) > Q$ for all $Q \in (0, \hat{Q})$ and $T(Q) < Q$ for all $Q \in (\hat{Q}, 1)$. We say that T is *strictly single-crossing from below* (SSC-B) if for all $\theta \geq \Theta$ there exists $\hat{Q} \in [0, 1]$ such that $T(Q) > Q$ for all $Q \in (\hat{Q}, 1)$ and $T(Q) < Q$ for all $Q \in (0, \hat{Q})$. If T is either SSC-A or SSC-B, we say that it is SSC. The left and right panels of Figure 3 illustrate examples of SSC-A and SSC-B transition maps as black solid lines.

Figure 3: Illustration: Steady States and Dynamics Under the SSC Property



Notes: In each subfigure, the solid line is the transition map T , the dashed line is the 45 degree line, the dotted line indicates the interior steady state \hat{Q} ($1/2$ in these examples), and the red arrows indicate the dynamics. The subfigures respectively correspond to SSC-A (“strict single crossing from above”) and SSC-B (“strict single crossing from below”), as defined in the text.

Lemma 1 (Steady States under the SSC Property). *If T is SSC, then there exist at most three deterministic steady states. These correspond to extreme pessimism $Q = 0$, extreme optimism $Q = 1$, and intermediate optimism $Q = \hat{Q}$. Moreover, when T is SSC-A: intermediate optimism is stable with a basin of attraction that includes $(0, 1)$; and whenever extreme optimism or extreme pessimism are steady states that do not coincide with \hat{Q} , they are unstable with respective basins of attraction $(0, \hat{Q})$ and $(\hat{Q}, 1)$. When T is SSC-B: whenever extreme optimism is a steady state, it is stable with basin of attraction $(\hat{Q}, 1]$; whenever extreme pessimism is a steady state it is stable with basin of attraction $[0, \hat{Q})$; and intermediate optimism is always unstable with basin of attraction $(\hat{Q}, 1)$.*

Proof. See Appendix A.8 □

In the SSC-A case there is a unique, (almost) globally stable steady state (left panel of Figure 3). In the SSC-B class, there exists a state-dependent criticality threshold $\hat{Q} \in [0, 1]$, below which the economy converges to extreme, self-fulfilling pessimism and above which the economy converges to extreme, self-fulfilling optimism (right panel of Figure 3). These two classes delineate two qualitatively different regimes for narrative dynamics: one with stable narrative convergence around a long-run steady state (SSC-A) and one with a strong role for initial conditions and hysteresis (SSC-B).

To gain insights into the determinants of the criticality threshold for optimism to “go viral,” we study our running LAV example:

Example 1 (continuing from p. 29). In the LAV class, when $\theta \geq (\theta_O, \theta_P)$ the criticality threshold is given by the unique solution to the equation:

$$\hat{Q} = u(2\hat{Q} - 1) + s\hat{Q} + r(a_0 + a_1 \log \theta + f(\hat{Q})) \quad (49)$$

Thus, under the approximation that $f(Q) = kQ$, we have that:

$$\hat{Q} = \frac{u - r(a_0 + a_1 \log \theta)}{2u + s + rk} \quad (50)$$

Hence, higher fundamentals, greater virality of optimism, greater effects of optimism, and greater associativeness lower the criticality threshold and make it easier for an epidemic of extreme optimism to take hold. 4

For the remainder of the analysis, we restrict attention to the SSC class, noting that (in view of our equilibrium characterization) this is an assumption solely on primitives.²⁸

6.3 Hump-shaped and Discontinuous Impulse Responses

We now study the impulse propagation mechanisms at work in the economy. We consider the responses of aggregate output and optimism in the economy to a one time positive shock to fundamentals from a steady state corresponding to $\theta = 1$:

$$\theta_t = \begin{cases} 1, & t = 0, \\ \hat{\theta}, & t = 1, \\ 1, & t \geq 2. \end{cases} \quad (51)$$

where $\hat{\theta} > 1$. We would like to understand when the impulse response to a one-time shock is *hump-shaped*, meaning that there exists a $\hat{t} \geq 2$ such that Y_t is increasing for $t \leq \hat{t}$ and decreasing thereafter. Moreover, we would like to understand how big a shock needs to be to send the economy from one steady state to another, as manifested as a discontinuity in the IRFs in the shock size $\hat{\theta}$.

In the SSC-A case, IRFs are continuous in the shock but can nevertheless display hump-shaped dynamics as a result of the endogenous evolution of optimism.

²⁸This is without a substantive loss of generality as we can always represent any non-SSC T_θ as the concatenation of a set of restricted functions that are SSC on their respective domains. Concretely, whenever T_θ is not SSC, we can represent its domain $[0, 1]$ as a collection of intervals I_j such that $\bigcup_{j \in J} I_j = [0, 1]$ and the restricted functions $T_{\theta,j} : I_j \rightarrow [0, 1]$ defined by the property that $T_{\theta,j}(Q) = T_\theta(Q)$ for all $Q \in I_j$ are either SSC-A or SSC-B for all $j \in J$. Thus, applying our results to these restricted functions, we have a complete description of the global dynamics.

Proposition 3 (SSC-A Impulse Response Functions). *In the SSC-A case, suppose that $Q_0 = \hat{Q}_1 \geq (0, 1)$. The impulse response to a one-time fundamental shock is given by:*

$$\log Y_t = \begin{cases} a_0 + f(\hat{Q}_1), & t = 0, \\ a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \quad Q_t = \begin{cases} \hat{Q}_1, & t = 1, \\ Q_2, & t = 2, \\ T_1(Q_{t-1}), & t \geq 3. \end{cases} \quad (52)$$

Moreover, $Q_2 = \hat{Q}_1 P_O(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1) + (1 - \hat{Q}_1) P_P(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1) > \hat{Q}_1$, Q_t is monotonically declining for all $t \geq 2$, and $Q_t \neq \hat{Q}_1$. The IRF is hump-shaped if and only if $\hat{\theta} < \exp \left\{ \frac{f(\hat{Q}_2) - f(\hat{Q}_1)}{a_1} \right\}$.

Proof. See Appendix A.9 □

All persistence in the IRF of output derives from persistence in the IRF of optimism. There is a hump in the IRF for output if the boom induced by optimism exceeds the direct effect of the shock.

This contrasts sharply from the SSC-B case, wherein impulse responses can be discontinuous in the shock size. To reduce repetition, the following proposition characterizes the IRFs from the pessimistic steady state; those from the optimistic steady state are analogous.

Proposition 4 (SSC-B Impulse Response Functions). *In the SSC-B case, suppose that $\theta_O < 1 < \theta_P$ and that $Q_0 = 0$. The impulse response of the economy to a one-time fundamental shock is given by:*

$$\log Y_t = \begin{cases} a_0, & t = 0, \\ a_0 + a_1 \log \hat{\theta}, & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \quad Q_t = \begin{cases} 0, & t = 1, \\ P_P(a_0 + a_1 \log \hat{\theta}, 0), & t = 2, \\ T_1(Q_{t-1}), & t \geq 3. \end{cases} \quad (53)$$

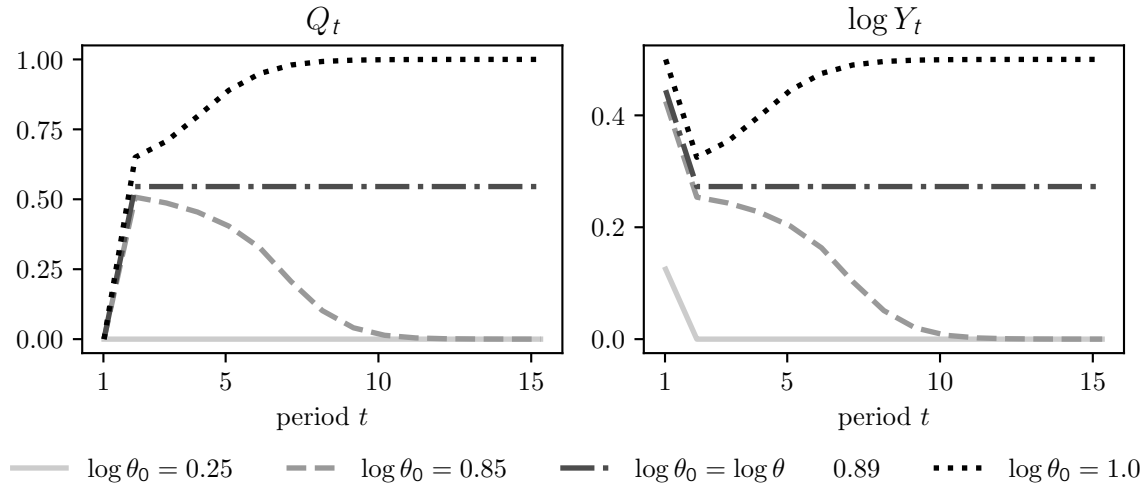
These impulse responses fall into the following four exhaustive cases:

1. $\hat{\theta} < \theta_P$, No Lift-O : $Q_t = 0$ for all $t \geq 1$.
2. $\hat{\theta} \geq (\theta_P, \theta)$, Transitory Impact: Q_t is monotonically declining for all $t \geq 2$ and $Q_t \neq 0$.
3. $\hat{\theta} = \theta$, Permanent (Knife-edge) Impact: $Q_t = \hat{Q}_1$ for all $t \geq 1$
4. $\hat{\theta} > \theta$, Permanent Impact, : Q_t is monotonically increasing for all $t \geq 2$ and $Q_t \neq 1$

where the critical shock threshold is $\theta = \exp \left\{ \frac{P_P^{-1}(\hat{Q}_1; 0) - a_0}{a_1} \right\} > \theta_P$. In the case of transitory impact, the output IRF is hump-shaped if and only if $\hat{\theta} < \exp \left\{ \frac{f(P_P(a_0 + a_1 \log \hat{\theta}, 0))}{a_1} \right\}$.

Proof. See Appendix A.10. □

Figure 4: Illustration of Deterministic IRFs in SSC-B Case



Notes: The plots show the deterministic impulse responses of Q_t and $\log Y_t$ in an illustrative model with LAV updating. The four initial conditions correspond to the four cases of Proposition 4.

To understand this result, we first inspect the IRFs. At time $t = 0$, the economy lies at a steady state of extreme pessimism with $\log \theta_0 = 0$ and so $\log Y_0 = a_0$. At time $t = 1$, the one-time productivity shock takes place and output jumps up to $\log Y_1 = a_0 + a_1 \log \hat{\theta}$ as everyone remains pessimistic. At time $t = 2$, agents observe that output rose in the previous period. As a result, a fraction $P_P(\log Y_1, 0)$ of the population becomes optimistic. For output, the one-time productivity shock has dissipated, so output is now given by its unshocked baseline a_0 plus the equilibrium output effect of optimism $f(Q_2)$. From this point, the IRF evolves deterministically and its long-run behavior depends solely on whether the fraction that became initially optimistic exceeds the criticality threshold \hat{Q}_1 that delineates the basins of attraction of the steady states of extreme optimism and extreme pessimism.

As a result, productivity shocks have the potential for the following four qualitatively distinct effects, described in Proposition 4 and illustrated numerically in Figure 4. First, if a shock is small and no agent is moved toward optimism, the shock has a one-period impact on aggregate output. Second, if some agents are moved to optimism by the transitory boost to output but this fraction lies below the criticality threshold, then output steadily declines back to its pessimistic steady state level as optimism was not sufficiently great to be self-fulfilling. Third, in the knife-edge case, optimism moves to a new (unstable) steady-state and permanently increases output. Fourth, when enough agents are moved to optimism by the initial boost to output, then the economy converges to the fully optimistic steady state and optimism is completely self-fulfilling.

6.4 Boom-Bust Cycles

Having characterized the deterministic impulse propagation mechanisms at work in the economy, we now turn to understanding the stochastic properties of the path of the economy.

To this end, we analytically study the period of boom and bust cycles: the expected time that it takes for the economy to move from a state of extreme pessimism to a state of extreme optimism, and *vice versa*. Formally, define these expected stopping times as:

$$T_{PO} = E_H [\min \{ \tau \in \mathbb{N} : Q = 1 \mid Q_0 = 0 \}], \quad T_{OP} = E_H [\min \{ \tau \in \mathbb{N} : Q = 0 \mid Q_0 = 1 \}] \quad (54)$$

where the expectation is taken under the true data generating process for the aggregate component of productivity H , which may or may not coincide with one of the narratives under consideration.

The following result provides sharp upper bounds, in the sense that they are attained for some H , on these stopping times as a function of deep structural parameters:

Proposition 5 (Period of Boom-Bust Cycles). *The expected regime-switching times satisfy the following inequalities:*

$$\begin{aligned} T_{PO} & \leq \frac{1}{H \left(\exp \left\{ \frac{P_P^y(1;0) - a_0}{a_1} \right\} \right)} \\ T_{OP} & \leq \frac{1}{H \left(\exp \left\{ \frac{P_O^y(0;1) - a_0 - f(1)}{a_1} \right\} \right)} \end{aligned} \quad (55)$$

where $P_P^y(x; Q) = \inf \{ Y : P_P(Y, Q) = x \}$ and $P_O^y(x; Q) = \sup \{ Y : P_O(Y, Q) = x \}$. Moreover, when $P_O^y(0;1) = P_P^y(1;0) = f(1)$, these bounds are tight in the sense that they are attained for some H .

Proof. See Appendix A.11 □

This result establishes that the economy regularly oscillates between times of booms and busts. We establish this result by postulating fictitious processes for optimism and showing that they bound, path-by-path, the true optimism process. This enables us to construct stopping times that dominate the true stopping times in the sense of first-order stochastic dominance and have expectations that can be computed analytically, thus providing the claimed bounds. We establish that these bounds are tight by constructing a family of distributions H such that the fictitious processes coincide always with the true processes.²⁹

²⁹We moreover show that elements of this family can be attained by taking the limit of normal mixtures with sufficiently dispersed means. Thus, for sufficiently dispersed μ_O and μ_P , we can therefore construct H

We can provide insights into the determinants of the period of boom-bust cycles from these analytical bounds. Concretely, consider the bound on the expected time to reach a bust from a boom. This bound is small when the quantity $H\left(\exp\left\{\frac{P_O^Y(0;1) a_0 f(1)}{a_1}\right\}\right)$ is large, which happens when there is a fat left tail of fundamentals, when it is relatively easier for optimists to switch to pessimism as measured by $P_O^Y(0;1)$, and when co-ordination motives are weak as measured by $f(1)$.

6.5 Comparison to Alternative Models

We conclude our theoretical analysis by comparing our theoretical results with those that would be obtained under a Bayesian benchmark and a setting with contrarian agents.

The Bayesian Benchmark. Consider an alternative model in which each agent i initially believes the optimistic model is correct with probability $\lambda_{i0} \in (0, 1)$, and subsequently updates this probability by observing aggregate output and applying Bayes' rule under rational expectations. Formally, this corresponds to the following law of motion for Q_t :

$$Q_{t+1} = \int_{[0,1]} P_i[\mu = \mu_{Oj} | \log Y_j g_{j=0}^t] di \quad (56)$$

where $P_i[\mu = \mu_{Oj}] = \lambda_{i0}$ for some $\lambda_{i0} \in (0, 1)$ for all $i \in [0, 1]$, and conditional probabilities are computed under rational expectations with knowledge of $f \lambda_{i0} g_{i \in [0,1]}$. We define the log-odds ratio of an agent's belief as $\Omega_{it} = \log \frac{\mu_{Oj}}{1-\mu_{Oj}}$. The following Lemma characterizes the dynamics of agents' subjective models under the Bayesian benchmark:

Proposition 6 (Dynamics under the Bayesian Benchmark). *Each agent's log-odds ratio follows a random walk with drift, $\Delta \Omega_{i,t+1} = a + \xi_t$, where $a = E_H \left[\frac{(\log \theta - p)^2 - (\log \theta - o)^2}{2} \right]$ and ξ_t is an IID, mean-zero random variable. The economy converges almost surely to either extreme optimism ($a > 0$) or extreme pessimism ($a < 0$). The dynamics of the economy are asymptotically described by:*

$$\log Y_t = \begin{cases} a_0 + a_1 \log \theta_t, & a < 0, \\ a_0 + a_1 \log \theta_t + f(1), & a > 0. \end{cases} \quad (57)$$

Thus, the economy does not feature steady state multiplicity, hump-shaped or discontinuous IRFs, or the possibility for boom-bust cycles.

Proof. See Appendix A.12. □

for which the bound is attained by taking weighted averages of the optimistic and pessimistic narratives and making the uncertainty under each sufficiently small.

The optimist fraction Q converges to either 0 or 1 in the long run because one model is unambiguously better-fitting, and this will be revealed with infinite data.³⁰ Thus, interesting (and empirically realistic) dynamics of optimism are only possible in this model away from the simple Bayesian benchmark.

Contrarianism, Endogenous Cycles, and Chaos. The baseline model can generate neither endogenous cycles nor chaotic dynamics without extrinsic shocks to fundamentals (as made formal by Lemma 1). This is because the probability that agents become optimistic is always increasing in the fraction of optimists in equilibrium. In this section, we relax this assumption and delineate the precise, testable conditions under which cyclical and chaotic dynamics occur.

We begin by defining cycles and chaos. There exists a cycle of period $k \geq \mathbb{N}$ if $Q = T^k(Q)$ and all elements of $\{Q, T(Q), \dots, T^{k-1}(Q)\}$ are non-equal. We will say that there are *chaotic dynamics* if there exists an uncountable set of points $S \subset [0, 1]$ such that (i) for every $Q, Q^\theta \in S$ such that $Q \neq Q^\theta$, we have that $\limsup_{t \rightarrow \infty} |T^t(Q) - T^t(Q^\theta)| > 0$ and $\liminf_{t \rightarrow \infty} |T^t(Q) - T^t(Q^\theta)| = 0$ and (ii) for every $Q \in S$ and periodic point $Q^\theta \in [0, 1]$, $\limsup_{t \rightarrow \infty} |T^t(Q) - T^t(Q^\theta)| > 0$. This definition of chaos is due to Li and Yorke (1975) and can be understood as saying that there is a large set of points such that the iterated dynamics starting from any two points in this set get both far apart and vanishingly close.

We will study the issue of cycles and chaos under the simplifying assumption that,³¹ in equilibrium, the induced probabilities that optimists and pessimists respectively become optimists are quadratic and given by:³²

$$\tilde{P}_O(Q) = a_O + b_O Q - cQ^2, \tilde{P}_P(Q) = a_P + b_P Q - cQ^2 \quad (58)$$

with parameters $(a_O, a_P, b_O, b_P, c) \in \mathbb{R}^5$ such that $P_O([0, 1]), P_P([0, 1]) \subset [0, 1]$. The parameters a_O and a_P index stubbornness, b_O and b_P capture both virality and associativeness (through the subsumed equilibrium map), and c captures any non-linearity.

The following result describes the potential dynamics:

³⁰In the proof, we also explicitly derive the evolution of the log-odds ratio favoring the “correct” model. The log-odds ratio converges linearly and so the odds ratio in favor of the better fitting model converges exponentially quickly.

³¹This simplifying assumption is without any qualitative loss as this model can demonstrate the full range of potential cyclical and chaotic dynamics.

³²This can be microfounded in a generalization our earlier LAV example (Example 1) by taking $P_i(\log Y, Q) = u_i + r_i \log Y + s_i Q - cQ^2$ for $i \in \{O, P\}$ and approximating $f(Q) = \frac{\alpha \delta^{OP}}{1 - \omega} Q$. In this case:

$$\tilde{P}_i(Q) = (u_i + r_i a_0 + r_i a_1 \log \theta) + \left(r_i \frac{\alpha \delta^{OP}}{1 - \omega} + s_i \right) Q - cQ^2$$

Proposition 7. *The following statements are true:*

1. *When $\tilde{P}_O = \tilde{P}_P$ and both are monotone, there are neither cycles of any period nor chaotic dynamics.*
2. *When \tilde{P}_O and \tilde{P}_P are linear, cycles of period 2 are possible, cycles of any period $k > 2$ are not possible, and chaotic dynamics are not possible.*
3. *Without further restrictions on \tilde{P}_O and \tilde{P}_P , cycles of any period $k \in \mathbb{N}$ and chaotic dynamics are possible.*

Proof. See Appendix A.13 □

The proof of this result follows a classic approach of recasting a quadratic difference equation as a logistic difference equation via topological conjugacy (see, *e.g.*, Battaglini, 2021; Deng, Khan, and Mitra, 2022). The restrictions on structural parameters implied by the hypotheses of the proposition then yield upper bounds on the possible logistic maps and allow us to characterize the possible dynamics using known results.

To understand this result, observe in our baseline case in which T is monotone that cycles and chaos are not possible. This is because there is no potential for optimism to sufficiently overshoot its steady state. By contrast, when \tilde{P}_O and \tilde{P}_P are either non-monotone or non-ranked, two-period cycles can take place where the economy undergoes endogenous boom-bust cycles with periods of high optimism and high output ushering in periods of low optimism and low output (and *vice versa*) as contrarians switch positions and consistently overshoot the (unstable) steady state. Finally, when \tilde{P}_O and \tilde{P}_P are non-linear and non-monotone, essentially any richness of dynamics can be achieved via erratic movements in optimism that are extremely sensitive to initial conditions.

Thus, modelling endogenous narratives with contrarianism complements the literature on endogenous cycles in macroeconomic models (see, *e.g.*, Boldrin and Woodford, 1990; Beaudry, Galizia, and Portier, 2020) by providing a further potential micro-foundation for the existence of endogenous cycles.

6.6 Extensions

Welfare Implications. So far, we have studied the positive implications of fluctuations in optimism. In Appendix B.1, we study the normative implications of optimism and provide conditions under which its presence is welfare improving, despite it being misspecified. Intuitively, optimism acts as an *ad valorem* price subsidy for firms, which induces firms to hire more and can undo distortions caused by market power.

Multi-dimensional Narratives and Persistent Fundamentals. In the conceptual framework and our measurement, we allowed for a general set of narratives. However, in our main theoretical analysis, we restricted attention to two different narratives where agents differ only in their optimism. In Appendix B.2, we extend the baseline model to allow for arbitrarily many narratives regarding the mean, persistence, and volatility of fundamentals, which is essentially exhaustive within the Gaussian class. We characterize quasi-linear equilibrium in this richer setting and show how past values of fundamentals interact non-linearly with the distribution of narratives in the population to determine aggregate output.

Persistent Idiosyncratic Shocks and Narrative Updating. In our empirical analysis, we found that firms that experience positive idiosyncratic shocks are more likely to become optimistic. In Appendix B.3, we extend the multi-dimensional narrative analysis to allow for persistent idiosyncratic states and updating that depends on realized idiosyncratic states. When idiosyncratic shocks are fully transitory, this is of no consequence and our equilibrium characterization is identical. However, when idiosyncratic shocks are persistent, the fact that narrative updating depends on idiosyncratic shock realizations induces statistical dependence between an agent’s narrative and their idiosyncratic productivity state. The quasi-linear equilibrium characterization follows closely from that of the multi-dimensional case with aggregate persistence, where the joint distribution of narratives and states replaces the marginal distribution of narratives.

Narratives in Games and the Role of Higher-Order Beliefs. We have studied narratively driven fluctuations in a business-cycle model, but our insights apply to co-ordination games much more generally. In Appendix B.4, we study viral narratives in beauty contests [Morris and Shin \(2002\)](#), in which agents’ best replies are a linear function of their expectations of fundamentals and the average actions of others. Many models of aggregative games in macroeconomics and finance can be recast as such games when (log-)linearized (for a review, see [Angeletos and Lian, 2016](#)). We characterize equilibrium in this context and show how optimism percolates through the hierarchy of higher-order beliefs about fundamentals.

7 Quantifying the Impact of Narratives

We now combine our model and empirical results to gauge the quantitative effects of narratives on business cycles. We find that optimism explains 32% of the reduction in GDP over the Early 2000s Recession, 18% over the Great Recession and, more generally, 66% of the medium-run (two-year) variation in output. Thus, according to our analysis, viral optimism explains a significant fraction of the U.S. business cycle.

Table 7: Model Calibration

Fixed	ϵ	Elasticity of substitution	2.6
	γ	Income effects in labor supply	0
	ψ	Frisch elasticity	0.4
	α	Returns to scale	1
Calibrated	μ_O	μ_P Belief effect of optimism	0.028
	κ	Signal-to-noise ratio	0.344
	ρ	Persistence of productivity	0.086
	σ	Std. dev. of the productivity innovation	0.011
	u	Stubbornness	0.104
	r	Associativeness	0.804
	s	Virality	0.290

Notes: “Fixed” parameters are externally set. “Calibrated” parameters are chosen to hit various moments. Both are described in the text.

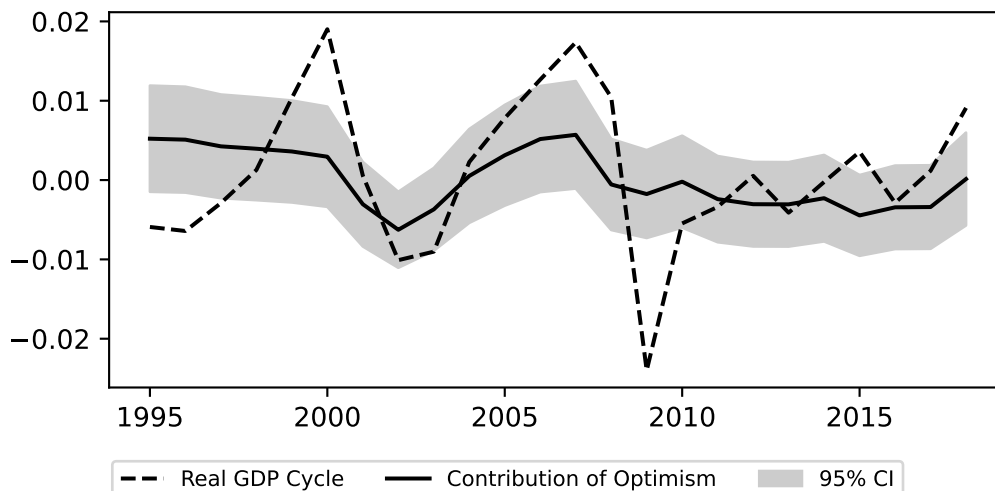
7.1 Calibrating the Model

To obtain numerical predictions from the model, we need to know (i) the static relationship between output and optimism; (ii) the updating probabilities for optimists and pessimists; and (iii) the data-generating process for fundamentals. We provide the model calibration in Table 7 and additional details in Online Appendix E.

First, we choose standard external calibrations of the macroeconomic preference and production parameters. We impose that intermediate goods firms have constant returns to scale or $\alpha = 1$, which has been argued by Basu and Fernald (1997), Foster, Haltiwanger, and Syverson (2008), and Flynn, Traina, and Gandhi (2019) to be a reasonable assumption for large U.S. firms. Second, as noted by Angeletos and La’O (2010), γ indexes wealth effects in labor supply, which are empirically very small (Cesarini, Lindqvist, Notowidigdo, and Östling, 2017). Hence, we set $\gamma = 0$ for our benchmark calibration. Third, we calibrate the Frisch elasticity of labor supply $\psi = 0.40$, corresponding to the central estimate in the review of estimates by Reichling and Whalen (2012). Finally, we calibrate the elasticity of substitution to match estimated markups from De Loecker, Eeckhout, and Unger (2020) of 60%, which implies that $\epsilon = 2.6$. Hence, altogether, this calibration implies an aggregate degree of strategic complementarity of $\omega = 0.49$.

Second, we calibrate f , the function governing the effect of optimism on aggregate output. As observed in Corollary 1, this requires only an estimate of the partial equilibrium effect of optimism on hiring, δ^{OP} , once the macroeconomic parameters have been chosen. We use our baseline estimate of $\hat{\delta}^{OP} = 0.0355$ (see Table 1). Moreover, this calibration places one

Figure 5: The Effect of Optimism on Output: Time Series



Notes: The “Real GDP Cycle” is calculated from a [Baxter and King \(1999\)](#) band-pass filter capturing periods between 6 and 32 quarters. The “Contribution of Optimism” is the model-implied effect of optimism on the log deviation of output from steady state. The 95% confidence interval incorporates uncertainty from the calibration of δ^{OP} using the delta method.

restriction on the parameters κ and $\mu_O - \mu_P$, respectively the signal-to-noise ratio and the impact of optimism on agents’ prior means.

Third, we calibrate the process for updating probabilities. We use our regression estimates of the LAV form (see Equation 33) which corresponds to the linear probability model (see Table 5).³³ This yields values of $u = 0.104$ for stubbornness (as $u_{opt} = 2u$), $r = 0.804$ for associativeness, and $s = 0.290$ for virality.

Fourth, and finally, we calibrate the process for fundamentals. To allow for persistence in both fundamentals as well as any unmodelled factors, we calibrate a case of the model with persistent fundamentals based on the analysis in Online Appendix B.2. Concretely, we suppose that $\log \theta_t$ is a Gaussian AR(1) process with persistence ρ and IID innovations $u_t \sim N(0, \sigma^2)$:

$$\log \theta_t = \rho \log \theta_{t-1} + u_t \quad (59)$$

To obtain the law of motion of aggregate output, we require three parameters (ρ, σ, κ) . We calibrate these to match the properties of *fundamental output*, defined as

$$\log Y_t^f = \log Y_t - f(Q_t) \quad (60)$$

³³This is with the caveat that our regression model does not necessarily yield probabilities between zero and one.

Table 8: The Effect of Optimism on US Recessions

Period	Change (%)		% Explained
	Detrended GDP	Optimism Component $f(Q_t)$	
2000-2002	-2.91	-0.92 (0.08)	31.65 (2.68)
2007-2009	-4.13	-0.75 (0.06)	18.06 (1.53)

Notes: Standard errors derive from the standard error from estimating δ^{OP} and are calculated using the delta method.

In Online Appendix E, we show that $\log Y_t^f$ follows an ARMA(1,1) with white noise process u_t . To calculate $\log Y_t^f$ in the data, we take $\log Y_t$ as band-pass filtered U.S. real GDP (Baxter and King, 1999), Q_t as our measured time series of aggregate optimism (see Figure 1), and f as our calibrated function.³⁴ We estimate by maximum-likelihood the ARMA(1,1) process for Y_t^f and then set (ρ, σ, κ) to exactly match the three estimated ARMA parameters. Upon obtaining κ , the restriction on κ and $\mu_O = \mu_P$ imposed by δ^{OP} yields the value of $\mu_O = \mu_P$.

7.2 The Effect of Optimism on U.S. GDP

We first use our calibration to estimate the contribution of measured aggregate optimism to business cycles. To do this, we take our measured aggregate optimism time series and compute its output effects via our calibrated f function. These estimates rely solely on our estimated partial equilibrium response of firms to optimism (reported in Table 1), the calibrated macroeconomic multiplier (from Step 1 of the previous subsection), and our measured aggregate optimism time series (reported in Figure 1).

In Figure 5, we plot the cyclical component of real GDP (dashed line) and the contribution of measured optimism toward output according to our model (solid line with grey 95% confidence interval). We observe that cyclical optimism explains a meaningful portion of fluctuations, particularly the booms of the mid-1990s and the mid-2000s and the busts of 2000-2002 and 2007-2009. Over each downturn, we can calculate the percentage of output reduction explained by the dynamics of optimism as

$$\% \text{ Explained} = 100 \frac{f(Q_{t_1}) - f(Q_{t_0})}{\log Y_{t_1} - \log Y_{t_0}} \quad (61)$$

³⁴We apply the Baxter and King (1999) band-pass filter to post-war quarterly U.S. real GDP data (Q1 1947 to Q1 2022). We use a lead-lag length of 12 quarters, a low period of 6 quarters, and a high period of 32 quarters. We then average these data to the annual level.

Table 9: Autocovariance Decomposition in the Calibrated Model

	Lag ℓ (years)		
	0	1	2
Autocovariance	0.75	0.31	0.07
Fraction $\hat{c}_Q(\ell)$	4.7%	17.1%	66.0%

Notes: Autocovariance is multiplied by 100^2 for ease of reading, so output is in units of 100 times log points.

where \hat{Q} is measured optimism, \hat{y} is the measured cyclical component of log real GDP, and (t_0, t_1) are the endpoints. We report these results in Table 8. The decline in the optimism component of GDP explains 31.65% of the output loss between 2000 and 2002, and 18.06% of the output loss between 2007 and 2009. These results imply that viral optimism is an important driver of economic fluctuations over the studied period.

7.3 Optimism and Economic Fluctuations

We next investigate more systematically how viral optimism contributes to the short-run dynamics of GDP in our model, using all information on viral spread and the estimated process for fundamentals. To this end, we calculate the percentage of the autocovariance of output at lag ℓ explained by optimism as:

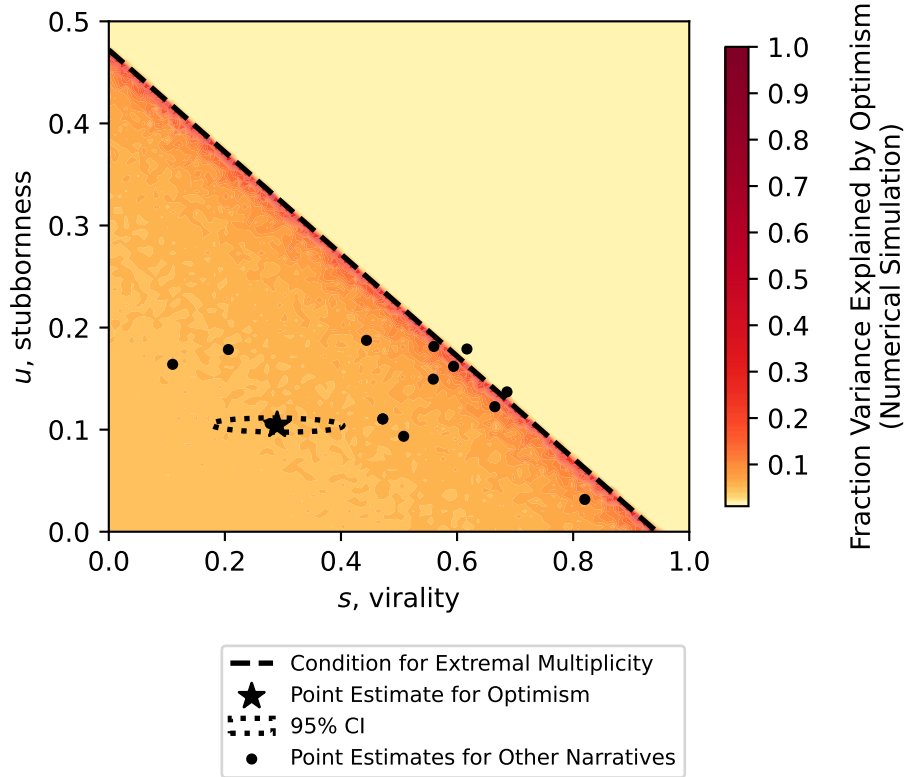
$$\hat{c}_Q(\ell) := 100 \frac{\text{Cov}[\log Y_t, \log Y_{t-\ell}^f] - \text{Cov}[\log Y_t^f, \log Y_{t-\ell}^f]}{\text{Cov}[\log Y_t, \log Y_{t-\ell}]} \quad (62)$$

We report our findings in Table 9. Optimism explains 4.9% of contemporary variance ($\ell = 0$), but this fraction increases with the lag. At one-year and two-year lags, optimism explains 17% and 66% of output covariance, respectively. Thus, most medium-frequency (two-year) dynamics are produced by viral optimism instead of fundamentals.

Counterfactual Effects Under Different Virality and Associativeness. We now study how our predictions depend on the extent of virality and stubbornness, two parameters that are somewhat imprecisely estimated. In Figure 6, we plot $\hat{c}_Q(0)$, the fraction of variance explained by optimism, for a range of alternative values of these parameters. We plot our point estimate of virality and stubbornness as a plus and its 95% confidence interval as a dotted ellipse. Within our confidence interval, we find that our baseline finding for $\hat{c}_Q(0)$ is numerically stable.

However, increases in virality and stubbornness sharply increase the variance effect of op-

Figure 6: Variance Decomposition for Different Values of Stubbornness and Virality



Notes: Calculations vary u and s , holding fixed all other parameters at their calibrated values. The shaded contours correspond to the fraction of variance explained by optimism, or $\hat{c}_Q(0)$ defined in Equation 62. The star is our calibrated value from Table 7, and the dotted lines are a 95% confidence set. The dots are calibrated values for other narratives from Figure A6. The dashed line is the condition of extremal multiplicity from Corollary 3.

timism as we approach the condition that delimits the SSC-A and SSC-B cases (as described theoretically by Equation 48 and plotted as the dashed black line). This is because, when this condition is close to holding or failing, all steady states are close to being unstable, so optimism is volatile. When the condition holds strongly, optimism explains essentially zero output variance in our simulations. This is because the economy (as a function of its initial history and random shocks) quickly settles into one of the extremal steady states, which are highly stable and very hard to leave.

Is this behavior economically unreasonable? While we can empirically reject such high stubbornness and virality for optimism, we cannot for our other estimated narratives. To emphasize this point, we plot our point estimates of (s, u) for the thirteen hiring-relevant narratives as black dots in Figure 6. Several of the thirteen plotted points are close to the condition for extremal multiplicity, and two are across it. Thus, while optimism features

dynamics well described by the SSC-A case, other narratives appear to be described by the SSC-B case and are therefore likely to feature hysteresis, among other phenomena.

Additional Sensitivity Analysis. We now consider how our calibration of macroeconomic parameters matters for our findings. Recall that $f(Q) = \frac{\delta^{OP}}{1-\beta} Q$, where $\frac{1}{1-\beta}$ is the general equilibrium demand multiplier in our economy, α indexes the returns to scale, and δ^{OP} is the partial equilibrium effect of optimism on hiring. Our baseline calibration implies a multiplier of 1.96, somewhat higher than what is implied by structural vector autoregression models which yield government spending multipliers between 1.0 and 1.3 (Caldara and Kamps, 2017) or estimated semi-structural models of demand multipliers which yield multipliers between 1.0 and 1.6 (Flynn, Patterson, and Sturm, 2022). In Table A12, we report sensitivity analysis to the macroeconomic parameters. Increasing substitutability ϵ and income effects in labor supply γ lowers multipliers and reduces the effect of optimism on output. Decreasing the returns to scale α lowers the multiplier and the translation of hiring effects into production and reduces the effect of optimism on output. Lowering the Frisch elasticity ψ raises multipliers and increases the effects of optimism.

8 Conclusion

This paper studies the macroeconomic implications of viral, belief-altering narratives. We develop a conceptual framework in which narratives form a factor structure of agents’ models of exogenous and, in equilibrium, endogenous variables and spread virally between agents. We measure proxies for narratives among US firms and find evidence that narratives are both decision-relevant and viral, consistent with our framework. We develop a business-cycle model that embeds these findings and find that narratives can generate hysteresis, impulse responses that are hump-shaped and discontinuous, and boom-bust cycles. When we calibrate the model to match the data, we find that the business-cycle implications of narratives are quantitatively significant: we estimate that measured declines in optimism account for approximately 32% of the peak-to-trough decline in output over the Early 2000s Recession and 18% over the Great Recession.

Our analysis leaves (at least) two important issues unexplored. First, we do not investigate the interaction among narratives in either affecting decisions or spreading. As a result, we do not speak to Shiller’s (2020) thesis that *narrative constellations* have greater impact than any single narrative. Second, we microfound neither the set of narratives nor what determines virality. In short, we do not model what “makes a narrative a narrative.” We view the study of these issues as important inputs into a richer theory of the macroeconomics of narratives.

References

- ACEMOGLU, D., V. CHERNOZHUKOV, AND M. YILDIZ (2016): “Fragility of asymptotic agreement under Bayesian learning,” *Theoretical Economics*, 11(1), 187–225.
- ADAM, K., A. MARCET, AND J. BEUTEL (2017): “Stock price booms and expected capital gains,” *American Economic Review*, 107(8), 2352–2408.
- ANDRE, P., I. HAALAND, C. ROTH, AND J. WOHLFART (2021): “Inflation narratives,” *CEPR Discussion Paper No. DP16758*.
- ANDRE, P., C. PIZZINELLI, C. ROTH, AND J. WOHLFART (2022): “Subjective models of the macroeconomy: Evidence from experts and a representative sample,” *Working Paper*.
- ANGELETOS, G.-M., Z. HUO, AND K. A. SASTRY (2021): “Imperfect macroeconomic expectations: Evidence and theory,” *NBER Macroeconomics Annual*, 35(1), 1–86.
- ANGELETOS, G.-M., AND J. LA’O (2010): “Noisy business cycles,” *NBER Macroeconomics Annual*, 24(1), 319–378.
- (2013): “Sentiments,” *Econometrica*, 81(2), 739–779.
- ANGELETOS, G.-M., AND C. LIAN (2016): “Incomplete information in macroeconomics: Accommodating frictions in coordination,” in *Handbook of Macroeconomics*, vol. 2, pp. 1065–1240. Elsevier.
- ANGELETOS, G.-M., AND A. PAVAN (2007): “Efficient use of information and social value of information,” *Econometrica*, 75(4), 1103–1142.
- ANGRIST, J. D. (2014): “The perils of peer effects,” *Labour Economics*, 30, 98–108.
- BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2016): “Measuring economic policy uncertainty,” *The Quarterly Journal of Economics*, 131(4), 1593–1636.
- BARBERIS, N., R. GREENWOOD, L. JIN, AND A. SHLEIFER (2018): “Extrapolation and bubbles,” *Journal of Financial Economics*, 129(2), 203–227.
- BASU, S., AND J. G. FERNALD (1997): “Returns to scale in US production: Estimates and implications,” *Journal of Political Economy*, 105(2), 249–283.
- BATTAGLINI, M. (2021): “Chaos and Unpredictability in Dynamic Social Problems,” Discussion paper, National Bureau of Economic Research.
- BAXTER, M., AND R. G. KING (1999): “Measuring business cycles: approximate band-pass filters for economic time series,” *Review of Economics and Statistics*, 81(4), 575–593.
- BEAUDRY, P., D. GALIZIA, AND F. PORTIER (2020): “Putting the cycle back into business cycle analysis,” *American Economic Review*, 110(1), 1–47.
- BEAUDRY, P., AND F. PORTIER (2006): “Stock prices, news, and economic fluctuations,” *American Economic Review*, 96(4), 1293–1307.

- BELLONI, A., D. CHEN, V. CHERNOZHUKOV, AND C. HANSEN (2012): “Sparse models and methods for optimal instruments with an application to eminent domain,” *Econometrica*, 80(6), 2369–2429.
- BENHABIB, J., P. WANG, AND Y. WEN (2015): “Sentiments and aggregate demand fluctuations,” *Econometrica*, 83(2), 549–585.
- BENHIMA, K. (2019): “Booms and busts with dispersed information,” *Journal of Monetary Economics*, 107, 32–47.
- BHANDARI, A., J. BOROVIČKA, AND P. HO (2019): “Survey data and subjective beliefs in business cycle models,” *Working Paper, Available at SSRN 2763942*.
- BLANCHARD, O. J., AND N. KİYOTAKI (1987): “Monopolistic competition and the effects of aggregate demand,” *American Economic Review*, 77(4), 647–666.
- BLEI, D. M., A. Y. NG, AND M. I. JORDAN (2003): “Latent dirichlet allocation,” *Journal of Machine Learning Research*, pp. 993–1022.
- BOHREN, J. A., AND D. N. HAUSER (2021): “Learning with heterogeneous misspecified models: Characterization and robustness,” *Econometrica*, 89(6), 3025–3077.
- BOLDRIN, M., AND M. WOODFORD (1990): “Equilibrium models displaying endogenous fluctuations and chaos: a survey,” *Journal of Monetary Economics*, 25(2), 189–222.
- BORDALO, P., N. GENNAIOLI, S. Y. KWON, AND A. SHLEIFER (2021): “Diagnostic bubbles,” *Journal of Financial Economics*, 141(3), 1060–1077.
- BRAY, M. (1982): “Learning, estimation, and the stability of rational expectations,” *Journal of Economic Theory*, 26(2), 318–339.
- BRAY, M. M., AND N. E. SAVIN (1986): “Rational expectations equilibria, learning, and model specification,” *Econometrica*, 54(5), 1129–1160.
- BURNSIDE, C., M. EICHENBAUM, AND S. REBELO (2016): “Understanding booms and busts in housing markets,” *Journal of Political Economy*, 124(4), 1088–1147.
- BYBEE, L., B. T. KELLY, A. MANELA, AND D. XIU (2021): “Business news and business cycles,” Discussion paper, National Bureau of Economic Research.
- CABALLERO, R. J., AND A. SIMSEK (2020): “A risk-centric model of demand recessions and speculation,” *The Quarterly Journal of Economics*, 135(3), 1493–1566.
- CALDARA, D., AND C. KAMPS (2017): “The analytics of SVARs: a unified framework to measure fiscal multipliers,” *The Review of Economic Studies*, 84(3), 1015–1040.
- CARROLL, C. D. (2001): “The epidemiology of macroeconomic expectations,” Discussion paper, NBER Working Paper 8695.
- CESARINI, D., E. LINDQVIST, M. J. NOTOWIDIGDO, AND R. ÖSTLING (2017): “The effect of wealth on individual and household labor supply: evidence from Swedish lotteries,” *American Economic Review*, 107(12), 3917–46.

- CHRISTIANO, L. J., C. L. ILUT, R. MOTTO, AND M. ROSTAGNO (2008): “Monetary policy and stock market boom-bust cycles,” *ECB working paper*.
- DE LOECKER, J., J. EECKHOUT, AND G. UNGER (2020): “The rise of market power and the macroeconomic implications,” *The Quarterly Journal of Economics*, 135(2), 561–644.
- DENG, L., M. A. KHAN, AND T. MITRA (2022): “Continuous unimodal maps in economic dynamics: On easily verifiable conditions for topological chaos,” *Journal of Economic Theory*, 201, 105446.
- ELIAZ, K., AND R. SPIEGLER (2020): “A model of competing narratives,” *American Economic Review*, 110(12), 3786–3816.
- ESPONDA, I., AND D. POUZO (2016): “Berk–Nash equilibrium: A framework for modeling agents with misspecified models,” *Econometrica*, 84(3), 1093–1130.
- FLYNN, J. P., C. PATTERSON, AND J. STURM (2022): “Fiscal policy in a networked economy,” Discussion paper, National Bureau of Economic Research.
- FLYNN, J. P., AND K. A. SASTRY (2021): “Attention Cycles,” Working paper, MIT.
- FLYNN, Z., J. TRAINA, AND A. GANDHI (2019): “Measuring markups with production data,” Working Paper 3358472, SSRN.
- FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2008): “Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?,” *American Economic Review*, 98(1), 394–425.
- FRICK, M., R. IJIMA, AND Y. ISHII (2020): “Stability and robustness in misspecified learning models,” *Cowles Foundation Discussion Paper*.
- FUDENBERG, D., G. LANZANI, AND P. STRACK (2021): “Limit points of endogenous misspecified learning,” *Econometrica*, 89(3), 1065–1098.
- GALÍ, J. (2008): *Monetary policy, in ation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.
- GENTRY, R. J., J. S. HARRISON, T. J. QUIGLEY, AND S. BOIVIE (2021): “A database of CEO turnover and dismissal in S&P 1500 firms, 2000–2018,” *Strategic Management Journal*, 42(5), 968–991.
- GENTZKOW, M., B. KELLY, AND M. TADDY (2019): “Text as data,” *Journal of Economic Literature*, 57(3), 535–74.
- HANSEN, L., AND T. J. SARGENT (2001): “Robust control and model uncertainty,” *American Economic Review*, 91(2), 60–66.
- HARRISON, J. M., AND D. M. KREPS (1978): “Speculative investor behavior in a stock market with heterogeneous expectations,” *The Quarterly Journal of Economics*, 92(2), 323–336.

- HASSAN, T. A., S. HOLLANDER, L. VAN LENT, AND A. TAHOUN (2019): “Firm-level political risk: Measurement and effects,” *The Quarterly Journal of Economics*, 134(4), 2135–2202.
- HELLWIG, C., AND L. VELDKAMP (2009): “Knowing what others know: Coordination motives in information acquisition,” *The Review of Economic Studies*, 76(1), 223–251.
- HOFFMAN, M., F. BACH, AND D. BLEI (2010): “Online learning for latent dirichlet allocation,” *Advances in Neural Information Processing Systems*, 23.
- ISABELLA, L. A. (1990): “Evolving interpretations as a change unfolds: How managers construe key organizational events,” *Academy of Management journal*, 33(1), 7–41.
- KAUSTIA, M., AND V. RANTALA (2021): “Common analysts: method for defining peer firms,” *Journal of Financial and Quantitative Analysis*, 56(5), 1505–1536.
- KEYNES, J. M. (1936): *The General Theory of Employment, Interest and Money*. Palgrave Macmillan.
- KOZLOWSKI, J., L. VELDKAMP, AND V. VENKATESWARAN (2020): “The tail that wags the economy: Beliefs and persistent stagnation,” *Journal of Political Economy*, 128(8), 2839–2879.
- LI, T.-Y., AND J. A. YORKE (1975): “Period Three Implies Chaos,” *The American Mathematical Monthly*, 82(10), 985–992.
- LOEWENSTEIN, J., W. OCASIO, AND C. JONES (2012): “Vocabularies and vocabulary structure: A new approach linking categories, practices, and institutions,” *Academy of Management Annals*, 6(1), 41–86.
- LORENZONI, G. (2009): “A theory of demand shocks,” *American Economic Review*, 99(5), 2050–84.
- LOUGHRAN, T., AND B. MCDONALD (2011): “When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks,” *The Journal of Finance*, 66(1), 35–65.
- MAENHOUT, P. J., A. VEDOLIN, AND H. XING (2021): “Robustness and dynamic sentiment,” *Available at SSRN 3798445*.
- MAITLIS, S. (2005): “The social processes of organizational sensemaking,” *Academy of management journal*, 48(1), 21–49.
- MAR CET, A., AND T. J. SARGENT (1989a): “Convergence of least-squares learning in environments with hidden state variables and private information,” *Journal of Political Economy*, 97(6), 1306–1322.
- (1989b): “Convergence of least squares learning mechanisms in self-referential linear stochastic models,” *Journal of Economic Theory*, 48(2), 337–368.
- MAXTED, P. (2020): “A macro-finance model with sentiment,” *Unpublished working paper. Harvard University*.

- MCDONALD, B. (2021): “Software Repository for Accounting and Finance,” Accessed June 2020. Available at: <https://sraf.nd.edu/>.
- MOLAVI, P. (2019): “Macroeconomics with learning and misspecification: A general theory and applications,” *Working Paper*.
- MORRIS, S., AND H. S. SHIN (2002): “Social value of public information,” *American Economic Review*, 92(5), 1521–1534.
- NICKELL, S. (1981): “Biases in dynamic models with fixed effects,” *Econometrica*, 49(6), 1417–1426.
- OTTONELLO, P., AND T. WINBERRY (2020): “Financial heterogeneity and the investment channel of monetary policy,” *Econometrica*, 88(6), 2473–2502.
- REICHLING, F., AND C. WHALEN (2012): “Review of estimates of the Frisch elasticity of labor supply,” *Working Paper*.
- SCHAAL, E., AND M. TASCHEREAU-DUMOUCHEL (2020): “Herding Through Booms and Busts,” *Working Paper*.
- SCHEINKMAN, J. A., AND W. XIONG (2003): “Overconfidence and speculative bubbles,” *Journal of political Economy*, 111(6), 1183–1220.
- SCHWARTZSTEIN, J., AND A. SUNDERAM (2021): “Using models to persuade,” *American Economic Review*, 111(1), 276–323.
- SEC (2011): “Investor Bulletin: How to Read a 10-K,” Article, Securities and Exchange Commission Office of Investor Education and Advocacy.
- SHILLER, R. J. (2017): “Narrative Economics,” *American Economic Review*, 107(4), 967–1004.
- (2020): *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.
- SPIEGLER, R. (2016): “Bayesian networks and boundedly rational expectations,” *The Quarterly Journal of Economics*, 131(3), 1243–1290.
- VAARA, E., S. SONENSHEIN, AND D. BOJE (2016): “Narratives as sources of stability and change in organizations: Approaches and directions for future research,” *Academy of Management Annals*, 10(1), 495–560.
- WEISSTEIN, E. W. (2001): “Logistic map,” *Wolfram MathWorld*.
- WOODFORD, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Appendices

A Omitted Derivations and Proofs

A.1 Derivation of Equations 4 and 5

We first provide two assumptions under which Equation 4 holds as a linear approximation. In what follows, we impose the technical requirements that X is convex, compact subset of \mathbb{R} , Y is a convex, compact subset of \mathbb{R}^n , Θ is a convex, compact subset of \mathbb{R}^m , and Ω is a convex, compact subset of \mathbb{R}^r . We first assume regularity on payoffs to ensure the sufficiency of first-order conditions for optimality:

Assumption 1. *The utility function u is strictly concave and twice continuously differentiable.*

We next assume that agents' information about the random fundamentals is generated by location experiments that are conditionally independent of the fundamental, agents' preference shifters and the narratives held by agents.

Assumption 2. *The agents' information sets are generated by location experiments, i.e., $s_{it} = \theta_t + \nu_{it}$, where ν_{it} is a zero-mean random variable that is independent of θ_t , ω_i , and λ_{it} .*

Under these two assumptions, we can derive the form of the regression equation and that, modulo any misspecification error, the conditional expectation function is linear.

Proposition 8. *Under Assumptions 1 and 2, we have that:*

$$x_{it} = \gamma_i + \chi_t + \sum_{k=1}^K \delta_k \lambda_{k;it} + \varepsilon_{it} + O(\|j(x_{it}, Y_t, \theta_t, Q_t, \omega_i, \nu_{it}, \lambda_{it})\|^2) \quad (63)$$

where ε_{it} is a zero mean random variable that is uncorrelated with γ_i , χ_t and λ_{it} . Thus, net of the misspecification error, the conditional expectation function is given by:

$$E[x_{it}|i, t, \lambda_{it}] = \gamma_i + \chi_t + \sum_{k=1}^K \delta_k \lambda_{k;it} \quad (64)$$

Proof. By Assumption 1, from the agents' problems (Equation 3), their best replies must solve the following first-order condition (where we suppress all individual and time subscripts):

$$E_x \left[u_x(x, \hat{Y}(\theta), \theta, \omega) | s \right] = 0 \quad (65)$$

We linearize this first-order conditions in (x, Y, θ, ω) around values $(\bar{x}, \bar{Y}, \bar{\theta}, \bar{\omega})$ which satisfy $E_{\lambda} [u_x(\bar{x}, \bar{Y}(\theta), \bar{\theta}, \bar{\omega})] = 0$. This gives

$$E_{\lambda} [u_{xx}(x - \bar{x}) + u_{xY}^{\theta}(Y - \bar{Y}) + u_x^{\theta}(\theta - \bar{\theta}) + u_{x\omega}^{\omega}(\omega - \bar{\omega})] + R = 0 \quad (66)$$

where the remainder R is $O(jj(x, Y, \theta, \omega, \nu, \lambda)jj^2)$. We can rearrange the above, and use the fact that ω is known to the agent, to write:

$$x = \bar{x} + \frac{1}{ju_{xxj}}u_{x\omega}^{\omega}(\omega - \bar{\omega}) + \frac{1}{ju_{xxj}}E_{\lambda} [u_{xY}^{\theta}(Y - \bar{Y}) + u_x^{\theta}(\theta - \bar{\theta})] + \frac{1}{ju_{xxj}}R \quad (67)$$

Moreover, we know that $\hat{Y} = \hat{Y}(Q, \theta)$. Thus, assuming that \hat{Y} is continuously differentiable, we may linearize $Y = \bar{Y} + Y_Q^{\theta}(Q - \bar{Q}) + Y^{\theta}(\theta - \bar{\theta}) + \hat{R}$, where $\bar{Y} = \hat{Y}(\bar{Q}, \bar{\theta})$ and \hat{R} is the error induced by the approximation of \hat{Y} , which is $O(jjY, Q, \theta jj^2)$. Substituting this approximation into Equation 67 gives

$$\begin{aligned} x &= \bar{x} + \frac{1}{ju_{xxj}}u_{x\omega}^{\omega}(\omega - \bar{\omega}) + \frac{1}{ju_{xxj}}E_{\lambda} [u_{xY}^{\theta}(Y_Q^{\theta}(Q - \bar{Q}) + Y^{\theta}(\theta - \bar{\theta})) + u_x^{\theta}(\theta - \bar{\theta})] + \tilde{R} \\ &= \gamma + \tilde{\chi} + E_{\lambda} [\tilde{\theta}js] + \tilde{R} \end{aligned} \quad (68)$$

where $\gamma = \bar{x} + \frac{1}{ju_{xxj}}u_{x\omega}^{\omega}(\omega - \bar{\omega}) + \frac{1}{ju_{xxj}}u_{xY}^{\theta}(Y_Q^{\theta}\bar{Q} + Y^{\theta}\bar{\theta}) + \frac{1}{ju_{xxj}}u_x^{\theta}\bar{\theta}$, $\tilde{\chi} = \frac{1}{ju_{xxj}}u_{xY}^{\theta}Y_Q^{\theta}Q$, and $\tilde{\theta} = (u_{xY}^{\theta}Y^{\theta} + u_x^{\theta})\theta$, $\tilde{R} = \frac{1}{ju_{xxj}}R + \hat{R}$. We now linearize the conditional expectation of $\tilde{\theta}$ to obtain:

$$E_{\lambda} [\tilde{\theta}js] = E_s \tilde{s} + E E_{\lambda} [\tilde{\theta}] + \tilde{R} \quad (69)$$

where \tilde{R} is the error induced by the approximation and is $O(jj(\theta, \nu, \lambda)jj^2)$, $\tilde{s} = (u_{xY}^{\theta}Y^{\theta} + u_x^{\theta})s = \tilde{\theta} + \tilde{\nu}$, with $\tilde{\nu} = (u_{xY}^{\theta}Y^{\theta} + u_x^{\theta})\nu$ independent of $\tilde{\theta}$ and of mean zero by Assumption 2. Defining $\chi = \tilde{\chi} + \kappa\tilde{\theta}$, $\varepsilon = E_s \hat{\nu}$ and $\delta_k = E E_k[\tilde{\theta}]$, where the expectation of $\tilde{\theta}$ when θ has distribution N_k (*i.e.*, that of narrative k), we may write:

$$x = \gamma + \chi + \sum_{k=1}^K \delta_k \lambda_k + \varepsilon + \bar{R} \quad (70)$$

where $\bar{R} = \tilde{R} + \tilde{R} = O(jj(x, Y, \theta, Q, \omega, \nu, \lambda)jj^2)$. Re-introducing subscripts, we have ω_i , Q_t , $\tilde{\theta}_t$, $\lambda_{k,it}$ and ε_{it} . Thus, we have the claimed regression equation:

$$x_{it} = \gamma_i + \chi_t + \sum_{k=1}^K \delta_k \lambda_{k,it} + \varepsilon_{it} + O(jj(x_{it}, Y_t, \theta_t, Q_t, \omega_i, \nu_{it}, \lambda_{it})jj^2) \quad (71)$$

As ε_{it} has zero mean and is uncorrelated with γ_i , χ_t and λ_{it} , the claimed formula for the conditional expectation function follows. \square

We now turn to the narrative updating rule. We impose the following assumption:

Assumption 3. *The updating rule P is continuously differentiable.*

We finally derive Equation 5 under this condition:

Proposition 9. *Under Assumption 3, we have that:*

$$P[\lambda_{it} = \lambda_j \lambda_{i:t-1}, Y_{t-1}, Q_{t-1}] = \zeta + u^\theta \lambda_{i:t-1} + r^\theta Y_{t-1} + s^\theta Q_{t-1} + O(jj(\lambda_{i:t-1}, Y_{t-1}, Q_{t-1})jj^2) \quad (72)$$

Proof. By definition we have that $P[\lambda_{it} = \lambda_j \lambda_{i:t-1}, Y_{t-1}, Q_{t-1}] = P(\lambda_{i:t-1}, Y_{t-1}, Q_{t-1})$. Linearizing this expression under Assumption 3, we immediately have:

$$P[\lambda_{it} = \lambda_j \lambda_{i:t-1} = \lambda^\theta, Y_{t-1}, Q_{t-1}] = \zeta + u^\theta + r^\theta Y_{t-1} + s^\theta Q_{t-1} + O(jj(\lambda_{i:t-1}, Y_{t-1}, Q_{t-1})jj^2) \quad (73)$$

Completing the proof. \square

A.2 Proof of Proposition 1

Proof. We guess and verify that there exists a unique quasi-linear equilibrium. That is, there exists a unique equilibrium of the following form for some parameters $a_0, a_1 \in \mathbb{R}$ and function $f : [0, 1] \rightarrow \mathbb{R}$:

$$\log Y(\theta, Q) = a_0 + a_1 \log \theta + f(Q) \quad (74)$$

To verify this conjecture, we need to compute best replies under this conjecture and show that when we aggregate these best replies that the conjecture is consistent and, moreover, that it is consistent for a unique triple (a_0, a_1, f) .

From the arguments in the main text, we need to compute two objects: $\log E_{it} \left[\theta_{it}^{\frac{1+\psi}{\alpha}} \right]$ and $\log E_{it} \left[Y_{it}^{\frac{1}{\epsilon}} \right]$. We can compute the first object directly. Conditional on a signal s_{it} and a narrative weight λ_{it} , we have that the distribution of the aggregate component of productivity is:

$$\log \theta_{it} | s_{it}, \lambda_{it} \sim N(\kappa s_{it} + (1 - \kappa)(\lambda_{it} \mu_O + (1 - \lambda_{it}) \mu_P), \sigma_{js}^2) \quad (75)$$

by the standard formula for the conditional distribution of jointly normal random variables,

where:

$$\kappa = \frac{1}{1 + \frac{\frac{\psi}{2}}{\theta}} \quad \text{and} \quad \sigma_{js}^2 = \frac{1}{\frac{1}{2} + \frac{1}{\frac{\psi}{\theta}}} \quad (76)$$

with κ being the signal-to-noise ratio and σ_{js}^2 the variance of fundamentals conditional on the signal. Thus, idiosyncratic productivity has conditional distribution given by:

$$\log \theta_{it} | s_{it}, \lambda_{it} \sim N(\log \gamma_i + \kappa s_{it} + (1 - \kappa)(\lambda_{it} \mu_O + (1 - \lambda_{it}) \mu_P), \sigma_{js}^2 + \sigma_{\epsilon}^2) \quad (77)$$

where we will denote the above mean by μ_{it} and variance by η^2 . Hence, rewriting and using the moment generating function of a normal random variable, we have that:

$$\begin{aligned} \log \mathbb{E}_{it} \left[\theta_{it}^{\frac{1+\psi}{\alpha}} \right] &= \log \mathbb{E}_{it} \left[\exp \left\{ \frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] \\ &= \frac{1+\psi}{\alpha} \mu_{it} + \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 \eta^2 \end{aligned} \quad (78)$$

Under our conjecture (Equation 74), we can moreover compute:

$$\begin{aligned} \log \mathbb{E}_{it} \left[Y_t^{\frac{1}{\epsilon}} \right] &= \log \mathbb{E}_{it} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) (a_0 + a_1 \log \theta_t + f(Q_t)) \right\} \right] \\ &= \left(\frac{1}{\epsilon} - \gamma \right) [a_0 + a_1 (\mu_{it} - \log \gamma_i) + f(Q_t)] + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 [\eta^2 + \sigma_{\epsilon}^2] \end{aligned} \quad (79)$$

Thus, we have that best replies under our conjecture are given by:

$$\begin{aligned} \log x_{it} &= \frac{1}{\frac{1}{1+\psi} + 1} \left[\log \left(\frac{1}{1+\psi} \right) + \frac{1+\psi}{\alpha} \mu_{it} + \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 \eta^2 \right. \\ &\quad \left. + \left(\frac{1}{\epsilon} - \gamma \right) [a_0 + a_1 (\mu_{it} - \log \gamma_i) + f(Q_t)] + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 [\eta^2 + \sigma_{\epsilon}^2] \right] \end{aligned} \quad (80)$$

To confirm the conjecture, we must now aggregate these levels of production and show that they are consistent with the conjecture. Performing this aggregation we have that:

$$\begin{aligned} \log Y_t &= \log \left[\left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \right] \\ &= \frac{\epsilon}{\epsilon-1} \log \mathbb{E}_t \left[\exp \left\{ \frac{\epsilon-1}{\epsilon} \log x_{it} \right\} \right] \\ &= \frac{\epsilon}{\epsilon-1} \log \mathbb{E}_t \left[\mathbb{E}_t \left[\exp \left\{ \frac{\epsilon-1}{\epsilon} \log x_{it} \right\} | \lambda_{it} \right] \right] \end{aligned} \quad (81)$$

Moreover, expanding the terms in Equation 80, we have that:

$$\begin{aligned}
\log x_{it} = & \frac{1}{\frac{1+\psi}{\alpha} + 1} \left[\log \left(\frac{1+\psi}{\alpha} \right) \right. \\
& + \frac{1+\psi}{\alpha} [\log \gamma_i + \kappa s_{it} + (1-\kappa) [\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P]] \\
& + \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 (\sigma_{js}^2 + \sigma_{\epsilon}^2) \\
& + \left(\frac{1}{\epsilon} - \gamma \right) [a_0 + a_1 (\kappa s_{it} + (1-\kappa) [\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P])] + f(Q_t) \\
& \left. + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2 \right]
\end{aligned} \tag{82}$$

which is, conditional on λ_{it} , normally distributed as both $\log \gamma_i$ and s_{it} are both normal.

Hence, we write $\log x_{it} | \lambda_{it} \sim N(\delta_t(\lambda_{it}), \hat{\sigma}^2)$, where:

$$\begin{aligned}
\delta_t(\lambda_{it}) = & \frac{1}{\frac{1+\psi}{\alpha} + 1} \left[\log \left(\frac{1+\psi}{\alpha} \right) \right. \\
& + \frac{1+\psi}{\alpha} [\mu + \kappa \log \theta_t + (1-\kappa) [\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P]] \\
& + \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 (\sigma_{js}^2 + \sigma_{\epsilon}^2) \\
& + \left(\frac{1}{\epsilon} - \gamma \right) [a_0 + a_1 (\kappa \log \theta_t + (1-\kappa) [\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P])] + f(Q_t) \\
& \left. + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2 \right]
\end{aligned} \tag{83}$$

and:

$$\hat{\sigma}^2 = \left(\frac{1}{\frac{1+\psi}{\alpha} + 1} \right)^2 \left[\left(\frac{1+\psi}{\alpha} \right)^2 \sigma^2 + \kappa^2 \left[\frac{1+\psi}{\alpha} + a_1 \left(\frac{1}{\epsilon} - \gamma \right) \right]^2 \sigma_{\epsilon}^2 \right] \tag{84}$$

Thus, we have that:

$$\mathbb{E}_t \left[\exp \left\{ \frac{\epsilon}{\epsilon} \frac{1}{\epsilon} \log x_{it} \right\} | \lambda_{it} \right] = \exp \left\{ \frac{\epsilon}{\epsilon} \frac{1}{\epsilon} \delta_t(\lambda_{it}) + \frac{1}{2} \left(\frac{\epsilon}{\epsilon} \frac{1}{\epsilon} \right)^2 \hat{\sigma}^2 \right\} \tag{85}$$

and so:

$$\begin{aligned}
\mathbb{E}_t \left[\mathbb{E}_t \left[\exp \left\{ \frac{\epsilon - 1}{\epsilon} \log x_{it} \right\} j \lambda_{it} \right] \right] &= Q_t \exp \left\{ \frac{\epsilon - 1}{\epsilon} \delta_t(1) + \frac{1}{2} \left(\frac{\epsilon - 1}{\epsilon} \right)^2 \hat{\sigma}^2 \right\} \\
&+ (1 - Q_t) \exp \left\{ \frac{\epsilon - 1}{\epsilon} \delta_t(0) + \frac{1}{2} \left(\frac{\epsilon - 1}{\epsilon} \right)^2 \hat{\sigma}^2 \right\} \\
&= \left[Q_t \exp \left\{ \frac{\epsilon - 1}{\epsilon} (\delta_t(1) - \delta_t(0)) \right\} + (1 - Q_t) \right] \exp \left\{ \frac{\epsilon - 1}{\epsilon} \delta_t(0) + \frac{1}{2} \left(\frac{\epsilon - 1}{\epsilon} \right)^2 \hat{\sigma}^2 \right\}
\end{aligned} \tag{86}$$

Yielding:

$$\log Y_t = \delta_t(0) + \frac{1}{2} \frac{\epsilon - 1}{\epsilon} \hat{\sigma}^2 + \frac{\epsilon}{\epsilon - 1} \log \left(Q_t \exp \left\{ \frac{\epsilon - 1}{\epsilon} (\delta_t(1) - \delta_t(0)) \right\} + (1 - Q_t) \right) \tag{87}$$

where we define $\alpha \delta^{OP} = \delta_t(1) - \delta_t(0)$ and compute:

$$\delta_t(1) - \delta_t(0) = \frac{1}{\frac{1+\psi}{\alpha} + 1} \left(\frac{1+\psi}{\alpha} + a_1 \left(\frac{1}{\epsilon} - \gamma \right) \right) (1 - \kappa) (\mu_O - \mu_P) = \alpha \delta^{OP} \tag{88}$$

and note that this is a constant. Finally, we see that $\delta_t(0)$ is given by:

$$\begin{aligned}
\delta_t(0) &= \frac{1}{\frac{1+\psi}{\alpha} + 1} \left[\log \left(\frac{1 - \frac{1}{\epsilon}}{1 + \frac{1}{\epsilon}} \right) + \frac{1 + \psi}{\alpha} (\mu - (1 - \kappa) \mu_P) - \frac{1}{2} \left(\frac{1 + \psi}{\alpha} \right)^2 (\sigma_{js}^2 + \sigma_-^2) \right. \\
&+ \left. \left(\frac{1}{\epsilon} - \gamma \right) (a_0 + a_1 (1 - \kappa) \mu_P) + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2 \right. \\
&+ \left. \left[\frac{1 + \psi}{\alpha} + a_1 \left(\frac{1}{\epsilon} - \gamma \right) \right] \kappa \log \theta_t + \left(\frac{1}{\epsilon} - \gamma \right) f(Q_t) \right]
\end{aligned} \tag{89}$$

By matching coefficients between Equations 87 and Equation 74, we obtain a_0 , a_1 , and f .

We first match coefficients on $\log \theta_t$ to obtain an equation for a_1 :

$$a_1 = \frac{\left[\frac{1+\psi}{\alpha} + a_1 \left(\frac{1}{\epsilon} - \gamma \right) \right] \kappa}{\frac{1+\psi}{\alpha} + 1} \tag{90}$$

Under our maintained assumption that $\frac{1}{\frac{1+\psi}{\alpha} + 1} \geq [0, 1)$, as $\kappa \geq [0, 1]$, we have that this has

a unique solution:

$$a_1 = \frac{\frac{1+\psi}{\frac{1+\psi}{\alpha} + \frac{1}{\epsilon}}}{1 + \frac{\left(\frac{1}{\epsilon}\right)}{\frac{1+\psi}{\alpha} + \frac{1}{\epsilon}}} \quad (91)$$

It is moreover positive.

Second, by collecting terms with Q_t we obtain an equation for f :

$$f(Q) = \frac{1}{1 + \frac{\gamma}{\epsilon}} f(Q) + \frac{\epsilon}{1} \log \left(1 + Q \left[\exp \left\{ \frac{\epsilon}{\epsilon} \alpha \delta^{OP} \right\} - 1 \right] \right) \quad (92)$$

which has a unique solution as $\frac{1}{1 + \frac{\gamma}{\epsilon}} \in [0, 1)$ and can be solved to yield:

$$f(Q) = \frac{1}{1 + \frac{\gamma}{\epsilon}} \log \left(1 + Q \left[\exp \left\{ \frac{\epsilon}{\epsilon} \alpha \delta^{OP} \right\} - 1 \right] \right) \quad (93)$$

where we observe that δ^{OP} depends only on primitive parameters and a_1 , for which we have already solved. Finally, by collecting constants, we obtain an equation for a_0 :

$$\begin{aligned} a_0 = & \frac{1}{1 + \frac{\gamma}{\epsilon}} \left[\log \left(\frac{1}{1 + \frac{\gamma}{\epsilon}} \right) + \frac{1 + \psi}{\alpha} (\mu + (1 - \kappa)\mu_P) - \frac{1}{2} \left(\frac{1 + \psi}{\alpha} \right)^2 (\sigma_{js}^2 + \sigma_-^2) \right. \\ & \left. + \left(\frac{1}{\epsilon} - \gamma \right) (a_0 + a_1(1 - \kappa)\mu_P) + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2 \right] + \frac{1}{2} \frac{\epsilon}{\epsilon} \hat{\sigma}^2 \end{aligned} \quad (94)$$

Solving this equation yields:

$$\begin{aligned} a_0 = & \frac{1}{1 + \frac{\gamma}{\epsilon}} \left[\frac{1}{1 + \frac{\gamma}{\epsilon}} \left[\log \left(\frac{1}{1 + \frac{\gamma}{\epsilon}} \right) + \frac{1 + \psi}{\alpha} (\mu + (1 - \kappa)\mu_P) - \frac{1}{2} \left(\frac{1 + \psi}{\alpha} \right)^2 (\sigma_{js}^2 + \sigma_-^2) \right. \right. \\ & \left. \left. + \left(\frac{1}{\epsilon} - \gamma \right) a_1(1 - \kappa)\mu_P + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2 \right] + \frac{1}{2} \frac{\epsilon}{\epsilon} \hat{\sigma}^2 \right] \end{aligned} \quad (95)$$

which we observe depends only on parameters, a_1 , and $\hat{\sigma}^2$. Moreover, $\hat{\sigma}^2$ depends only on parameters and a_1 . Thus, given that we have solved for a_1 , we have now recovered a_0 , a_1 , and f uniquely and verified that there exists a unique quasi-linear equilibrium. Finally, to obtain the formula for the best reply of agents, simply substitute a_0 , a_1 , and f into Equation 82 and label the coefficients as in the claim. \square

A.3 Proof of the Claims in Remark 1

We now prove the claims made in Remark 1. We have already shown that there exists a unique quasi-linear equilibrium. More generally, we seek to rule out an equilibrium of any other form. To do so, we show that there is a unique equilibrium when fundamentals are bounded by some $M \in \mathbb{R}$, $\log \theta_t \in [-M, M]$, $\log \gamma_i \in [-M, M]$, $\log \tilde{\theta}_{it} \in [-M, M]$, and $\varepsilon_{it} \in [-M, M]$.

Lemma 2. *When fundamentals are bounded, there exists a unique equilibrium*

Proof. To this end, we can recast any equilibrium function $\log Y(\theta, q)$ as one that solves the fixed point in Equation 41. In the case where fundamentals are bounded, this can be accomplished by demonstrating that the implied fixed-point operator is a contraction by verifying Blackwell's sufficient conditions. More formally, consider the space of bounded, real-valued functions \mathcal{C} under the L^1 -norm and consider the operator $V_M : \mathcal{C} \rightarrow \mathcal{C}$ given by:

$$V_M(g)(\theta, Q) = \frac{\epsilon}{\epsilon - 1} \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \frac{-1}{1 + \psi} \left(\log \left(\frac{1 - \psi}{1 + \psi} \right) \right. \right. \right. \\ \left. \left. \left. \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \frac{1 + \psi}{\alpha} \log \theta_{it} \right\} \right] + \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \right) \right\} \right] \quad (96)$$

The following two conditions are sufficient for this operator to be a contraction: (i) monotonicity: for all $g, h \in \mathcal{C}$ such that $g \leq h$, we have that $V_M(g) \leq V_M(h)$ (ii) discounting: there exists a parameter $c \in [0, 1)$ such that for all $g \in \mathcal{C}$ and $a \in \mathbb{R}_+$ and $V_M(g + a) \leq V_M(g) + ca$. Thus, as the space of bounded functions under the L^1 -norm is a complete metric space, if Blackwell's conditions hold, then by the Banach fixed-point theorem, there exists a unique fixed point of the operator V_M .

To complete this argument, we now verify (i) and (ii). To show monotonicity, observe that $\frac{1}{\epsilon} - \gamma \geq 0$ as $\omega \geq 0$ and recall that $\epsilon > 1$. Thus, we have that:

$$\log E_{(s; \mathcal{Q})} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \leq \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) h \right\} \right] \quad (97)$$

for all (s, Q) . And so $V_M(g)(\theta, Q) \leq V_M(h)(\theta, Q)$ for all (θ, Q) . To show discounting, observe that:

$$\log E_{(s; \mathcal{Q})} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) (g + a) \right\} \right] = \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] + \left(\frac{1}{\epsilon} - \gamma \right) a \quad (98)$$

And so:

$$\begin{aligned}
V_M(g+a)(\theta, Q) &= \frac{\epsilon}{\epsilon-1} \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \frac{-1}{1+\frac{\psi}{\alpha}} \left(\log \left(\frac{1-\frac{\psi}{\alpha}}{1+\frac{\psi}{\alpha}} \right) \right. \right. \right. \\
&\quad \left. \left. \left. \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] + \log E_{(s; \mathcal{Q})} \left[\exp \left\{ \left(\frac{1}{\epsilon} \quad \gamma \right) g \right\} \right] + \left(\frac{1}{\epsilon} \quad \gamma \right) a \right\} \right] \right] \\
&= V_M(g)(\theta, Q) + \omega a
\end{aligned} \tag{99}$$

where $\omega \in [0, 1)$ by assumption. Note that the modulus of contraction ω is precisely the claimed strategic complementarity parameter in Equation 27. This verifies equilibrium uniqueness. \square

Away from the case with bounded fundamentals, the above strategy cannot be used to demonstrate uniqueness. Even though the fixed-point operator still satisfies Blackwell's conditions, the relevant function space now becomes any L^p -space for $p \geq (1, \infty)$ and the sup-norm over such spaces can be infinite, making Blackwell's conditions insufficient for V to be a contraction. In this case, we show that the unique quasi-linear equilibrium in the unbounded fundamentals case is an appropriately-defined ϵ -equilibrium for any $\epsilon > 0$. Let the unique quasi-linear equilibrium we have guessed and verified be $\log Y$. We say that g is a ϵ -equilibrium if

$$\|g - V_M(g)\|_p < \epsilon \tag{100}$$

where $\|\cdot\|_p$ is the L^p -norm. In words, g is a ϵ -equilibrium if its distance from being a fixed point is at most ϵ . The following Lemma establishes that $\log Y$ is a ϵ -equilibrium for bounded fundamentals for any $\epsilon > 0$ for some bound M :

Lemma 3. *For every $\epsilon > 0$, there exists an $M \in \mathbb{N}$ such that $\log Y$ is a ϵ -equilibrium.*

Proof. Now extend from \mathcal{C} , $V_M : L^p(\mathbb{R}) \rightarrow L^p(\mathbb{R})$ as in Equation 96. We observe that V_M is continuous in the limit in M in the sense that $V_M(g) \rightarrow V(g)$ as $M \rightarrow \infty$ for all $g \in L^p(\mathbb{R})$. This observation follows from noting that both $\log E_{(s; \mathcal{Q})} [\exp \{ \frac{1+\psi}{\alpha} \log \theta_{it} \}]$ and $\log E_{(s; \mathcal{Q})} [\exp \{ (\frac{1}{\epsilon} \quad \gamma) g \}]$ are convergent pointwise for $M \rightarrow \infty$ for all (s, Q) . In Proposition 1, we showed that $V(\log Y) = \log Y$. Thus, we have that: $V_M(\log Y) \rightarrow V(\log Y) = \log Y$, which implies that:

$$\lim_{M \rightarrow \infty} \| \log Y - V_M(\log Y) \|_p = 0 \tag{101}$$

which implies that for every $\varepsilon > 0$, there exists a $\bar{M} \geq N$ such that:

$$|E[\log Y] - V_M(\log Y)| < \varepsilon \quad \forall M \geq N : M > \bar{M} \quad (102)$$

Completing the proof. \square

A.4 Proof of Corollary 1

Proof. From Equation 82, we may express:

$$\begin{aligned} \log x_{it} = & \text{cons} + b_3 f(Q_t) + \frac{1+\psi}{\alpha} (\log \gamma_i + \kappa S_{it}) + \left(\frac{1}{\varepsilon} - \gamma \right) \frac{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}{1 - \frac{\left(\frac{1}{\varepsilon}\right)}{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}} \kappa S_{it} \\ & \left[\frac{1}{\frac{1}{1+\psi} + 1} + \left(\frac{1}{\varepsilon} - \gamma \right) \frac{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}{1 - \frac{\left(\frac{1}{\varepsilon}\right)}{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}} \right] (1 - \kappa) (\mu_O - \mu_P) \lambda_{it} \end{aligned} \quad (103)$$

We substitute this expression into $\log L_{it} = \frac{1}{\alpha} (\log x_{it} - \log \theta_{it})$ to write

$$\begin{aligned} \log L_{it} = & \frac{1}{\alpha} \log \theta_{it} + \text{cons} + c_4 f(Q_t) + \text{cons}_i \\ & + \frac{1}{\alpha} \left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\varepsilon} - \gamma \right) \frac{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}{1 - \frac{\left(\frac{1}{\varepsilon}\right)}{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}} \right] \kappa \log \theta_t \\ & + \frac{1}{\alpha} \left[\frac{1}{\frac{1}{1+\psi} + 1} + \left(\frac{1}{\varepsilon} - \gamma \right) \frac{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}{1 - \frac{\left(\frac{1}{\varepsilon}\right)}{\frac{1+\psi}{\alpha} \frac{1}{\alpha + \frac{1}{\varepsilon}}}} \right] (1 - \kappa) (\mu_O - \mu_P) \lambda_{it} \\ & + \xi_{it}^\theta \end{aligned} \quad (104)$$

where $\xi_{it}^\theta \sim N(0, \sigma^2)$ and IID. Comparing the above with the definition of $\alpha \delta^{OP}$ in Equation 88, we see that the coefficient on $\log \theta_t$ in the above expression is δ^{OP} . We finally observe from Equation 93 that $f(Q)$ depends on $(\varepsilon, \gamma, \psi, \alpha)$ and δ^{OP} . Hence, given $(\varepsilon, \gamma, \psi, \alpha)$, f is identified uniquely from the studied regression estimate. \square

A.5 Proof of Corollary 2

Proof. This is immediate by substituting Equation 42 into Equation 32. \square

A.6 Proof of Proposition 2

Proof. We prove the three claims in sequence.

(i) The map $T : [0, 1] \rightarrow [0, 1]$ is continuous as f , P_O and P_P are continuous for all $\theta \geq \Theta$ functions. Moreover, it maps a compact set to a compact set. Thus, by Brouwer's fixed point theorem, there exists a Q such that $Q = T(Q)$ for all $\theta \geq \Theta$.

(ii) To characterize the existence of extremal steady states, observe that $Q = 1$ is a steady state for θ if and only if $T(1) = P_O(a_0 + a_1 \log \theta + f(1), 1) = 1$ and $Q = 0$ is a steady state for θ if and only if $T(0) = P_P(a_0 + a_1 \log \theta, 0) = 0$. Thus, $Q = 1$ is a steady state if and only if $P_O^{-1}(1; 1) = a_0 + a_1 \log \theta + f(1)$ and $Q = 0$ is a steady state if and only if $P_P^{-1}(0; 0) = a_0 + a_1 \log \theta$.

(iii) To analyze the stability of the extremal steady states, observe that if $T'(Q) < 1$ at a steady state Q , then Q is stable. When it exists (which it does almost everywhere), we have that:

$$\begin{aligned} T'(Q) &= P_O(a_0 + a_1 \log \theta + f(Q), Q) - P_P(a_0 + a_1 \log \theta + f(Q), Q) \\ &\quad + Q \frac{d}{dQ} P_O(a_0 + a_1 \log \theta + f(Q), Q) + (1 - Q) \frac{d}{dQ} P_P(a_0 + a_1 \log \theta + f(Q), Q) \end{aligned} \quad (105)$$

Thus, for $\theta < \theta_P$:

$$\begin{aligned} T'(0) &= P_O(a_0 + a_1 \log \theta, 0) - P_P(a_0 + a_1 \log \theta, 0) \\ &\quad + \frac{d}{dQ} P_P(a_0 + a_1 \log \theta + f(Q), Q) \Big|_{Q=0} \\ &= P_O(a_0 + a_1 \log \theta, 0) \end{aligned} \quad (106)$$

where the second equality follows by observing that all of P_P , $\frac{\partial P_P}{\partial \log Y}$, and $\frac{\partial P_P}{\partial Q}$ are zero for $\theta < \theta_P$. Thus, we have that $T'(0) < 1$ when $P_O(a_0 + a_1 \log \theta, 0) < 1$. Moreover, for $\theta < \theta_P$, we have that: $P_O(a_0 + a_1 \log \theta, 0) - P_O(a_0 + a_1 \log \theta_P, 0) = P_O(P_P^{-1}(0; 0), 0)$. Thus, a sufficient condition for $T'(0) < 1$ for $\theta < \theta_P$ is that $P_O(P_P^{-1}(0; 0), 0) < 1$.

For $\theta > \theta_O$, we have that:

$$\begin{aligned} T'(1) &= P_O(a_0 + a_1 \log \theta + f(1), 1) - P_P(a_0 + a_1 \log \theta + f(1), 1) \\ &\quad + \frac{d}{dQ} P_O(a_0 + a_1 \log \theta + f(1), 1) \Big|_{Q=1} \\ &= P_P(a_0 + a_1 \log \theta + f(1), 1) \end{aligned} \quad (107)$$

where the second equality follows again by observing that all of P_O , $\frac{\partial P_O}{\partial \log Y}$, and $\frac{\partial P_O}{\partial Q}$ are zero

for $\theta > \theta_O$. Hence, we have that $T^\theta(1) < 1$ when $P_P(a_0 + a_1 \log \theta + f(1), 1) > 0$. For $\theta > \theta_O$ we have that $P_P(a_0 + a_1 \log \theta + f(1), 1) - P_P(a_0 + a_1 \log \theta_O + f(1), 1) = P_P(P_O^{-1}(1, 1), 1)$. Thus, a sufficient condition for $T^\theta(1) < 1$ for $\theta > \theta_O$ is that $P_P(P_O^{-1}(1, 1), 1) > 0$. \square

A.7 Proof of Corollary 3

Proof. By Proposition 2, the extremal steady states coexist if and only if $\theta \geq [\theta_O, \theta_P]$, which is non-empty if and only if $\theta_O \leq \theta_P$ which is equivalent to $P_O^{-1}(1; 1) \leq P_P^{-1}(0; 0) \leq f(1)$. \square

A.8 Proof of Lemma 1

Proof. Fix $\theta \geq \Theta$. We first study the SSC-A case. By SSC-A of T we have that there exists $\hat{Q} \in [0, 1]$ such that $T(Q) > Q$ for all $Q \in (0, \hat{Q})$ and $T(Q) < Q$ for all $Q \in (\hat{Q}, 1)$. As T is continuous we have that $T(\hat{Q}) = \hat{Q}$. Consider now some $Q_0 \in (0, 1)$ such that $Q_0 \neq \hat{Q}$. We have that $T(Q_0) > \hat{Q}$ if $Q_0 < \hat{Q}$ and $T(Q_0) < \hat{Q}$ if $Q_0 > \hat{Q}$. Hence, there exists at most one $Q \in (0, 1)$ such that $T(Q) = Q$. Thus, there exist at most three steady states $Q = 0$, $Q = \hat{Q}$, and $Q = 1$.

To find the basins of attraction of these steady states, fix $Q_0 \in (0, 1)$ and consider the sequence $\{T^n(Q_0)\}_{n \in \mathbb{N}}$. For a steady state Q , its basin of attraction is:

$$B(Q) = \left\{ Q_0 \in [0, 1] : \lim_{n \rightarrow \infty} T^n(Q_0) = Q \right\} \quad (108)$$

First, consider $Q_0 \in [0, \hat{Q})$. We now show by induction that $T^n(Q_0) \leq T^{n-1}(Q_0)$ for all $n \in \mathbb{N}$. Consider $n = 1$. We have that $T(Q_0) > Q_0$ as T is SSC-A and $Q_0 < \hat{Q}$. Suppose now that $T^n(Q_0) \leq T^{n-1}(Q_0)$. We have that:

$$T^{n+1}(Q_0) = T(T^n(Q_0)) \leq T(T^{n-1}(Q_0)) = T^n(Q_0) \quad (109)$$

by monotonicity of T , which proves the inductive hypothesis. Observe moreover that the sequence $\{T^n(Q_0)\}_{n \in \mathbb{N}}$ is bounded as $T^n(Q_0) \in [0, 1]$ for all $n \in \mathbb{N}$. Hence, by the monotone convergence theorem, $\lim_{n \rightarrow \infty} T^n(Q_0)$ exists. Toward a contradiction, suppose that $Q_0^* = \lim_{n \rightarrow \infty} T^n(Q_0) > \hat{Q}$. By SSC-A of T we have that $T(Q_0^*) > Q_0^*$, but this contradicts that $Q_0^* = \lim_{n \rightarrow \infty} T^n(Q_0)$. Thus, we have that $Q_0^* \leq \hat{Q}$. Hence, $(0, \hat{Q}) \subset B(\hat{Q})$. Second, consider $Q_0 = \hat{Q}$. We have that $T(\hat{Q}) = \hat{Q}$. Thus, $Q_0^* = \hat{Q}$. Hence, $\hat{Q} \in B(\hat{Q})$. Third, consider $Q_0 \in (\hat{Q}, 1]$. Following the arguments of the first part, we have that $(\hat{Q}, 1) \subset B(\hat{Q})$. Thus, $(0, 1) \subset B(\hat{Q})$. Moreover, if $Q = 0$ or $Q = 1$ are steady states, they can only have basins of attraction in $[0, 1] \setminus B(\hat{Q})$, which implies that they are unstable

and can only have basins of attraction $f0g$ and $f1g$.

The analysis of the SSC-B case follows similarly. By SSC-B of T we have that there exists $\hat{Q} \in [0, 1]$ such that $T(Q) > Q$ for all $Q \in (\hat{Q}, 1)$ and $T(Q) < Q$ for all $Q \in (0, \hat{Q})$. As T is continuous, we have that $T(\hat{Q}) = \hat{Q}$. Consider now some $Q_0 \in (0, 1)$ such that $Q_0 \neq \hat{Q}$. Observe that $T(Q_0) < \hat{Q}$ if $Q_0 < \hat{Q}$ and $T(Q_0) > \hat{Q}$ if $Q_0 > \hat{Q}$. Hence, there exists at most one $Q \in (0, 1)$ such that $T(Q) = Q$. Thus, there exist at most three steady states $Q = 0$, $Q = \hat{Q}$, and $Q = 1$.

To find the basins of attraction of these steady states, first consider $Q_0 \in (0, \hat{Q})$. We now show by induction that $T^n(Q_0) \leq T^{n-1}(Q_0)$ for all $n \in \mathbb{N}$. Consider $n = 1$. We have that $T(Q_0) < Q_0$ as T is SSC-B and $Q_0 < \hat{Q}$. Suppose now that $T^n(Q_0) \leq T^{n-1}(Q_0)$. We have that:

$$T^{n+1}(Q_0) = T(T^n(Q_0)) \leq T(T^{n-1}(Q_0)) = T^n(Q_0) \quad (110)$$

by monotonicity of T , which proves the inductive hypothesis. Observe moreover that the sequence $fT^n(Q_0)g_{n \in \mathbb{N}}$ is bounded as $T^n(Q_0) \in [0, 1]$ for all $n \in \mathbb{N}$. Hence, by the monotone convergence theorem, $\lim_{n \rightarrow \infty} T^n(Q_0)$ exists. Finally, toward a contradiction, suppose that $Q_0^* = \lim_{n \rightarrow \infty} T^n(Q_0) > 0$. By SSC-B of T we have that $T(Q_0^*) < Q_0^*$, but this contradicts that $Q_0^* > 0$. Thus, we have that $Q_0^* = 0$. Hence, $[0, \hat{Q}) \in B(0)$. Second, consider $Q_0 = \hat{Q}$. We have that $T(\hat{Q}) = \hat{Q}$. Thus, $Q_0^* = \hat{Q}$. Hence $\hat{Q} \in B(\hat{Q})$. Third, consider $Q_0 \in (\hat{Q}, 1]$. By the exact arguments of the first part, we have that $(\hat{Q}, 1] \in B(1)$. Observing $B(0)$, $B(\hat{Q})$, and $B(1)$ are disjoint completes the proof. \square

A.9 Proof of Proposition 3

Proof. By Proposition 1 and substituting the form of the shock process from Equation 51, we obtain the formula for the output IRF. For the fraction of optimists, we see that:

$$\begin{aligned} Q_2 &= \hat{Q}_1 P_O(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1) + (1 - \hat{Q}_1) P_P(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1) \\ &> \hat{Q}_1 P_O(a_0 + f(\hat{Q}_1), \hat{Q}_1) + (1 - \hat{Q}_1) P_P(a_0 + f(\hat{Q}_1), \hat{Q}_1) = \hat{Q}_1 \end{aligned} \quad (111)$$

and $Q_t = T_1(\log Y_{t-1}, Q_{t-1})$ for $t \geq 3$ by iterating forward. That Q_t monotonically declines to \hat{Q}_1 follows from Lemma 1 as we are in the SSC-A case. The hump shape is obtained if $\log Y_1 = \log Y_2$. This corresponds to

$$\log Y_1 = a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1) \quad a_0 + f(\hat{Q}_2) = \log Y_2 \quad (112)$$

which rearranges to the desired expression. \square

A.10 Proof of Proposition 4

Proof. We first derive the IRF functions. The formula for the output IRF follows Proposition 3. For the IRF for the fraction of optimists, we simply observe that $Q_0 = Q_1 = 0$ and $Q_2 = P_P(a_0 + a_1 \log \hat{\theta}, 0)$, and that $Q_t = T_1(Q_{t-1})$ for $t \geq 3$ by iterating forward.

We now describe the properties of the IRFs as a function of the size of the initial shock $\hat{\theta}$. First, observe that $Q_2 = P_P(a_0 + a_1 \log \hat{\theta}, 0)$. Thus, we have that $Q_1 = 0$ if and only if $P_P^{-1}(0; 0) \leq a_0 + a_1 \log \hat{\theta}$ which holds if and only if $\hat{\theta} \leq \theta_P$. For any $\hat{\theta} > \theta_P$ it follows that $Q_2 > 0$. As we lie in the SSC class, by Lemma 1, we have that the steady states $Q = 0$, $Q = 1$, and $Q = \hat{Q}_1$ have basins of attraction given by $[0, \hat{Q}_1)$, $(\hat{Q}_1, 1]$, $f\hat{Q}_1g$. Thus, if $Q_2 < \hat{Q}_1$, we have monotone convergence of Q_t to 0. If $Q_2 = \hat{Q}_1$, then $Q_t = \hat{Q}_t$ for all $t \geq 2$. If $Q_2 > \hat{Q}_1$, we have monotone convergence of Q_t to 1. Moreover, the threshold for $\hat{\theta}$ such that $Q_2 = \hat{Q}_1$ is $\exp\left\{\frac{P_P^{-1}(\hat{Q}_1; 0) - a_0}{a_1}\right\}$.

Finally, to find the condition such that the IRF is hump-shaped, we observe that this occurs if and only if $f(Q_2) > a_1 \log \hat{\theta}$ as Q_t is monotonically decreasing for $t \geq 2$, which is precisely the claimed condition. \square

A.11 Proof of Proposition 5

Proof. We prove this result by first constructing fictitious processes for optimism that bound above and below the true optimism process for all realizations of $f\theta_tg_{t \geq 2\mathbb{N}}$ before the stopping time. We can then use this to bound the stopping times' distributions in the sense of first-order stochastic dominance and use this fact to bound the expectations.

First, consider the case where we seek to bound $\tau_{PO} = \min\{t \geq 2 : Q_t = 1, Q_0 = 0\}$. In the model, we have that $Q_{t+1} = T(Q_t, \theta_t)$. Fix a path of fundamentals $f\theta_tg_{t \geq 2\mathbb{N}}$ and define the fictitious \bar{Q} process as:

$$\bar{Q}_{t+1} = \mathbb{1}[T(\bar{Q}_t, \theta_t) = 1] \quad (113)$$

with $\bar{Q}_0 = 0$. We prove by induction that $\bar{Q}_t \leq Q_t$ for all $t \geq 2$. Consider first the base case that $t = 1$:

$$\bar{Q}_1 = \mathbb{1}[T(0, \theta_0) = 1] \leq T(0, \theta_0) = Q_1 \quad (114)$$

Toward the inductive hypothesis, suppose that $\bar{Q}_{t-1} \leq Q_{t-1}$. Then we have that:

$$\bar{Q}_t = \mathbb{1}[T(\bar{Q}_{t-1}, \theta_{t-1}) = 1] \leq \mathbb{1}[T(Q_{t-1}, \theta_{t-1}) = 1] \leq T(Q_{t-1}, \theta_{t-1}) = Q_t \quad (115)$$

where the first inequality follows by the property that $T(\cdot, \theta)$ is a monotone increasing function.

As $\bar{Q}_t = Q_t$ for all $t \geq N$, we have that:

$$\bar{\tau}_{PO} = \min\{t \geq N : \bar{Q}_t = 1, \bar{Q}_0 = 0\} = \min\{t \geq N : Q_t = 1, Q_0 = 0\} = \tau_{PO} \quad (116)$$

Else, we would have at $\bar{\tau}_{PO}$ that $Q_{-\bar{\tau}_{PO}} < \bar{Q}_{-\bar{\tau}_{PO}}$, which is a contradiction.

We now have a pathwise upper bound on τ_{PO} . We now characterize the distribution of the bound. Observe that the possible sample paths for $f\bar{Q}_t g_{t \geq N}$ until stopping are given by the set:

$$G_{PO} = \{f(0^{(n-1)}, 1)g : n \geq 1\} \quad (117)$$

Moreover, conditional on $\bar{Q}_{t-1} = 0$, the distribution of \bar{Q}_t is independent of $f\theta_s g_{s \leq t-1}$. Thus, the fictitious stopping time $\bar{\tau}_{PO}$ has a geometric distribution with parameter given by $P[Q_{t+1} = 1 | Q_t = 0]$. This parameter is given by:

$$\begin{aligned} P[Q_{t+1} = 1 | Q_t = 0] &= P[P_P(a_0 + a_1 \log \theta_t, 0) = 1] \\ &= P\left[\theta_t \exp\left\{\frac{P_P^y(1; 0) - a_0}{a_1}\right\}\right] \\ &= 1 - H\left(\exp\left\{\frac{P_P^y(1; 0) - a_0}{a_1}\right\}\right) \end{aligned} \quad (118)$$

Thus, we have established a stronger result and provided a distributional bound on the stopping time:

$$T_{PO} \stackrel{FOSD}{\leq} \bar{\tau}_{PO} \sim \text{Geo}\left(1 - H\left(\exp\left\{\frac{P_P^y(1; 0) - a_0}{a_1}\right\}\right)\right) \quad (119)$$

An immediate corollary is that:

$$T_{PO} = E[T_{PO}] \quad E[\bar{\tau}_{PO}] = \frac{1}{1 - H\left(\exp\left\{\frac{P_P^y(1; 0) - a_0}{a_1}\right\}\right)} \quad (120)$$

We can apply appropriately adapted arguments for the other case, where we now define:

$$\underline{Q}_{t+1} = 1 - [T(\underline{Q}_t, \theta_t) \leq 0] \quad (121)$$

with $\underline{Q}_0 = 1$. In this case, by an analogous induction we have that $\underline{Q}_t = Q_t$ for all $t \geq N$ for all sequences $f\theta_t g_{t \geq N}$. And so, we have that if \underline{Q}_t has reached 0 then so too has Q_t . Thus,

$T_{OP} = T_{OP}$. The possible sample paths in this case are:

$$G_{OP} = f(1^{(n-1)}, 0)g : n \geq 1g \quad (122)$$

So again the stopping time has a geometric distribution, this time with parameter:

$$\begin{aligned} \mathbb{P}[Q_{t+1} = 0 | Q_t = 1] &= \mathbb{P} \left[\theta_t \exp \left\{ \frac{P_O^y(0;1) a_0 f(1)}{a_1} \right\} \right] \\ &= H \left(\exp \left\{ \frac{P_O^y(0;1) a_0 f(1)}{a_1} \right\} \right) \end{aligned} \quad (123)$$

And so we have:

$$T_{OP} = \frac{1}{H \left(\exp \left\{ \frac{P_O^y(0;1) a_0 f(1)}{a_1} \right\} \right)} \quad (124)$$

It remains to show that these bounds are tight. To do so, we derive a law H such that $Q_t = \bar{Q}_t = \underline{Q}_t$ for all $t \geq \mathbb{N}$. Concretely, define the set:

$$\Theta = \left(1, \exp \left\{ \frac{P_O^y(0;1) a_0 f(1)}{a_1} \right\} \right) \cup \left[\exp \left\{ \frac{P_P^y(1;0) a_0}{a_1} \right\}, 1 \right) \quad (125)$$

and suppose that θ takes values only in this set, where the two sub-intervals are disjoint as $P_O^y(0;1) = P_P^y(1;0) = f(1)$. In this case, starting from $Q_t = 1$, the only possible values for Q_{t+1} are zero and one. Moreover, starting from $Q_t = 0$, the only possible values for Q_{t+1} are zero and one. Thus, in either case, $Q_t = \bar{Q}_t = \underline{Q}_t$ pathwise and $T_{OP} = T_{OP}$ and $T_{PO} = T_{PO}$. It is worth noting that such a distribution can be obtained by considering a limit of normal-mixture distributions. Concretely, suppose that H is derived as a mixture of two normal distributions $N(\mu_A, \sigma^2)$ and $N(\mu_B, \sigma^2)$ for $\mu_A < \exp \left\{ \frac{P_O^y(0;1) a_0 f(1)}{a_1} \right\}$ and $\mu_B > \exp \left\{ \frac{P_P^y(1;0) a_0}{a_1} \right\}$. Taking the limit as $\sigma \rightarrow 0$, the support of H converges to being contained within Θ . \square

A.12 Proof of Proposition 6

Proof. The equilibrium Characterization of Proposition 1 still holds. Moreover, Q_0 is known to all agents. Thus, they can identify θ_0 as:

$$\theta_0 = \frac{\log Y_0 a_0 f(Q_0)}{a_1} \quad (126)$$

Thus, we have that $\lambda_{i1} = P[\mu = \mu_O | \theta_0, \lambda_{i0}]$. Moreover, all agents know that $Q_1 = \int_{[0,1]} \lambda_{i1} di$. Thus, agents can sequentially identify θ_t by observing $fY_j g_t$ by computing:

$$\theta_t = \frac{\log Y_t - a_0 - f(Q_t)}{a_1} \quad (127)$$

Thus, we can describe the evolution of agent's beliefs by computing:

$$\lambda_{i;t+1} = P[\mu = \mu_O | f\theta_j g_{j=1}^t] = \lambda_{i;t+1} = P[\mu = \mu_O | fY_j g_{j=1}^t] \quad (128)$$

By application of Bayes rule, we obtain:

$$\lambda_{i;t+1} = P[\mu = \mu_O | \theta_t, \lambda_{i;t}] = \frac{f_O(\theta_t) \lambda_{i;t}}{f_O(\theta_t) \lambda_{i;t} + f_P(\theta_t) (1 - \lambda_{i;t})} \quad (129)$$

which implies that:

$$\begin{aligned} \frac{\lambda_{i;t+1}}{1 - \lambda_{i;t+1}} &= \frac{f(\log \theta_t | \mu = \mu_O)}{f(\log \theta_t | \mu = \mu_P)} \frac{\lambda_{i;t}}{1 - \lambda_{i;t}} \\ &= \exp \left\{ \frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{\sigma^2} \right\} \frac{\lambda_{i;t}}{1 - \lambda_{i;t}} \end{aligned} \quad (130)$$

Defining $\Omega_{it} = \log \frac{\lambda_{i;t}}{1 - \lambda_{i;t}}$ and $a = E_H \left[\frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{2} \right]$ and $\xi_t = \frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{2}$, we then have that:

$$\begin{aligned} \Omega_{i;t+1} &= \Omega_{i;t} + \frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{\sigma^2} \\ &= \Omega_{i;t} + a + \xi_t \end{aligned} \quad (131)$$

which is a random walk with drift. Iterating, dividing by t , and applying the law of large numbers, we obtain:

$$\frac{\Omega_{i;t}}{t} = \frac{1}{t} \Omega_{i;0} + \frac{t-1}{t} a + \frac{1}{t} \sum_{i=1}^t \xi_i \stackrel{a.s.}{\rightarrow} a \quad (132)$$

Hence, almost surely, we have that $Q_t \rightarrow 1$ if $a > 0$ and $Q_t \rightarrow 0$ if $a < 0$.

Hence, the dynamics are asymptotically described by Proposition 1 with $Q_t = 1$ if $a > 0$ and $Q_t = 0$ if $a < 0$. The resulting properties for output follow immediately. \square

A.13 Proof of Proposition 7

Proof. The dynamics of optimism are characterized by the transition map

$$\begin{aligned} T(Q) &= Q(a_O + b_O Q - cQ^2) + (1 - Q)(a_P + b_P Q - cQ^2) \\ &= a_P + (a_O - a_P + b_P)Q - (c + b_P - b_O)Q^2 \end{aligned} \quad (133)$$

where we define $\omega_0 = a_P$, $\omega_1 = (a_O - a_P + b_P)$, $\omega_2 = (c + b_P - b_O)$ for simplicity. We first show that the dynamics described by T are topologically conjugate to those of the logistic map $\tilde{T}(x) = \eta x(1 - x)$ with

$$\eta = 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)} \quad (134)$$

Two maps $T : [0, 1] \rightarrow [0, 1]$ and $T^\theta : [0, 1] \rightarrow [0, 1]$ are topologically conjugate if there exists a continuous, invertible function $h : [0, 1] \rightarrow [0, 1]$ such that $T^\theta \circ h = h \circ T$. If T is topologically conjugate to T^θ and we know the orbit of T^θ , we can compute the orbit of T via the formula:

$$T^k(Q) = (h^{-1} \circ T^{\theta k} \circ h)(Q) \quad (135)$$

Hence, we can prove the properties of interest using known properties of the map \tilde{T} as well as the mapping from the deeper parameters of T to the parameters of \tilde{T} .

To show the topological conjugacy of T and \tilde{T} , we proceed in three steps:

1. T is topically topologically conjugate to the quadratic map $\hat{T}(Q) = Q^2 + k$ for appropriate choice of k . We guess the following homeomorphism $\hat{h}(Q) = \hat{\alpha} + \hat{\beta}Q$. Plugging \hat{h} in \hat{T} , we have that:

$$\hat{T}(\hat{h}(Q)) = (k + \hat{\alpha}^2) + 2\hat{\alpha}\hat{\beta}Q + \hat{\beta}^2Q^2 \quad (136)$$

Inverting \hat{h} and applying it to this expression yields:

$$\hat{h}^{-1}(\hat{T}(\hat{h}(Q))) = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}} + 2\hat{\alpha}Q + \hat{\beta}Q^2 \quad (137)$$

To verify topological conjugacy, we need to show that $T(Q) = \hat{h}^{-1}(\hat{T}(\hat{h}(Q)))$. Matching coefficients, this is the case if and only if:

$$\omega_0 = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}}, \omega_1 = 2\hat{\alpha}, \omega_2 = -\hat{\beta} \quad (138)$$

We therefore have that:

$$k = \hat{\beta}\omega_0 + \hat{\alpha}(1 - \hat{\alpha}) = \omega_2\omega_0 + \frac{\omega_1}{2} \left(1 - \frac{\omega_1}{2}\right) \quad (139)$$

with $\hat{h}(Q) = \frac{1-\omega_1}{2} - \omega_2Q$.

2. \hat{T} is topologically conjugate to \check{T} for appropriate choice of η . We guess the following homeomorphism $\check{h}(Q) = \check{\alpha} + \check{\beta}Q$. Plugging \check{h} in \check{T} , we obtain:

$$\check{T}(\check{h}(Q)) = \eta(\check{\alpha}(1 - \check{\alpha}) + \check{\beta}(1 - 2\check{\alpha})Q - \check{\beta}^2Q^2) \quad (140)$$

Inverting \check{h} and applying it, we obtain:

$$\check{h}^{-1}(\check{T}(\check{h}(Q))) = \frac{\eta\check{\alpha}(1 - \check{\alpha}) - \check{\alpha}}{\check{\beta}} + \eta(1 - 2\check{\alpha})Q - \eta\check{\beta}Q^2 \quad (141)$$

Matching coefficients, we find:

$$k = \frac{\eta\check{\alpha}(1 - \check{\alpha}) - \check{\alpha}}{\check{\beta}}, 0 = \eta(1 - 2\check{\alpha}), 1 = -\eta\check{\beta} \quad (142)$$

We therefore obtain that:

$$k = \eta(\check{\alpha} - \eta(1 - \check{\alpha})) = \frac{\eta}{2} \left(1 - \frac{\eta}{2}\right) \quad (143)$$

which implies that $\eta = 1 + \sqrt{1 - 4k}$ with $\check{h}(Q) = \frac{1}{2} - \frac{1}{1 + \sqrt{1 - 4k}}Q$.

3. T is topologically conjugate to \check{T} for appropriate choice of η . We now compose the mappings proved in steps 1 and 2 to show

$$T = \hat{h}^{-1} \check{h}^{-1} \check{T} \check{h} \hat{h} \quad (144)$$

with

$$\begin{aligned} \eta &= 1 + \sqrt{1 - 4 \left(\omega_2\omega_0 + \frac{\omega_1}{2} \left(1 - \frac{\omega_1}{2}\right) \right)} = 1 + \sqrt{(\omega_1 - 1)^2 + 4\omega_2\omega_0} \\ &= 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)} \end{aligned} \quad (145)$$

and therefore that T is topologically conjugate to \check{T} .

Having showed the conjugacy of T to \check{T} , we now find bounds on η implied by each case and use this conjugacy to derive the implications for possible dynamics. The following points prove each claim 1-3 in the original Proposition.

1. $\tilde{P}_O = \tilde{P}_P$ and both are monotone. Thus, T is increasing and there cannot be cycles or chaos. This implies that $\eta < 3$ (see [Weisstein, 2001](#), for reference).
2. \tilde{P}_O and \tilde{P}_P are linear. It suffices to show that we can attain $\eta > 3$ but that η must be less than $1 + \sqrt[3]{6}$ (see [Weisstein, 2001](#), for reference). In this case, $c = 0$. This is in addition to the requirements that $\max_{Q \in [0;1]} \tilde{P}_i(Q) = 1$ and $\min_{Q \in [0;1]} \tilde{P}_i(Q) = 0$ for $i \in \{O, P\}$, which can be expressed as:

$$\begin{aligned} \max_{Q \in [0;1]} \tilde{P}_i(Q) &= \max \left\{ a_i, a_i + b_i - c, \left(a_i + \frac{b_i^2}{4c} \right) \mid [0 \leq b_i \leq 2c] \right\} = 1 \\ \min_{Q \in [0;1]} \tilde{P}_i(Q) &= \min \{ a_i, a_i + b_i - c \} = 0 \end{aligned} \quad (146)$$

The maximal value of η consistent with these restrictions can therefore be obtained by solving the following program:

$$\begin{aligned} \max_{(a_O; a_P; b_O; b_P) \in \mathbb{R}^4} & (a_O - a_P + b_P - 1)^2 + 4a_P(b_P - b_O) \\ \text{s.t.} & \max \{ a_O, a_O + b_O \} = 1, \max \{ a_P, a_P + b_P \} = 1 \\ & \min \{ a_O, a_O + b_O \} = 0, \min \{ a_P, a_P + b_P \} = 0 \end{aligned} \quad (147)$$

Exact solution of this program via Mathematica yields that the maximum value is 5. This implies that the maximum value of η is $1 + \sqrt[3]{5} \approx 3.23$, which is greater than 3 but less than $1 + \sqrt[3]{6}$. Moreover, this maximum is attained at $a_O = 0, a_P = 1, b_O = 0, b_P = 1$.

3. No further restrictions on \tilde{P}_O and \tilde{P}_P . We can attain $\eta = 4$ by setting $a_O = a_P = 0, b_O = b_P = 4, c = 4$. Thus, cycles of any period $k \in \mathbb{N}$ and chaotic dynamics can occur (see [Weisstein, 2001](#), for reference).

□

Online Appendix

for “The Macroeconomics of Narratives” by Flynn and Sastry

Contents

B	Model Extensions	76
B.1	Welfare Implications	76
B.2	Multi-Dimensional Narratives and Persistent States	80
B.3	Equilibrium Characterization When Updating Depends on Idiosyncratic Shocks and Idiosyncratic Shocks are Persistent	83
B.4	Narratives in Games and the Role of Higher-Order Beliefs	84
B.5	Model with Firm Dynamics	87
C	Additional Details on Textual Data	89
C.1	Obtaining and Processing 10-Ks	89
C.2	Obtaining and Processing Conference Call Text	90
C.3	Measuring Positive and Negative Words	90
D	Additional Details on Firm Fundamentals Data	92
D.1	Compustat: Data Selection	92
D.2	Compustat: Calculation of TFP	92
E	Additional Details on Model Quantification	96
E.1	Solution of Model With Persistent Fundamentals	96
E.2	Calibration Methodology	98
E.3	Simulation Methodology	99
F	Our Analysis and Shiller’s <i>Narrative Economics</i>	100
F.1	The Modelling of Narratives	100
F.2	Our Work and Shiller’s Seven Propositions	102
F.3	The Perennial Economic Narratives: Our Empirical Findings	105
G	Additional Tables and Figures	106

B Model Extensions

In this appendix, we first study the normative implications of narrative fluctuations (B.1). Second, we extend the baseline model to allow for multidimensional narratives and persistent fundamentals and characterize equilibrium dynamics (B.2). Third, we further extend the baseline model to allow for persistent idiosyncratic fundamentals and narrative updating that depends on realizations of idiosyncratic fundamentals (B.3). Fourth, we highlight the role of higher-order beliefs and show how our analysis could generalize to other settings by deriving a similar law of motion for optimism in abstract, linear beauty contest games (Morris and Shin, 2002) (B.4). Finally, we sketch an extension of our abstract framework to allow for persistent idiosyncratic states and adjustment costs (B.5).

B.1 Welfare Implications

We derive the normative implications of narratives for the economy. To this end, the following result characterizes welfare along any path for the fraction of optimists in the population and the conditions under which a steady state of extreme optimism is preferred to one of extreme pessimism:

Proposition 10 (Narratives and Welfare). *For any path of aggregate optimism $\mathbf{Q} = fQ_t g_{t=0}^1$, aggregate welfare is given by*

$$U(\mathbf{Q}) = U_C \sum_{t=0}^1 \beta^t \exp f(1 - \gamma) f(Q_t) g \tag{148}$$

$$U_L \sum_{t=0}^1 \beta^t (Q_t \exp f(1 + \psi) d_2 g + (1 - Q_t) \exp f(1 + \psi) d_3 f(Q_t) g)$$

for some positive constants U_C , U_L , d_2 and d_3 that are provided in the proof of the result. Thus, there is higher welfare in an optimistic steady state than a pessimistic steady state if and only if

$$\frac{U_C}{U_L} \frac{\exp f(1 - \gamma) f(1) g - 1}{\exp f(1 + \psi) (d_2 + d_3 f(1)) g - 1} > 1 \tag{149}$$

Moreover, when the pessimistic narrative is correctly specified, extreme optimism is welfare-equivalent to an ad valorem price subsidy for intermediate goods producers of:

$$\tau = \exp \left\{ (1 - \omega) \left(\frac{1 + \psi}{\alpha} \frac{\alpha}{\epsilon} + \frac{1}{\epsilon} \right) f(1) \right\} - 1 \tag{150}$$

Proof. We have that welfare for any path of optimism $\mathbf{Q} = fQ_t g_{t2N}$ is given by:

$$U(\mathbf{Q}) = \sum_{t=0}^T \beta^t \left(\mathbb{E}_H \left[\frac{C_t(Q_t, \theta_t)^1}{1 - \gamma} \right] \mathbb{E}_H \left[\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1 + \psi} di \right] \right) \quad (151)$$

By market clearing, we have that $C_t = Y_t$ for all t . Thus, using the formula for equilibrium aggregate output from Proposition 1 and our assumption that $\log \theta_t$ is Gaussian under H , we have that the consumption component of welfare is given by:

$$\begin{aligned} \mathbb{E}_H \left[\frac{C_t^1(Q_t, \theta_t)}{1 - \gamma} \right] &= \mathbb{E}_H \left[\frac{1}{1 - \gamma} \exp f(1 - \gamma) \log Y(Q_t, \theta) g \right] \\ &= \mathbb{E}_H \left[\frac{1}{1 - \gamma} \exp f(1 - \gamma) (a_0 + a_1 \log \theta + f(Q_t)) g \right] \\ &= \frac{1}{1 - \gamma} \exp \left\{ (1 - \gamma) (a_0 + a_1 \mu_H + f(Q_t)) + \frac{1}{2} a_1^2 \sigma_H^2 \right\} \\ &= \frac{1}{1 - \gamma} \exp \left\{ (1 - \gamma) (a_0 + a_1 \mu_H) + \frac{1}{2} a_1^2 \sigma_H^2 \right\} \exp f(1 - \gamma) f(Q_t) g \\ &= U_C \exp f(1 - \gamma) f(Q_t) g \end{aligned} \quad (152)$$

From Proposition 1, we moreover have that labor employed by each firm can be written as:

$$L_{it} = d_1 \log \theta_t + d_2 \lambda_{it} + d_3 f(Q_t) + v_{it} \quad (153)$$

where v_{it} is Gaussian and IID over i . Hence given θ and Q_t :

$$\begin{aligned} &\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1 + \psi} di \\ &= \frac{1}{1 + \psi} (Q_t \exp f(1 + \psi) d_2 g + (1 - Q_t)) \\ &\quad \exp \left\{ (1 + \psi) (d_1 \log \theta + \mu_v + d_3 f(Q_t)) + \frac{1}{2} (1 + \psi)^2 \sigma_v^2 \right\} \end{aligned} \quad (154)$$

Hence, the expectation over θ is given by:

$$\begin{aligned} &\mathbb{E}_H \left[\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1 + \psi} di \right] \\ &= \frac{1}{1 + \psi} (Q_t \exp f(1 + \psi) d_2 g + (1 - Q_t)) \\ &\quad \exp f(1 + \psi) d_3 f(Q_t) g \exp \left\{ (1 + \psi) (d_1 \mu_H + \mu_v) + \frac{1}{2} (1 + \psi)^2 (\sigma_v^2 + d_1^2 \sigma_H^2) \right\} \\ &= U_L (Q_t \exp f(1 + \psi) d_2 g + (1 - Q_t)) \exp f(1 + \psi) d_3 f(Q_t) g \end{aligned} \quad (155)$$

And so total welfare under narrative path \mathbf{Q} is given by:

$$\begin{aligned}
U(\mathbf{Q}) = & U_C \sum_{t=0}^T \beta^t \exp f(1 - \gamma) f(Q_t) g \\
& U_L \sum_{t=0}^T \beta^t (Q_t \exp f(1 + \psi) d_2 g + (1 - Q_t) \exp f(1 + \psi) d_3 f(Q_t) g)
\end{aligned} \tag{156}$$

The final inequality follows by noting that $f(0) = 0$ and rearranging this expression.

Now consider the benchmark model but where, without loss of generality, all agents are pessimistic $Q_t = 0$ and a planner levies an *ad valorem* subsidy. That is, when the consumer price is $p_{it}^C = Y_t^{\frac{1}{\epsilon}} x_{it}^{\frac{1}{\epsilon}}$, the price received by the producer is $p_{it}^P = (1 + \tau) p_{it}^C$. Under this subsidy, each producer's first-order condition is:

$$\begin{aligned}
\log x_{it} = & \frac{1}{\frac{1}{1+\psi} + \frac{1}{\alpha} + \frac{1}{\epsilon}} \left(\log \left(\frac{1}{1+\psi} \right) \log E_{it} \left[\exp \left\{ \frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] \right. \\
& \left. + \log E_{it} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) \log Y_t \right\} \right] \right) + \Xi(\tau)
\end{aligned} \tag{157}$$

where $\Xi(\tau) = \frac{1}{\frac{1}{1+\psi} + \frac{1}{\alpha} + \frac{1}{\epsilon}} \log(1 + \tau)$. By identical arguments to Proposition 1, we have that there is a unique quasi-linear equilibrium, where:

$$\log Y(\theta, \tau) = a_0 + a_1 \log \theta + \frac{1}{1 + \frac{1}{\omega}} \Xi(\tau) \tag{158}$$

and a_0 and a_1 are as in Proposition 1. Hence, in this equilibrium we have that:

$$\log x_{it}(\tau) = \log x_{it}(0) + \frac{1}{1 + \frac{1}{\omega}} \Xi(\tau) \tag{159}$$

Which implies that:

$$\log L_{it}(\tau) = \log L_{it}(0) + \frac{1}{\alpha} \frac{1}{1 + \frac{1}{\omega}} \Xi(\tau) \tag{160}$$

And so, welfare under the subsidy τ is given by:

$$\begin{aligned}
U(\tau) = & U_C \sum_{t=0}^T \beta^t \exp \left\{ \left(1 - \gamma \right) \frac{1}{1 + \frac{1}{\omega}} \Xi(\tau) \right\} \\
& U_L \sum_{t=0}^T \beta^t \exp \left\{ \left(1 + \psi \right) d_3 \frac{1}{1 + \frac{1}{\omega}} \Xi(\tau) \right\}
\end{aligned} \tag{161}$$

as $d_3 = 1$. Hence:

$$U(1) = U(\tau) \tag{162}$$

where τ is such that $\frac{1}{1-\omega} \Xi(\tau) = f(1)$. Hence:

$$\tau = \exp \left\{ (1 - \omega) \left(\frac{1 + \psi}{\alpha} + \frac{1}{\epsilon} \right) f(1) \right\} \tag{163}$$

Completing the proof. □

This result is not only useful for our quantitative analysis as it permits exact computation of welfare without approximation, but also sheds light on the potential for non-fundamental optimism to increase aggregate welfare. Concretely, in the presence of the product market monopoly and labor market monopsony distortions, intermediates goods firms under-hire labor and under-produce goods. As a result, if irrational optimism causes them to produce more output, but not so much that the household over-supplies labor, then it has the potential to be welfare improving. The final part of the proposition then reduces this question to assessing if the implied optimism-equivalent subsidy is less than the welfare-optimal subsidy. Thus, optimism in the economy can serve the role of undoing monopoly frictions and thereby has the potential to be welfare-improving, even when misspecified.

Welfare Effects and Equivalent Subsidies. Here, we use the calibrated model described in Section 7 to quantify the welfare effects of optimism. We calculate the average payoff of the representative household under three scenarios. The first corresponds to the calibrated narrative dynamics in simulation, under the assumption that the pessimistic model is correctly specified.³⁵ The second is a counterfactual scenario with permanent extreme optimism, or $Q_t = 1$ for all t . The third is a counterfactual scenario with permanent extreme pessimism, or $Q_t = 0$ for all t , and an *ad valorem* subsidy of τ to all producers. We use the third scenario to translate the first and second into payoff-equivalent subsidies. We find that both viral and extreme optimism are welfare-increasing relative to extreme pessimism in autarky (i.e, $\tau = 0$). In payoff units, they correspond respectively to equivalent subsidies of 1.33% and 2.59%. Our finding of an overall positive welfare effect for viral optimism suggests that, in our macroeconomic calibration, losses from inducing misallocation are more than compensated by level increases in output.

³⁵Relative to the positive analysis, the normative analysis requires two additional model parameters. We set the idiosyncratic component of productivity to have unit mean and zero variance.

B.2 Multi-Dimensional Narratives and Persistent States

Our baseline model featured two narratives regarding the mean of fundamentals and transitory fundamentals, but we live in a world of many competing narratives regarding many aspects of reality and potentially persistent fundamentals. In this extension, we broaden our analysis to study a class of three-dimensional narratives, which is essentially exhaustive within the Gaussian class. Concretely, suppose that agents believe that the aggregate component of fundamentals follows:

$$\log \theta_t = (1 - \rho)\mu + \rho \log \theta_{t-1} + \sigma \nu_t \quad (164)$$

with $\nu_t \sim N(0, 1)$ and IID. Narratives now correspond to a vector of (μ, ρ, σ) , indexing the mean, persistence and variance of the process for fundamentals. The set of narratives can therefore be represented by $f(\mu_k, \rho_k, \sigma_k)g_{k \in 2K}$. We restrict that agents to place dirac weights on this set, so that they only ever believe one narrative at a time, and let $Q_{t,k}$ be the fraction of agents who believe narrative $(\mu_k, \rho_k, \sigma_k)$ at time t . Finally, we assume that agents face the same signal-to-noise ratio κ , regardless of the narrative that they hold.³⁶ Together, these assumptions ensure that agents' posteriors are normal and place a common weight on narratives when agents form their expectations of fundamentals.

By modifying the functional guess-and-verify arguments from Proposition 1, we characterize equilibrium output in this setting in the following result:

Proposition 11 (Equilibrium Characterization with Multi-Dimensional Narratives and Persistence). *There exists a quasi-linear equilibrium:*

$$\log Y(\log \theta_t, \log \theta_{t-1}, Q_t) = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1}) \quad (165)$$

for some $a_1 > 0$, $a_2 \geq 0$, and f . In this equilibrium, the distribution of narratives in the population evolves according to:

$$Q_{t+1,k} = \sum_{k' \in 2K} Q_{t,k'} P_{k'}(k, a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1}), Q_t) \quad (166)$$

Proof. We follow the same steps as in the proof of Proposition 1, appropriately adapted to this richer setting. First, we guess an equilibrium of the form:

$$\log Y(\log \theta_t, \log \theta_{t-1}, Q_t) = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1}) \quad (167)$$

³⁶Formally, this means that the variance of the noise in agents' signals satisfies $\sigma_{\varepsilon,k}^2 \propto \sigma_k^2$ across narratives.

To verify that this is an equilibrium, we need to compute agents' best replies under this conjecture, aggregate them, and show that they are consistent with this guess once aggregated.

We first find agent's posterior beliefs given narrative weights. Let E denote the standard basis for \mathbb{R}^K with k -th basis vector denoted by $e_k = \underbrace{f0, \dots, 0}_{k-1}, 1, \underbrace{0, \dots, 0g}_{K-k}$. We have that $\lambda_{it} = e_k$ for some $k \in K$. Under this narrative loading, we have that agent's posteriors are given by:

$$\log \theta_{it} / \lambda_{it}, s_{it} \sim N(\log \gamma_i + \kappa s_{it} + (1 - \kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma_{js}^2(\lambda_{it}) + \sigma_-^2) \quad (168)$$

with:

$$\begin{aligned} \mu(e_k, \theta_{t-1}) &= (1 - \rho_k)\mu_k + \rho_k \log \theta_{t-1} \\ \sigma_{js}^2(e_k) &= \frac{1}{\frac{1}{2} + \frac{1}{\epsilon, k}} \quad \kappa = \frac{1}{1 + \frac{\frac{\epsilon, k}{2}}{k}} \end{aligned} \quad (169)$$

for all $k \in K$, where κ does not depend on k as $\sigma_{\epsilon, k}^2 \propto \sigma_k^2$. Hence, we can compute agents' best replies by evaluating:

$$\log \mathbb{E}_{it} \left[\theta_{it}^{\frac{1+\psi}{\alpha}} \right] = \frac{1+\psi}{\alpha} (\log \gamma_i + \kappa s_{it} + (1 - \kappa)\mu(\lambda_{it}, \theta_{t-1})) + \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 (\sigma_{js}^2(\lambda_{it}) + \sigma_-^2) \quad (170)$$

$$\begin{aligned} \log \mathbb{E}_{it} \left[Y_{it}^{\frac{1}{\epsilon}} \right] &= \left(\frac{1}{\epsilon} - \gamma \right) (a_0 + a_1 (\kappa s_{it} + (1 - \kappa)\mu(\lambda_{it}, \theta_{t-1})) + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1})) \\ &+ \frac{1}{2} \left(\frac{1}{\epsilon} - \gamma \right)^2 a_1^2 \sigma_{js}^2(\lambda_{it}) \end{aligned} \quad (171)$$

By substituting this into agents' best replies, we obtain:

$$\begin{aligned} \log x_{it} &= \frac{1}{\frac{1}{1+\psi} + 1} \left[\log \left(\frac{1}{1+\psi} \right) + \frac{1+\psi}{\alpha} [\log \gamma_i + \kappa s_{it} + (1 - \kappa)\mu(\lambda_{it}, \theta_{t-1})] \right. \\ &\quad \left. + \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 (\sigma_{js}^2(\lambda_{it}) + \sigma_-^2) \right. \\ &\quad \left. + \left(\frac{1}{\epsilon} - \gamma \right) [a_0 + a_1 (\kappa s_{it} + (1 - \kappa)\mu(\lambda_{it}, \theta_{t-1})) + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1})] \right. \\ &\quad \left. + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2(\lambda_{it}) \right] \end{aligned} \quad (172)$$

which we observe is conditional normally distributed as $\log x_{it}/\lambda_{it} \sim N(\delta_t(\lambda_{it}), \hat{\sigma}^2)$ with $\hat{\sigma}^2$ as in Equation 84 and:

$$\begin{aligned} \delta_t(e_k) = & \frac{1}{\frac{1}{1+\psi} + 1} \left[\log \left(\frac{1}{\frac{1}{1+\psi}} \right) \right. \\ & + \frac{1+\psi}{\alpha} [\log \gamma_t + \kappa \log \theta_t + (1-\kappa)\mu(e_k, \theta_{t-1})] \\ & + \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 (\sigma_{js}^2(e_k) + \sigma^2) \\ & + \left(\frac{1}{\epsilon} - \gamma \right) [a_0 + a_1 (\kappa \log \theta_t + (1-\kappa)\mu(e_k, \theta_{t-1})) + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1})] \\ & \left. + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2(e_k) \right] \end{aligned} \quad (173)$$

for all $k = K$. Aggregating these best replies, using Equation 85, we obtain that:

$$\begin{aligned} \log Y_t = & \frac{\epsilon}{\epsilon-1} \log \mathbb{E}_t \left[\mathbb{E}_t \left[\exp \left\{ \frac{\epsilon-1}{\epsilon} \log x_{it} \right\} j \lambda_{it} \right] \right] \\ = & \frac{\epsilon}{\epsilon-1} \log \left(\sum_k Q_{t,k} \exp \left\{ \frac{\epsilon-1}{\epsilon} \delta_t(e_k) + \frac{1}{2} \left(\frac{\epsilon-1}{\epsilon} \right)^2 \hat{\sigma}^2 \right\} \right) \\ = & \delta_t(e_1) + \frac{1}{2} \frac{\epsilon-1}{\epsilon} \hat{\sigma}^2 + \frac{\epsilon}{\epsilon-1} \log \left(\sum_k Q_{t,k} \exp \left\{ \frac{\epsilon-1}{\epsilon} (\delta_t(e_k) - \delta_t(e_1)) \right\} \right) \end{aligned} \quad (174)$$

where $\hat{\sigma}^2$ is a constant, $\delta_t(e_1)$ depends linearly on $\log \theta_t$ and $\log \theta_{t-1}$ and $\delta_t(e_k) - \delta_t(e_1)$ does not depend on $\log \theta_t$ for all $k = K$ and can therefore be written as $\delta_{k1}(\theta_{t-1})$. Moreover, by matching coefficients, we obtain that a_1 is the same as in the proof of Proposition 1. And we find that f must satisfy:

$$f(Q, \theta_{t-1}) = \frac{1}{\frac{1}{1+\psi} + 1} \gamma f(Q, \theta_{t-1}) + \frac{\epsilon}{\epsilon-1} \log \left(\sum_k Q_{t,k} \exp \left\{ \frac{\epsilon-1}{\epsilon} \delta_{k1}(\theta_{t-1}) \right\} \right) \quad (175)$$

and so:

$$f(Q, \theta_{t-1}) = \frac{1}{1 - \frac{1}{\frac{1}{1+\psi} + 1} \gamma} \log \left(\sum_k Q_{t,k} \exp \left\{ \frac{\epsilon-1}{\epsilon} \delta_{k1}(\theta_{t-1}) \right\} \right) \quad (176)$$

Completing the proof. \square

In the multidimensional narrative case with persistence, the past value of fundamentals interacts non-linearly with the cross-sectional narrative distribution in affecting aggregate

output. However, without more structure, the properties of the dynamics generated by this multi-dimensional system are essentially unrestricted.

B.3 Equilibrium Characterization When Updating Depends on Idiosyncratic Shocks and Idiosyncratic Shocks are Persistent

We now extend the analysis from Section B.2 to the case where agents' idiosyncratic states drive narrative updating and are persistent. Concretely, in that setting, we let P_{k^θ} depend on $(Y_t, Q_t, \tilde{\theta}_{it})$ and idiosyncratic productivity shocks evolve according to an AR(1) process:

$$\log \tilde{\theta}_{it} = \rho_- \log \tilde{\theta}_{i,t-1} + \zeta_{it} \quad (177)$$

where $0 < \rho_- < 1$ and $\zeta_{it} \sim N(0, \sigma^2)$. We let F_- denote the stationary distribution of $\tilde{\theta}_{it}$, which coincides with the cross-sectional marginal distribution of $\tilde{\theta}_{it}$ for all $t \geq N$.

The additional theoretical complication these two changes induce is that the marginal distribution of narratives Q_t is now insufficient for describing aggregate output. This is because narratives λ_{it} and idiosyncratic fundamentals $\tilde{\theta}_{it}$ are no longer independent as λ_{it} and $\tilde{\theta}_{it}$ both depend on $\tilde{\theta}_{i,t-1}$. The relevant state variable is now the joint distribution of narratives and idiosyncratic productivity $\check{Q}_t \in \Delta(\Lambda \times \mathbb{R})$. We denote the marginals as Q_t and F_- , and the conditional distribution of narratives given $\tilde{\theta}$ as $\check{Q}_{t,kj^-} = \frac{Q_{t,k}(\cdot)}{f_{\tilde{\theta}}(\cdot)}$.

Proposition 12 (Equilibrium Characterization with Multi-Dimensional Narratives, Aggregate and Idiosyncratic Persistence, and Idiosyncratic Narrative Updating). *There exists a quasi-linear equilibrium:*

$$\log Y(\log \theta_t, \log \theta_{t-1}, \check{Q}_t) = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(\check{Q}_t, \theta_{t-1}) \quad (178)$$

for some $a_1 > 0$, $a_2 \leq 0$, and f .

Proof. This proof follows closely that of Proposition 11. Under narrative loading λ_{it} , we have that the agent's posterior regarding $\log \theta_{it}$ is given by:

$$\log \theta_{it} | \tilde{\theta}_{i,t-1}, \lambda_{it}, s_{it} \sim N\left(\log \gamma_i + \rho_- \tilde{\theta}_{i,t-1} + \kappa s_{it} + (1 - \kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma_{js}^2(\lambda_{it}) + \sigma^2\right) \quad (179)$$

where $\mu(\lambda_{it}, \theta_{t-1})$, κ , and $\sigma_{js}^2(\lambda_{it})$ are as in Proposition 11. Then substitute $\log \gamma_i + \rho_- \tilde{\theta}_{i,t-1}$ for $\log \gamma_i$ and follow the Proof of Proposition 11 until the aggregation step (Equation 174).

We now instead have that:

$$\begin{aligned}
\log Y_t &= \frac{\epsilon}{\epsilon - 1} \log \mathbb{E}_t \left[\mathbb{E}_t \left[\exp \left\{ \frac{\epsilon - 1}{\epsilon} \log x_{it} \right\} f_{\tilde{\theta}_{it-1}, \lambda_{it}} \right] \right] \\
&= \frac{\epsilon}{\epsilon - 1} \log \mathbb{E}_t \left[\exp \left\{ \frac{\epsilon - 1}{\epsilon} \delta_t(e_k, \tilde{\theta}_{it-1}) + \frac{1}{2} \left(\frac{\epsilon - 1}{\epsilon} \right)^2 \hat{\sigma}^2 \right\} \right] \\
&= \frac{\epsilon}{\epsilon - 1} \log \left(\int \sum_k \check{Q}_{t:kj^-} \exp \left\{ \frac{\epsilon - 1}{\epsilon} \delta_t(e_k, \tilde{\theta}) + \frac{1}{2} \left(\frac{\epsilon - 1}{\epsilon} \right)^2 \hat{\sigma}^2 \right\} dF_-(\tilde{\theta}) \right) \quad (180) \\
&= \delta_t(e_1, 0) + \frac{1}{2} \frac{\epsilon - 1}{\epsilon} \hat{\sigma}^2 \\
&\quad + \frac{\epsilon}{\epsilon - 1} \log \left(\int \sum_k \check{Q}_{t:kj^-} \exp \left\{ \frac{\epsilon - 1}{\epsilon} \left(\delta_t(e_k, \tilde{\theta}) - \delta_t(e_1, 0) \right) \right\} dF_-(\tilde{\theta}) \right)
\end{aligned}$$

Again, $\hat{\sigma}^2$ is a constant and $\delta_t(e_1, 0)$ depends linearly on $\log \theta_t$, $\log \theta_{t-1}$, $\log \tilde{\theta}_{it-1}$ and $\delta_t(e_k, \tilde{\theta})$. $\delta_t(e_1)$ does not depend on $\log \theta_t$ for all $k = K$. Thus, we may write it as $\delta_{k1}(\theta_{t-1}, \tilde{\theta})$. Again, a_1 is the same as in Proposition 1. By the same steps as in Proposition 11, we then have that:

$$f(\check{Q}, \theta_{t-1}) = \frac{\frac{1}{1 + \frac{\epsilon}{\alpha} \alpha + \frac{1}{\epsilon}}}{1} \log \left(\int \sum_k \check{Q}_{t:kj^-} \exp \left\{ \frac{\epsilon - 1}{\epsilon} \delta_{k1}(\theta_{t-1}, \tilde{\theta}) \right\} dF_-(\tilde{\theta}) \right) \quad (181)$$

Completing the proof. □

B.4 Narratives in Games and the Role of Higher-Order Beliefs

We have studied a micro-founded business-cycle model, but the basic insights extend much more generally to abstract, linear beauty contest games. Importantly, these settings provide us with an ability to disentangle the dual roles of narratives in affecting both agents' first-order and higher-order beliefs about fundamentals.

Concretely, suppose that agents' best replies are given by the following beauty contest form (see, *e.g.*, Morris and Shin, 2002):

$$x_{it} = \alpha \mathbb{E}_{it}[\theta_t] + \beta \mathbb{E}_{it}[Y_t] \quad (182)$$

where $\alpha > 0$ and $\beta \in [0, 1)$. This linear form for best replies is commonly justified by (log-)linearization of some underlying best response function (see, *e.g.*, Angeletos and Pavan, 2007). For example, log-linearization of the agents' best replies in the baseline model of this section yields such an equation with $\beta = \omega$ and all variables above standing in for their

log-counterparts. Moreover, suppose that aggregation is linear so that $Y_t = \int_{[0,1]} x_{it} di$. This can similarly be justified via an appropriate first-order expansion of non-linear aggregators. Finally, we let the structure of narratives be as before.

Toward characterizing equilibrium, we define the average expectations operator:

$$\bar{E}_t[\theta_t] = \int_{[0,1]} E_{it}[\theta_t] di \quad (183)$$

and the higher-order average expectations operator for $k \geq 1$ as:

$$\bar{E}_t^k[\theta_t] = \int_{[0,1]} E_{it}[\bar{E}_t^{k-1}[\theta_t]] di \quad (184)$$

Moreover, we observe by recursive substitution that equilibrium aggregate output is given by:

$$Y_t = \alpha \sum_{k=1}^{\infty} \beta^k \bar{E}_t^k[\theta_t] \quad (185)$$

We can therefore solve for the unique equilibrium by computing the hierarchy of higher-order expectations. We can do this in closed-form by observing that agents' idiosyncratic first-order beliefs are given by:

$$E_t[\theta_t | s_{it}, \lambda_{it}] = \kappa s_{it} + (1 - \kappa)(\lambda_{it} \mu_O + (1 - \lambda_{it}) \mu_P) \quad (186)$$

which allows us to compute average first-order expectations of fundamentals as:

$$\bar{E}_t[\theta_t] = \kappa \theta_t + (1 - \kappa)(Q_t \mu_O + (1 - Q_t) \mu_P) \quad (187)$$

which is a weighted-average between true fundamentals and the average impact of narratives on agents' priors. By taking agents' expectations over this object and averaging, we compute higher-order average expectations as:

$$\bar{E}_t^k[\theta_t] = \kappa^k \theta_t + (1 - \kappa^k)(Q_t \mu_O + (1 - Q_t) \mu_P) \quad (188)$$

which is again a weighted average between the state and agents' priors, but now with a geometrically increasing weight on narratives as we consider higher-order average beliefs.

The following result characterizes aggregate output and agents' best replies in the unique equilibrium:

Proposition 13 (Narratives and Higher-Order Beliefs). *There exists a unique equilibrium. In this unique equilibrium, aggregate output is given by:*

$$Y_t = \frac{\alpha}{1-\beta} \left(\frac{(1-\beta)\kappa}{1-\beta\kappa} \theta_t + \frac{1-\kappa}{1-\beta\kappa} (Q_t \mu_O + (1-Q_t) \mu_P) \right) \quad (189)$$

Moreover, agents' actions follow:

$$\begin{aligned} x_{it} = & \alpha \frac{1}{1-\beta\kappa} [\kappa \theta_t + \kappa \varepsilon_{it} + (1-\kappa) (\lambda_{it} \mu_O + (1-\lambda_{it}) \mu_P)] \\ & + \beta \frac{\alpha}{1-\beta} \frac{1-\kappa}{1-\beta\kappa} (Q_t \mu_O + (1-Q_t) \mu_P) \end{aligned} \quad (190)$$

Proof. To substantiate the arguments in the main text, by aggregating Equation 182, we obtain that:

$$Y_t = \alpha \bar{E}_t[\theta_t] + \beta \bar{E}_t[Y_t] \quad (191)$$

Thus, by recursive substitution k times we obtain that:

$$Y_t = \alpha \sum_{j=1}^k \beta^{j-1} \bar{E}_t^j[\theta_t] + \beta^k \bar{E}_t^k[Y_t] \quad (192)$$

Moreover, we have that:

$$\bar{E}_t^j[\theta_t] = \kappa^j \theta_t + (1-\kappa^j) (Q_t \mu_O + (1-Q_t) \mu_P) \quad (193)$$

and thus that:

$$\begin{aligned} \alpha \sum_{j=1}^k \beta^{j-1} \bar{E}_t^j[\theta_t] &= \alpha \sum_{j=1}^k \beta^{j-1} (\kappa^j \theta_t + (1-\kappa^j) (Q_t \mu_O + (1-Q_t) \mu_P)) \\ &= \alpha \sum_{j=1}^k \beta^{j-1} (Q_t \mu_O + (1-Q_t) \mu_P) + \alpha \beta^{-1} \sum_{j=1}^k (\beta \kappa)^j [\theta_t - (Q_t \mu_O + (1-Q_t) \mu_P)] \end{aligned} \quad (194)$$

Hence:

$$\begin{aligned} \lim_{k \rightarrow \infty} \alpha \sum_{j=1}^k \beta^{j-1} \bar{E}_t^j[\theta_t] &= \frac{\alpha}{1-\beta} (Q_t \mu_O + (1-Q_t) \mu_P) + \\ & \frac{\alpha \kappa}{1-\beta \kappa} [\theta_t - (Q_t \mu_O + (1-Q_t) \mu_P)] \end{aligned} \quad (195)$$

We therefore have that there is a unique equilibrium if $\lim_{k \rightarrow \infty} \beta^k \bar{E}_t^k[Y_t] = 0$. Hellwig and

Veldkamp (2009) show in Proposition 1 of their supplementary material that all equilibria differ on a most a measure zero set of fundamentals. In this setting, this implies that $\lim_{\kappa \rightarrow 1} \beta^k \bar{E}_t^k[Y_t] = c$ for some $c \in \mathbb{R}$ for almost all $\theta \in \Theta$. Hence, the equilibrium is given by:

$$\begin{aligned} Y_t &= \frac{\alpha}{1-\beta} (Q_t \mu_O + (1-Q_t) \mu_P) + \frac{\alpha \kappa}{1-\beta \kappa} [\theta_t - (Q_t \mu_O + (1-Q_t) \mu_P)] + c \\ &= \frac{\alpha}{1-\beta} \left(\frac{(1-\beta) \kappa}{1-\beta \kappa} \theta_t + \frac{1-\kappa}{1-\beta \kappa} (Q_t \mu_O + (1-Q_t) \mu_P) \right) + c \end{aligned} \quad (196)$$

But then we have that $c = 0$ by computing $\lim_{\kappa \rightarrow 1} \beta^k \bar{E}_t^k[Y_t] = 0$ under this equilibrium.

Finally, to solve for individual actions under this equilibrium, we compute:

$$\begin{aligned} x_{it} &= \alpha E_{it}[\theta_t] + \beta E_{it}[Y_t] \\ &= \alpha E_{it}[\theta_t] + \beta E_{it} \left[\frac{\alpha}{1-\beta} \left(\frac{(1-\beta) \kappa}{1-\beta \kappa} \theta_t + \frac{1-\kappa}{1-\beta \kappa} (Q_t \mu_O + (1-Q_t) \mu_P) \right) \right] \\ &= \left(\alpha + \beta \frac{\alpha}{1-\beta} \frac{(1-\beta) \kappa}{1-\beta \kappa} \right) E_{it}[\theta_t] + \beta \frac{\alpha}{1-\beta} \frac{1-\kappa}{1-\beta \kappa} (Q_t \mu_O + (1-Q_t) \mu_P) \\ &= \alpha \frac{1}{1-\beta \kappa} (\kappa s_{it} + (1-\kappa) (\lambda_{it} \mu_O + (1-\lambda_{it}) \mu_P)) \\ &\quad + \beta \frac{\alpha}{1-\beta} \frac{1-\kappa}{1-\beta \kappa} (Q_t \mu_O + (1-Q_t) \mu_P) \end{aligned} \quad (197)$$

Completing the proof. \square

This result allows us to see how narratives affect output by propagating up through the hierarchy of higher-order beliefs. Concretely, we have that the static impulse response of output to a contemporaneous shock to the fraction of optimists in the population is given by:

$$\frac{\partial Y_t}{\partial Q_t} = \frac{\alpha}{1-\beta} \frac{1-\kappa}{1-\beta \kappa} (\mu_O - \mu_P) \quad (198)$$

which is composed of the relative importance of fundamentals $\frac{1}{1-\beta}$, the impact of prior beliefs on the entire hierarchy of higher-order beliefs about exogenous and endogenous outcomes $\frac{1}{1-\beta \kappa}$ and the difference between the two narratives $\mu_O - \mu_P$. Moreover, this result shows how our earlier regression equation holds in equilibrium in this linearized setting.

B.5 Model with Firm Dynamics

We now sketch an augmentation of our baseline conceptual model of the firm from which we derived our earlier estimating equations (see Appendix A.1) to allow for persistent idiosyn-

cratic states and adjustment costs. This allows us to more formally justify why controlling for firm productivity and lagged labor is sufficient to account for the presence of adjustment costs to first-order.

In every period t , each firm i still takes an action $x_{it} \in X$. Their objective function still takes as an input their action, aggregate outcomes $Y_t \in Y$, and aggregate fundamentals θ_t (which in analogy to the previous appendix sections, we allow to follow a first-order (continuous) Markov process). However, they now have idiosyncratic fundamentals $\tilde{\theta}_{it}$, which follow a first-order (continuous) Markov process. Moreover, their actions are subject to adjustment costs $\Phi : \mathbb{R} \rightarrow \mathbb{R}_+$ equal to $\Phi(x - x_{t-1})$ when their last action was x_{t-1} . Thus, we let their flow utility be $u(x, Y, \theta, \tilde{\theta}) = \Phi(x - x_{t-1})$. The firm discounts the future at rate $\beta_i \in [0, 1)$. The aggregate state variables in period t are the distribution of x_{it-1} in the population F_{t-1}^x , the distribution of narratives in the population Q_t , and the level of current and past aggregate fundamentals θ_t and θ_{t-1} . Thus, equilibrium aggregate output is described by some function $\hat{Y}(F_{t-1}^x, Q_t, \theta_t, \theta_{t-1})$. Moreover, observe at time t that the following are the state variables for a firm: (i) the level of idiosyncratic productivity in the previous period $\tilde{\theta}_{it-1}$ (ii) the level of aggregate productivity in the previous period θ_{t-1} (iii) the firm's action in the previous period x_{it-1} (iv) the narrative entertained by the agent λ_{it} (v) their current signal about fundamentals s_{it} , and (vi) the additional aggregate states (F_{t-1}^x, Q_t) .

We can therefore represent any firm policy function as:

$$x_{it} = g(x_{it-1}, \theta_{t-1}, \tilde{\theta}_{it-1}, F_{t-1}^x, Q_t, \lambda_{it}, s_{it}) \quad (199)$$

If this is differentiable, we may linearize it to obtain:

$$x_{it} = \gamma_i + \chi_t + \sum_{k=1}^K \delta_k \lambda_{k:it} + \gamma \theta_{it-1} + \omega x_{it-1} + \varepsilon_{it} \quad (200)$$

where the aggregate fixed effect now absorbs $(F_{t-1}^x, Q_t, \theta_t, \theta_{t-1})$, θ_{it-1} capture agents' idiosyncratic expectations of future fundamentals, and x_{it-1} captures their adjustment costs.

C Additional Details on Textual Data

C.1 Obtaining and Processing 10-Ks

Here, we describe our methodology for obtaining and processing raw data on 10-K filings. We start with raw html files downloaded directly from the SEC’s EDGAR (Electronic Data Gathering, Analysis, and Retrieval) system. Each of these files corresponds to a single 10-K filing. Each file is identified by its unique accession number. In its heading, each file also contains the end-date for the period the report concerns (*e.g.*, 12/31/2018 for a FY 2018 ending in December), and a CIK (Central Index Key) firm identifier from the SEC. We use standard linking software provided by Wharton Research Data Services (WRDS) to link CIK numbers and fiscal years to the alternative firm identifiers used in data on firm fundamentals and stock prices. We have, in our original dataset, 182,259 files.

We follow the following steps to turn each document, now identified by firm and year, into a bag-of-words representation:

1. *Cleaning raw text.* We first translate the document into unformatted text. Specifically, we follow the following steps in order:
 - (a) Removing hyperlinks and other web addresses
 - (b) Removing html formatting tags encased in the brackets <>
 - (c) Making all text lowercase
 - (d) Removing extra spaces, tabs, and new lines.
 - (e) Removing punctuation
 - (f) Removing non-alphabetical characters
2. *Removing stop words.* Following standard practice, we remove “stop words” which are common in English but do not convey specific meaning in our analysis. We use the default English stop word list in the `nl tk` Python package. Example stopwords include articles (“a”, “the”), pronouns (“I”, “my”), prepositions (“in”, “on”), and conjunctions (“and”, “while”).
3. *Lemmatizing documents.* Again following standard practice, we use lemmatization software to reduce words to their common roots. We use the default English-language lemmatizer of the `spacy` Python package. The lemmatizer uses both the word’s identity and its content to transform sentences. For instance, when each is used as a verb, “meet,” “met,” and “meeting” are commonly lemmatized to “meet.” But if the software predicts that “meeting” is used as a noun, it will be lemmatized as the noun “meeting.”

4. *Estimating a bigram model.* We estimate a bigram model to group together commonly co-occurring words as single two-word phrases. We use the `phrases` function of the `gensim` package. The bigram modeler groups together words that are almost always used together. For instance, if our original text data set were the 10-Ks of public firms Nestlé and General Mills, the model may determine that “ice” and “cream,” which almost always appear together, are part of a bigram “ice_cream.”
5. *Computing the bag of words representation.* Having now expressed each document as a vector of clean words (*i.e.*, single words and bigrams), we simply collapse these data to frequencies.

Finally, note that our procedure uses *all* of the non-formatting text in the 10K. This includes all sections of the documents, and does not limit to the Management Discussion and Analysis (MD&A) section. This is motivated by the fact that management’s discussion is not limited to one section [SEC \(2011\)](#). Moreover, prior literature has found that textual analysis of the entire 10-K versus the MD&A section tends to closely agree, and that limiting scope to the MD&A section has limited practical benefits due to the trade-off of limiting the amount of text per document ([Loughran and McDonald, 2011](#)).

C.2 Obtaining and Processing Conference Call Text

We obtain the full text of sales and earnings conference calls from 2002 to 2014 from the Fair Disclosure (FD) Wire service. The original sample includes 261,034 documents, formatted as raw text. We next subset to documents that have reported firm names and stock tickers, which are automatically associated with documents by Lexis Nexis. When matches are probabilistic, we use the first (highest probability) match.³⁷ We finally restrict to firms that are listed on one of three US stock exchanges: the NYSE, the NASDAQ, or the NYSE-MKT (Small Cap). We finally connect tickers to the firm identifiers in our fundamentals data using the master cross-walk available on Wharton Research Data Services (WRDS). These operations together reduce the sample size to 158,810 calls. We clean these data by conducting steps 1-3 described above in [Appendix C.1](#). We then calculate positive word counts, negative word counts, and optimism exactly as described in the main text for the 10-K data.

C.3 Measuring Positive and Negative Words

To calculate sets of positive and negative 10K words, we use the updated dictionary available online at [McDonald \(2021\)](#) as of June 2020. This dictionary includes substantial updates

³⁷In the essentially zero-measure cases in which there is a tie, we take the alphabetically first ticker.

relative to the dictionaries associated with the original [Loughran and McDonald \(2011\)](#) publication. These changes are reviewed in the *Documentation* available at [McDonald \(2021\)](#).

The Loughran-McDonald dictionary includes 2345 negative words and 347 positive words. The dictionary is constructed to include multiple forms of each relevant word. For instance, the first negative root “abandon” is listed as: “abandon,” “abandoned,” “abandoning,” “abandonment,” “abandonments,” and “abandons.” To ensure consistency with our own lemmatization procedure, we first map each unique word to all of its possible lemmas using the `getAllLemmas` function of the `lemminflect` Python package, which is an extension to the `spacy` package we use for lemmatization. We then construct a new list of negative words by combining the original list of negative words with all new, unique lemmas to which a negative word mapped (and similarly for positive words). This procedure results in new lists of 2411 negative words and 366 positive words, which map exactly to the words that appear in our cleaned bag of words representation. We list the top ten most common positive and negative words from this cleaned set in [Table A1](#). In particular, to make the table most legible, we first associate words with their lemmas, then count the sum of document frequencies for each associated word (which may exceed one), and then print the most common word associated with the lemma.

D Additional Details on Firm Fundamentals Data

D.1 Compustat: Data Selection

Our data selection criteria and variable definitions are identical to those used in [Flynn and Sastry \(2021\)](#). In this Appendix, we review essential points. We refer the reader to the Appendix material of [Flynn and Sastry \(2021\)](#) for certain details.

Our dataset is Compustat Annual Fundamentals. Our main variables of interest are defined in Appendix Table [A13](#). We restrict the sample to firms based in the United States, reporting statistics in US Dollars, and present in the “Industrial” dataset. We exclude firms whose 2-digit NAICS is 52 (Finance and Insurance) or 22 (Utilities). This filter eliminates firms in two industries that, respectively, may have highly non-standard production technology and non-standard market structure.

We summarize our definitions of major “input and output” variables in Appendix Table [A13](#). For labor choice, we measure the number of employees. For materials expenditure, we measure the sum of reported variable costs (COGS) and sales and administrative expense (xsga) net of depreciation (dp).³⁸ As in [Ottonello and Winberry \(2020\)](#) and [Flynn and Sastry \(2021\)](#), we use a perpetual inventory method to calculate the value of the capital stock. We start with the first reported observation of gross value of plant, property, and equipment and adding net investment or the differences in net value of plant, property, and equipment. Note that, because all subsequent analysis is conditional on industry-by-time fixed effects, it is redundant at this stage to deflate materials and capital expenditures by industry-specific deflators.

We categorize the data into 44 sectors. These are defined at the 2-digit NAICS level, but for the Manufacturing (31-33) and Information (51) sectors, which we classify at the 3-digit level to achieve better balance of sector size. More summary information about these industries is provided in Appendix F of [Flynn and Sastry \(2021\)](#).

D.2 Compustat: Calculation of TFP

When calculating firms’ Total Factor Productivity, we restrict attention to a subset of our sample that fulfills the following inclusion criteria:

1. Sales, material expenditures, and capital stock are strictly positive;
2. Employees exceed 10;

³⁸A small difference from [Flynn and Sastry \(2021\)](#) is that, in assessing the firms’ costs and later calculating TFP, we do not “unbundle” materials expenditures on labor and non-labor inputs using supplemental data on annual wages.

3. Acquisitions as a proportion of assets (aqc over at) does not exceed 0.05.

The first ensures that all companies meaningfully report all variables of interest for our production function estimation; the second applies a stricter cut-off to eliminate firms that are very small, and lead to outlier estimates of productivity and choices. The third is a simple screening device for large acquisitions which may spuriously show up as large innovations in firm choices and/or productivity.

Our method for recovering total factor productivity is based on cost shares. In brief, we use cost shares for materials and labor to back out production elasticities, and treat the elasticity of capital as the implied “residual” given an assumed mark-up $\mu > 1$ (in our baseline, $\mu = 4/3$) and constant physical returns to scale. The exact procedure is the following:

1. For all firms in industry j , calculate the estimated materials share:

$$\text{Share}_{M:j^0} = \frac{\sum_{i:j(i)=j^0} \sum_t \text{MaterialExpenditure}_{it}}{\sum_{i:j(i)=j^0} \sum_t \text{Sales}_{it}} \quad (201)$$

2. If $\text{Share}_{M:j^0} < \mu^{-1}$, then set

$$\begin{aligned} \alpha_{M:j^0} &= \mu \cdot \text{Share}_{M:j^0} \\ \alpha_{K:j^0} &= 1 - \alpha_{M:j^0} - \alpha_{L:j^0} \end{aligned} \quad (202)$$

3. Otherwise, adjust shares to match the assumed returns to scale, or set

$$\begin{aligned} \alpha_{M:j^0} &= 1 \\ \alpha_{K:j^0} &= 0 \end{aligned} \quad (203)$$

To translated our production function estimates into productivity, we calculate a “Sales Solow Residual” $\tilde{\theta}_{it}$ of the following form:

$$\log \tilde{\theta}_{it} = \log \text{Sales}_{it} - \frac{1}{\mu} (\alpha_{M:j(i)} \log \text{MatExp}_{it} + \alpha_{K:j(i)} \log \text{CapStock}_{it}) \quad (204)$$

We finally define our estimate $\log \hat{\theta}$ as the previous net of industry-by-time fixed effects

$$\log \hat{\theta}_{it} = \log \tilde{\theta}_{it} - \chi_{j(i):t} \quad (205)$$

Theoretical Interpretation. The aforementioned method recovers physical productivity (“TFPQ”) under the assumptions, consistent with our quantitative model, that firms operate

constant-returns to scale technology and face an isoelastic, downward-sloping demand curve of *known* elasticity (equivalently, they charge a known markup). The idea is that, given the known markup, we can impute firms' (model-consistent) costs as a fixed fraction of sales and then calculate the theoretically desired cost shares. Here, we describe the simple mathematics.

There is a single firm i operating in industry j with technology

$$Y_i = \theta_i M_i^{\alpha_j} K_i^{1-\alpha_j} \quad (206)$$

They act as a monopolist facing the demand curve

$$p_i = Y_i^{-\frac{1}{\epsilon}} \quad (207)$$

for some inverse elasticity $\epsilon > 1$. Observe that this is, up to scale, the demand function faced by monopolistically competitive intermediate goods producers in our model. The firm's revenue is therefore $p_i Y_i = Y_i^{1-\frac{1}{\epsilon}}$. Finally, the firm can buy materials at industry-specific price q_j and rent capital at rate r_j . The firm's program for profit maximization is therefore

$$\max_{M_i, K_i} \left\{ (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} - q_j M_i - r_j K_i \right\} \quad (208)$$

We first justify our formulas for the input shares (Equation 202). To do this, we solve for the firm's optimal input choices. This is a concave problem, in which first-order conditions are necessary and sufficient. These conditions are

$$\begin{aligned} q_j &= M_i^{-1} \alpha_j \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \\ r_j &= K_i^{-1} (1 - \alpha_j) \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \end{aligned} \quad (209)$$

Re-arranging, and substituting in $p_i = Y_i^{-\frac{1}{\epsilon}}$, we derive

$$\begin{aligned} \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{q_j M_i}{p_i Y_i} \\ 1 - \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{r_j K_i}{p_i Y_i} \end{aligned} \quad (210)$$

Or, in words, that the materials elasticity is $\frac{\epsilon}{\epsilon - 1}$ times the ratio of materials input expenditures to sales. Observe also that, by re-arranging the two first-order conditions, we can

write expressions for production and the price

$$Y = \left(\left(\frac{\epsilon - 1}{\epsilon} \right) \theta_i \left(\frac{\alpha_j}{q_j} \right) \left(\frac{1 - \alpha_j}{r_j} \right)^{1 - \alpha_j} \right) p = \left(\frac{\epsilon}{\epsilon - 1} \right) \theta_i^{-1} \left(\frac{q_j}{\alpha_j} \right)^{\alpha_j} \left(\frac{r_j}{1 - \alpha_j} \right)^{1 - \alpha_j} \quad (211)$$

and observe that $\theta_i^{-1} \left(\frac{q_j}{\alpha_j} \right)^{\alpha_j} \left(\frac{r_j}{1 - \alpha_j} \right)^{1 - \alpha_j}$ is the firm's marginal cost. Hence, we can define $\mu = \frac{\epsilon}{\epsilon - 1} > 1$ as the firm's markup and write the shares as required:

$$\alpha = \mu \frac{q_j M_i}{p_i Y_i} \quad (212)$$

Finally, we now apply Equations 204 and 205 to calculate productivity. Assume that we observe materials expenditure $q_j M_i$ and capital value $p_{\kappa j} K_i$, where $p_{\kappa j}$ is an (unobserved) price of capital. We find

$$\log \tilde{\theta}_i = \left(1 - \frac{1}{\epsilon} \right) (\log \theta_i - \alpha \log q_j - (1 - \alpha) \log p_{\kappa j}) \quad (213)$$

We finally observe that the industry-level means are

$$\chi_j = \left(1 - \frac{1}{\epsilon} \right) (\log \bar{\theta}_j - \alpha \log q_j - (1 - \alpha) \log p_{\kappa j}) \quad (214)$$

where $\log \bar{\theta}_j$ is the mean of $\log \theta_i$ over the industry. Hence,

$$\log \hat{\theta}_i = \left(1 - \frac{1}{\epsilon} \right) (\log \theta_i) \quad (215)$$

or our measurement captures physical TFP, up to scale.

E Additional Details on Model Quantification

E.1 Solution of Model With Persistent Fundamentals

We first provide the exact solution of the model when fundamentals follow an AR(1) process. We build on the analysis of Online Appendix B.2, which allows for (among other features) persistent fundamentals.

Law of Motion for Output. Log aggregate productivity follows the process

$$\log \theta_t = (1 - \rho)\mu + \rho \log \theta_{t-1} + \sigma \nu_t \quad (216)$$

with $\nu_t \sim N(0, 1)$ IID. We continue to assume, as in our main analysis, that there are two narratives associated with high and low values of μ , $\mu_O > \mu_P$, while the true value is $\mu = 0$. Proposition 11 establishes that equilibrium can be written as (f does not depend on θ_{t-1} here as all agents believe persistence is ρ)

$$\log Y_t = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t) \quad (217)$$

where we normalize $a_0 = 0$. We define the *fundamental component of output* as $\log Y_t - f(Q_t)$:

$$\log Y_t^f = a_1 \log \theta_t + a_2 \log \theta_{t-1} \quad (218)$$

Subtracting $\rho \log Y_{t-1}^f$ from both sides, the above becomes an ARMA(1, 1) process:

$$\log Y_t^f - \rho \log Y_{t-1}^f = a_1 \sigma \nu_t + a_2 \sigma \nu_{t-1} \quad (219)$$

It remains to solve for the coefficients (a_1, a_2). In particular, Equations 173 and 174 gives the fixed-point equation which these coefficients must solve. We can simplify these fixed point equations considerably in the case with optimism and pessimism about means and compute δ_{tk} for $k \in \{O, P\}$:

$$\begin{aligned} \delta_{tk} = \frac{1}{\frac{1+\psi}{1+\psi} + 1} & \left[\log \left(\frac{1-\kappa}{1+\psi} \right) + \frac{1+\psi}{\alpha} [\log \gamma_i + \kappa \log \theta_t + (1-\kappa)((1-\rho)\mu_k + \rho \log \theta_{t-1})] \right. \\ & \frac{1}{2} \left(\frac{1+\psi}{\alpha} \right)^2 (\sigma_{js}^2 + \sigma_{\epsilon}^2) + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma \right)^2 \sigma_{js}^2 \\ & \left. + \left(\frac{1}{\epsilon} - \gamma \right) [a_0 + a_1 (\kappa \theta_t + (1-\kappa)((1-\rho)\mu_k + \rho \log \theta_{t-1})) + a_2 \log \theta_{t-1} + f(Q_t)] \right] \end{aligned} \quad (220)$$

Here, we have used the fact that posterior variances and perceived persistence are the same for the two narratives, and the fact that $\mu(e_k, \theta_{t-1}) = (1-\rho)\mu_k + \rho \log \theta_{t-1}$. Therefore,

$$\alpha \delta^{OP} := \delta_{t,O} - \delta_{t,P} = \frac{1}{\frac{1+\psi}{\alpha} + 1} \left(\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \right) (1-\kappa)(1-\rho)(\mu_O - \mu_P) \quad (221)$$

is the (time-invariant) average difference in actions between optimists and pessimists, as we identify in the data.

Next, taking $\delta_t(e_1) = \delta_{t,P}$ and Q_t as the fraction of optimists, we write Equation 174 as

$$\log Y_t = \delta_{t,P} + \frac{1}{2} \frac{\epsilon}{\epsilon-1} \hat{\sigma}^2 + \frac{\epsilon}{\epsilon-1} \log \left(Q_t \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha \delta^{OP} \right\} \right) \quad (222)$$

Substituting in the expression for $\delta_{t,P}$, we can write the above up to a constant C that does not depend on $(\log \theta_t, \log \theta_{t-1}, Q_t)$ as

$$\begin{aligned} \log Y_t = C &+ \frac{1}{\frac{1+\psi}{\alpha} + 1} \left[\frac{1+\psi}{\alpha} \kappa + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \kappa \right] \log \theta_t \\ &+ \frac{1}{\frac{1+\psi}{\alpha} + 1} \left[\frac{1+\psi}{\alpha} (1-\kappa) \rho + \left(\frac{1}{\epsilon} - \gamma \right) a_1 (1-\kappa) \rho \right] \log \theta_{t-1} \\ &+ \frac{\epsilon}{\epsilon-1} \log \left(Q_t \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha \delta^{OP} \right\} \right) \end{aligned} \quad (223)$$

To obtain the coefficients in our desired representation, first note that we can write

$$f(Q_t) = \frac{\epsilon}{\epsilon-1} \log \left(Q_t \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha \delta^{OP} \right\} \right) - \frac{\epsilon}{\epsilon-1} \log \left(\frac{1}{2} \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha \delta^{OP} \right\} \right) \quad (224)$$

This is the same form as our main analysis, with a normalization such that $f(1/2) = 0$. Next, from matching coefficients, and noting the definition of $\omega = (1/\epsilon - \gamma)/((1+\psi - \alpha)/\alpha + 1/\epsilon)$,

$$a_1 = \frac{1}{1 - \kappa \omega} \frac{\frac{1+\psi}{\alpha} \kappa}{\frac{1+\psi}{\alpha} + 1} \quad (225)$$

Finally, from matching coefficients for a_2 ,

$$a_2 = \frac{1}{\frac{1+\psi}{\alpha} + 1} \left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \right] (1-\kappa) \rho \quad (226)$$

Updating Rule. We use the Linear-Associative-Viral Updating rule introduced as Example 1, with a normalization. More specifically, we assume transition probabilities

$$P_O^H(\log Y, Q) = [u + r \log Y + sQ + C_P]_0^1 \quad \text{and} \quad P_P^H(\log Y, Q) = [u + r \log Y + sQ + C_P]_0^1 \quad (227)$$

We choose C_P such that an economy with neutral fundamentals ($\log \theta_t = \log \theta_{t-1} = 0$) and equal optimists and pessimists ($Q = 1/2$) continues to have equal optimists and pessimists. Specifically, this implies $C_P = \frac{1-s}{2}$

E.2 Calibration Methodology

To calibrate the model, we proceed in four steps.

1. *Setting macro parameters.* We set $(\epsilon, \gamma, \psi, \alpha)$ to values based on the literature. We describe our selection in Section 7.1 and report the parameters in Table 7.
2. *Calibrating the effect of optimism on output.* We observe that, conditional on $(\epsilon, \gamma, \psi, \alpha)$ and an estimate of δ^{OP} , we have identified $f(Q_t)$ as defined in Equation 224. We take our estimate of δ^{OP} from column 1 in Table 1. This regression identifies δ^{OP} for the reasons described in Corollary 1.
3. *Calibrating the updating rule.* The coefficients of the LAV updating model is estimated in column 1 of Table 5.
4. *Calibrating the statistical properties of fundamentals (κ, ρ, σ) .*
 - (a) *Computing fundamental output.* We construct a cyclical component of output, $\log \hat{Y}_t$, as band-pass filtered U.S. real GDP (Baxter and King, 1999).³⁹ We apply our estimated function \hat{f} to our measured time series of optimism (see Figure 1) to get an estimated optimism component of output. we then calculate

$$\log \hat{Y}_t^f = \log \hat{Y}_t + \hat{f}(\hat{Q}_t) \quad (228)$$

- (b) *Estimating the ARMA representation.* Using our 24 annual observations of $\log \hat{Y}_t^f$, we estimate a Gaussian-errors ARMA(1,1) model via maximum likelihood. Our point estimates are

$$\log \hat{Y}_t^f - 0.086 \log \hat{Y}_{t-1}^f = .0078(\nu_t + .32 \nu_{t-1}) \quad (229)$$

³⁹Specifically, we filter to post-war quarterly U.S. real GDP data (Q1 1947 to Q1 2022). We use a lead-lag length of 12 quarters, a low period of 6 quarters, and a high period of 32 quarters. We then average these data to the annual level.

This implies $\rho = 0.086$, $a_1\sigma = .0078$, and $a_2\sigma = .32$. ρ is therefore identified immediately.

- (c) *Calibrating (κ, σ)* . We search non-linearly for values of (κ, σ) that satisfy $a_1\sigma = .0078$ and $a_2\sigma = .32$. There is a unique such pair, reported in Table 7, which also is therefore the maximum likelihood estimate of (κ, σ) .

E.3 Simulation Methodology

For all simulation results, we simulate a time path of 5,000,000 years starting from the initial condition $Q_1 = 0.5$ and $\log \theta_1 = 0.0$. We then calculate statistics as time-averages from year 1,000,000 onwards.

F Our Analysis and Shiller’s *Narrative Economics*

Shiller (2017, 2020) introduces the notion of narrative economics and identifies “Seven Propositions of Narrative Economics” as a basis for the theoretical and empirical investigation of narratives. In this section, we discuss our work, how our modelling of narratives relates to Shiller’s ideas, and the relationship of our modelling, measurement, and results with these propositions. In the process, we highlight how these propositions have informed our analysis, discuss how our analysis contributes new insights, and propose avenues for future work to more fully understand narratives and the macroeconomy.

F.1 The Modelling of Narratives

We first describe how our modelling and measurement of narratives are designed to capture the salient features of narratives that Shiller (2020) introduces in the preface:

In using the term narrative economics, I focus on two elements: (1) the word-of-mouth contagion of ideas in the form of stories and (2) the efforts that people make to generate new contagious stories or to make stories more contagious. First and foremost, I want to examine how narrative contagion affects economic events. The word narrative is often synonymous with story. But my use of the term reflects a particular modern meaning given in the Oxford English Dictionary: “a story or representation used to give an explanatory or justificatory account of a society, period, etc.” Expanding on this definition, I would add that stories are not limited to simple chronologies of human events. A story may also be a song, joke, theory, explanation, or plan that has emotional resonance and that can easily be conveyed in casual conversation.

To map this verbal definition to our framework (see Section 2 for the formal details and notation), consider the following simple verbal rationale for our modelling approach. There is a latent space of economic fundamentals (demand for goods) and endogenous outcomes (aggregate output). An agent has some beliefs about economic fundamentals and corresponding endogenous outcomes (π). They are told the following simple story by another agent about the economy: “I didn’t hire (x) because aggregate output (Y) will be low because demand (θ) is weak.” This might cause the agent to believe this story that demand is weak and adopt this narrative (placing weight λ on the implied distribution of fundamentals). Moreover, if many of their friends tell them the story, they might be more positively inclined to believe it. Of course, the agent doesn’t listen to the story blindly: they can see if demand was previously low (and might even have information about demand from their personal economic

activities l) and might regard such a story as fanciful if their own experience contradicts this claim. At the end of this process of contemplation, they update their own weight on the narrative (via P) and arrive at their new belief (π^l).

Of course, the actual realization of output depends on the circulation of narratives in the population (Q). If an agent believes the “demand is weak” narrative, then they will curtail their hiring. Knowing this, other agents – even if they do not believe that latent demand is weak – will believe that others will curtail their hiring, so that realized demand will be weak. Then, knowing this, all agents further cut hiring. This paradox of thrift induces a hierarchy of higher-order expectations regarding realized demand induced by the distribution of narratives. This converges to a fixed point ($Y(Q) : \Theta \rightarrow \mathcal{Y}$) describing the mapping of demand into aggregate output under the prevailing circulation of narratives.

Thus, while the primitive narrative began as a story about the strength of demand, in equilibrium it takes on a meaning as not only describing exogenous economic outcomes, but also endogenous economic outcomes. Concretely, given an equilibrium mapping, the narrative endogenously induces the joint belief $N \in \Delta(\Theta \times \mathcal{Y})$ given by $N(\theta, Y) = N(\theta) \mathbb{1}[Y = Y(\theta)]$. Importantly, the distribution of narratives Q and endogenous outcomes Y then shape the distribution of narratives tomorrow Q^l . The resulting joint dynamics of narratives and endogenous outcomes are the subject of the theoretical and quantitative analysis of this paper.

To ensure that we measure narratives, trace their impact on decisions and study their spread, we operationalize this empirically by measuring narratives in agents’ use of language by employing natural language processing methods that we have described in Section 3.1. This allows us to use our framework to test if these text-based proxies for narratives shape actions and spread across agents. As described in Section 4, we find strong evidence of these premises.

We are, however, essentially silent about the more fundamental determinants of how something comes to be a narrative or what makes a narrative contagious. As a result, we do not speak to the issue of narrative generation suggested by Shiller. We do have one empirical result that hints that firms use narratives to persuade financial analysts. In our IBES strategy, we found that optimistic firms significantly overestimate their sales relative to pessimistic firms. However, we found much weaker evidence that analysts believe that firms are performing this overestimation. As a result, we take this as tentative evidence that firms manage to persuade analysts of their predictions. We view further exploration of this issue as an interesting angle for future work.

F.2 Our Work and Shiller’s Seven Propositions

Proposition 1: Epidemics Can Be Fast or Slow, Big or Small. The model developed in this paper allows for various speeds of narrative dynamics as well as their size and economic impact. Shiller, drawing on the epidemiology literature, identifies two parameters as particularly important in determining these features: the contagion rate and the recovery rate. Viewed through this lens, our structural model from Section 5 postulates a recovery rate of $1 - P_O(\log Y_t, Q_t)$ and a contagion rate of $P_P(\log Y_t, Q_t)$. Thus, the fundamental parameters determining stubbornness u , associativeness r and virality s are key determinants of the speed and size of narrative epidemics within our model.

Yet further, by moving beyond a purely epidemiological model and studying the two-way feedback between narratives and endogenous outcomes, we endogenize these rates as equilibrium outcomes by characterizing the equilibrium map $Q_t, \theta_t \mathcal{V} Y_t$. Thus, the parameters of P_O and P_P as well as those determining the information and strategic interaction in the economy affect the contagion and recovery rates in ways that we have characterized. Most interestingly, beyond affecting the quantitative features of narrative dynamics (such as speed and size), accounting for the dynamic complementarity of narratives affects their qualitative features. Concretely, these same parameters delineate whether the economy is stable, has a unique steady state, features hysteresis, or hump-shaped and discontinuous impulse responses.

Moreover, we have used our measurement and empirical strategies to place empirical discipline on these parameters. By so doing, we have been able to provide ballpark figures for the likely quantitative importance of the narratives we have uncovered in our data. Future work may lever alternative data sources and identification strategies to study different narratives or more precisely estimate the parameters that we have studied.

Proposition 2: Important Economic Narratives May Comprise a Very Small Percentage of Popular Talk. Our empirical analysis found that very little of the total variation in narrative discussion is at the aggregate level (See Appendix Table A5). For example, only 1% of the variation in optimism is captured in the aggregate time series. Indeed, even for the 75% percentile of our estimated topic narratives, the fraction of variance explained by the time series is less than 10%. Thus, movements in the intensity with which narratives are discussed appear to be largely a cross-sectional phenomenon. As we have shown, this does not at all mean that aggregate movements are unimportant: measured movements in aggregate optimism can account for between 1/6 and 1/3 of GDP movements over significant economic events. Thus, just as idiosyncratic income risk is much larger than aggregate GDP risk, idiosyncratic narrative variation is much larger than aggregate narrative

variation.

This echoes the observational account of Shiller that important economic narratives may not feature prominently in popular talk and underlines the conclusion that even if movements in aggregate narratives are not a large fraction of what agents think or discuss, they can nevertheless be critical for understanding economic fluctuations.

Proposition 3: Narrative Constellations Have More Impact Than Any One Narrative. A central concept in Shiller’s analysis is that of the *narrative constellation*: a grouping of narratives around some basic idea that reinforces contagion. This is a concept about which we are theoretically silent. However, our empirical analysis is designed to allow for the possibility of narrative constellations. Take our analysis of optimism, for example. Our measure does not necessarily capture one coherent economic narrative regarding the strength of the outlook of the economy. What it instead captures is the total sentiment expressed by firms, averaging across the various underlying narratives that they may be adopting at any one instant. We investigate formally the extent to which our data support this by using our more granular narratives as instruments for optimism (see Appendix Table A9, columns 4) and find similar results to our baseline analysis.

Moreover, our analysis of Shiller’s narratives allows us to pick up narrative constellations to the precise extent that Shiller discusses the words comprising the underlying narratives in these constellations in his own analysis. Finally, our topic analysis allows us to pick up narrative constellations to the extent that narratives are used jointly in individual documents. Thus, while we neither explicitly model constellations nor test the reinforcing effect of constellations hypothesized by Shiller, we do account for their existence in our measurement and empirical exercises. We view the analysis of this hypothesis as an interesting avenue for future work.

Proposition 4: The Economic Impact of Narratives May Change Through Time. Shiller suggests that the impact of economic narratives has the potential to change through time. We evaluate this hypothesis in the context of our study in Section 4.1 of how measured optimism affects hiring. Specifically, we consider our baseline regression model

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(l);t} + \varepsilon_{it} \quad (230)$$

Our baseline estimate, in column 1 of Table 1, is $\hat{\delta}^{OP} = 0.0355$. We now consider a variant in which the coefficient δ^{OP} varies for each year 1996 τ 2019 in our sample:⁴⁰

$$\Delta \log L_{it} = \sum_{=1996}^{2017} \delta^{OP}(\text{opt}_{it} \mid [t = \tau]) + \gamma_i + \chi_{j(t):t} + \varepsilon_{it} \quad (231)$$

We show the coefficient series of δ^{OP} graphically in Figure A7. We observe no strong pattern of a trend or business cycle in the coefficient estimates. We interpret this as evidence that the main narrative studied in our empirical analysis has relatively stable effects on decisions over time.

Proposition 5: Truth Is Not Enough to Stop False Narratives. Shiller emphasizes that narrative epidemics can take place even when patently divorced from fundamentals. Our theoretical analysis shares this feature. Namely, when the virality of a narrative is high, this can swamp any adverse effects on contagion stemming from outcomes that do not fit the narrative. This is made especially clear by our Proposition 2, in which multiple steady states of narrative penetration can coexist even when one narrative (or even both narratives) are false.

Proposition 6: Contagion of Economic Narratives Builds on Opportunities for Repetition. Increased exposure to a narrative is likely to cause agents to pick it up. We find evidence that agents are both more likely to retain a narrative they currently have and that exposure to others holding the same narrative increases the chance that an agent both picks up and retains a narrative. These findings are consistent with Shiller’s hypothesis that oft-repeated narratives are more likely to result in epidemics. However, we do not explore the idea that repeated exposure through time is likely to increase the persistence or virality of a narrative. We view this as an interesting avenue for future work.

Proposition 7: Narratives Thrive on Attachment: Human Interest, Identity, and Patriotism. We do not investigate the idea that more interesting narratives are more likely to be viral in this paper. Studying this idea requires a deeper model for the origins of the stubbornness, associativeness and virality of narratives, which we do not attempt to provide. We merely measure these parameters and take them as given. Of course, this renders our analysis vulnerable to a form of the Lucas critique: if a policymaker attempted to use our estimates as a guide for how they could affect the economy via manipulating narratives, these coefficients could change if they fail to mimic the human interest, identity, or patriotism that drove attachment to the narrative. While this issue is unimportant for our positive analysis,

⁴⁰The number of firms with data reported for 2019 is very small, so our sample essentially ends in 2018.

an understanding of the deeper origins of narrative success is an interesting avenue for future work – especially if a policymaker were to seek to guide narratives to achieve certain economic outcomes.

F.3 The Perennial Economic Narratives: Our Empirical Findings

Shiller (2020) identifies nine perennial economic narratives based on an historical analysis. These narratives correspond to:

1. Panic versus Confidence
2. Frugality versus Conspicuous Consumption
3. The Gold Standard versus Bimetallism
4. Labor-Saving Machines Replace Many Jobs
5. Automation and Artificial Intelligence Replace Almost All Jobs
6. Real Estate Booms and Busts
7. Stock Market Bubbles
8. Boycotts, Profiteers, and Evil Businesses
9. The Wage-Price Spiral and Evil Labor Unions

We have measured the presence of these narratives in firm language and studied which of these matter for firm decisionmaking. Under our baseline LASSO specification, we found that two of these narratives are relevant for firm hiring: Labor-Saving Machines Replace Many Jobs and Stock Market Bubbles. Discussion of both is positively associated with firm hiring (see column 1 of Appendix Table A9). Moreover, we find evidence of virality and stubbornness in firms holding onto these narratives (see Appendix Figure A6).

G Additional Tables and Figures

List of Tables

A1	The Twenty Most Common Positive and Negative Words	107
A2	The Twenty Most Common Words for Each Shiller Chapter	108
A3	The Ten Most Common Words for Each Selected Topic	109
A4	Persistence and Cyclicity of Narratives	110
A5	Variance Decomposition of Narratives	113
A6	The Effect of Optimism on Hiring, Conference-Call Measurement	114
A7	The Effect of Optimism on All Inputs	114
A8	The Effect of Optimism on Stock Prices: High-Frequency Analysis	115
A9	The Effect of Narratives on Hiring: All Narratives	116
A10	The Virality of Optimism, CEO Change Strategy	117
A11	The Virality and Associativeness of Optimism, Instrumented With Other Narratives	118
A12	Sensitivity Analysis for Quantitative Analysis	120
A13	Data Definitions in Compustat	120

List of Figures

A1	Time-Series for Positive, Negative, and Their Difference	110
A2	Time-Series for Shiller's Perennial Economic Narratives	111
A3	Time-Series for the Selected LDA Topics	112
A4	Heterogeneous Effects of Optimism on Hiring	113
A5	Dynamic Relationship between Optimism and Firm Fundamentals, Conference-Call Measurement	115
A6	The Virality and Associativeness of Other Identified Narratives	119

Table A1: The Twenty Most Common Positive and Negative Words

Positive		Negative	
Lemmatized Word	Total DF	Lemmatized Word	Total DF
well	1.92	loss	1.75
good	1.82	decline	1.22
benefit	1.56	disclose	1.21
high	1.54	subject	1.18
gain	1.34	terminate	1.15
advance	1.25	omit	1.14
achieve	1.14	defer	1.11
improve	1.05	claim	1.10
improvement	0.88	concern	1.07
opportunity	0.86	default	1.03
satisfy	0.86	limitation	0.96
lead	0.84	delay	0.93
enhance	0.73	deficiency	0.85
enable	0.71	fail	0.85
able	0.63	losses	0.82
best	0.62	damage	0.81
gains	0.56	weakness	0.79
improvements	0.53	adversely	0.76
opportunities	0.49	against	0.75
resolve	0.49	impairment	0.72

Notes: The twenty most common lemmatized words among the 230 positive words and 1354 negative words. Total document frequencies are the sums of document frequencies for all words that map to that root (*e.g.*, “benefit,” “benefited,” “benefiting,” “benefits,” “benefitted,” and “benefitting” for “benefit”), and can thus sum to more than one.

Table A2: The Twenty Most Common Words for Each Shiller Chapter

Panic		Frugality		Gold Standard		Labor-Saving Machines		Automation and AI		Real Estate		Stock Market		Boycotts		Wage-Price Spiral	
Lem. Word	DF	Lem. Word	DF	Lem. Word	DF	Lem. Word	DF	Lem. Word	DF	Lem. Word	DF	Lem. Word	DF	Lem. Word	DF	Lem. Word	DF
bank	1.29	help	0.44	standard	1.46	replac	1.21	replac	1.21	price	2.29	chapter	0.37	price	2.29	countri	0.53
consum	0.80	hous	0.37	book	0.63	technolog	1.16	appear	0.69	appear	0.69	peopl	0.21	pro t	1.70	labor	0.37
appear	0.69	home	0.35	money	0.48	produc	1.07	show	0.58	real	0.47	specul	0.19	good	0.81	in at	0.32
forecast	0.59	famili	0.35	paper	0.28	appear	0.69	labor	0.37	nd	0.45	drop	0.13	consum	0.80	strong	0.32
show	0.58	buy	0.32	run	0.26	power	0.68	question	0.31	estat	0.45	play	0.11	start	0.50	union	0.30
economi	0.40	display	0.23	peopl	0.21	book	0.63	world	0.29	hous	0.37	great	0.11	fall	0.41	world	0.29
run	0.26	peopl	0.21	metal	0.14	save	0.45	human	0.28	citi	0.37	depress	0.10	buy	0.32	ask	0.22
suggest	0.22	lost	0.20	eastern	0.14	labor	0.37	univers	0.26	home	0.35	peak	0.10	in at	0.32	wage	0.22
peopl	0.21	job	0.18	almost	0.13	electr	0.32	run	0.26	land	0.34	today	0.10	world	0.29	peopl	0.21
around	0.20	explain	0.11	depress	0.10	innov	0.30	vacat	0.24	buy	0.32	news	0.08	wage	0.22	mile	0.15
weather	0.20	phrase	0.11	young	0.10	world	0.29	ask	0.22	world	0.29	warn	0.07	peopl	0.21	war	0.12
con d	0.19	unemploy	0.11	today	0.10	human	0.28	suggest	0.22	movement	0.24	get	0.07	cut	0.16	great	0.11
concept	0.17	great	0.11	depositor	0.08	problem	0.26	peopl	0.21	never	0.21	decad	0.06	shop	0.14	depress	0.10
wall	0.15	depress	0.10	man	0.08	run	0.26	autom	0.21	peopl	0.21	reaction	0.06	behavior	0.12	peak	0.10
gur	0.12	fashion	0.10	want	0.07	mechan	0.25	worker	0.21	specul	0.19	memori	0.06	war	0.12	happen	0.09
unemploy	0.11	postpon	0.10	popular	0.07	wage	0.22	around	0.20	tend	0.18	storm	0.06	campaign	0.12	recess	0.09
great	0.11	cultur	0.10	decad	0.06	peopl	0.21	job	0.18	social	0.16	newspap	0.05	play	0.11	tri	0.09
depress	0.10	car	0.10	gold	0.06	worker	0.21	machin	0.16	search	0.16	moral	0.04	explain	0.11	popular	0.07
grew	0.10	consumpt	0.09	newspap	0.05	job	0.18	answer	0.15	popul	0.15	narr	0.04	great	0.11	get	0.07
today	0.10	fault	0.09	idea	0.05	machin	0.16	war	0.12	apart	0.14	sudden	0.04	depress	0.10	wrote	0.07

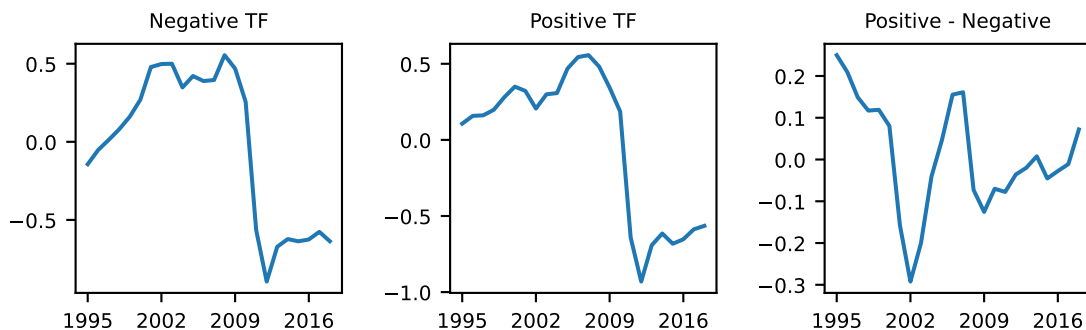
Notes: The twenty most common lemmatized words among the 100 words that typify each Shiller narrative. Total document frequencies are the sums of document frequencies for all words that map to that root (*e.g.*, “benefit,” “benefited,” “benefiting,” “benefits,” “benefitted,” and “benefitting” for “benefit”), and can thus sum to more than one.

Table A3: The Ten Most Common Words for Each Selected Topic

Topic 1		Topic 2		Topic 3		Topic 4		Topic 5		Topic 6	
Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight
lease	0.047	business	0.052	value	0.088	advertising	0.029	nancial	0.051	stock	0.049
tenant	0.042	public	0.025	fair	0.082	retail	0.028	control	0.02	compensation	0.039
landlord	0.03	combination	0.024	loss	0.024	brand	0.018	internal	0.019	tax	0.039
lessee	0.017	merger	0.023	investment	0.024	credit	0.018	material	0.013	share	0.028
rent	0.016	class	0.015	asset	0.022	consumer	0.017	a ect	0.012	income	0.023
lessor	0.014	o er	0.014	debt	0.02	distribution	0.016	o cer	0.011	average	0.019
property	0.012	share	0.013	gain	0.019	card	0.015	base	0.01	expense	0.018
term	0.011	account	0.011	credit	0.019	marketing	0.015	information	0.01	asset	0.016
day	0.009	ordinary	0.01	level	0.017	food	0.013	make	0.01	outstanding	0.016
provide	0.008	private	0.01	nancial	0.016	store	0.013	business	0.01	weight	0.015
Topic 7		Topic 8		Topic 9		Topic 10		Topic 11			
Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight		
gaming	0.035	debt	0.039	reorganization	0.048	court	0.038	technology	0.018		
service	0.029	credit	0.039	bankruptcy	0.047	settlement	0.027	revenue	0.017		
network	0.022	facility	0.037	plan	0.044	district	0.021	development	0.015		
wireless	0.021	senior	0.028	predecessor	0.036	certain	0.019	business	0.013		
local	0.019	interest	0.026	successor	0.027	litigation	0.016	customer	0.012		
cable	0.015	agreement	0.021	chapter	0.021	action	0.016	stock	0.012		
provide	0.014	cash	0.019	asset	0.019	complaint	0.012	product	0.012		
equipment	0.013	rate	0.016	court	0.018	damage	0.011	support	0.009		
access	0.013	term	0.016	cash	0.016	approximately	0.011	market	0.009		
video	0.012	certain	0.014	certain	0.014	case	0.01	service	0.008		

Notes: The ten most common words (lemmatized bigrams) in each topic estimated by LDA and selected by our LASSO procedure as relevant for hiring (see Section 4.1). Weights correspond to relative importance for scoring the document and add up to one over all words.

Figure A1: Time-Series for Positive, Negative, and Their Difference



Notes: These variables are used in the construction of binary optimism.

Table A4: Persistence and Cyclicity of Narratives

Narrative N_t	Correlation with			
	N_{t-1}	u_{t-1}	u_t	u_{t+1}
Optimism	0.754	-0.283	-0.368	-0.287
Topic Narratives (25th)	0.810	-0.430	-0.307	-0.210
Topic Narratives (median)	0.935	0.003	-0.143	-0.092
Topic Narratives (75th)	0.965	0.339	0.252	0.077
Shiller Narratives (25th)	0.792	-0.379	-0.379	-0.367
Shiller Narratives (median)	0.805	0.043	0.088	-0.034
Shiller Narratives (75th)	0.884	0.541	0.422	0.246

Notes: Calculated with annual data from 1995 to 2017 for optimism and the topics, and 1995 to 2017 for the Shiller Narratives. u_t is the U.S. unemployment rate. The quantiles for Shiller Narratives and Topic Narratives are the quantiles of the distribution of the variable in that column within each set of narratives.

Figure A2: Time-Series for Shiller's Perennial Economic Narratives

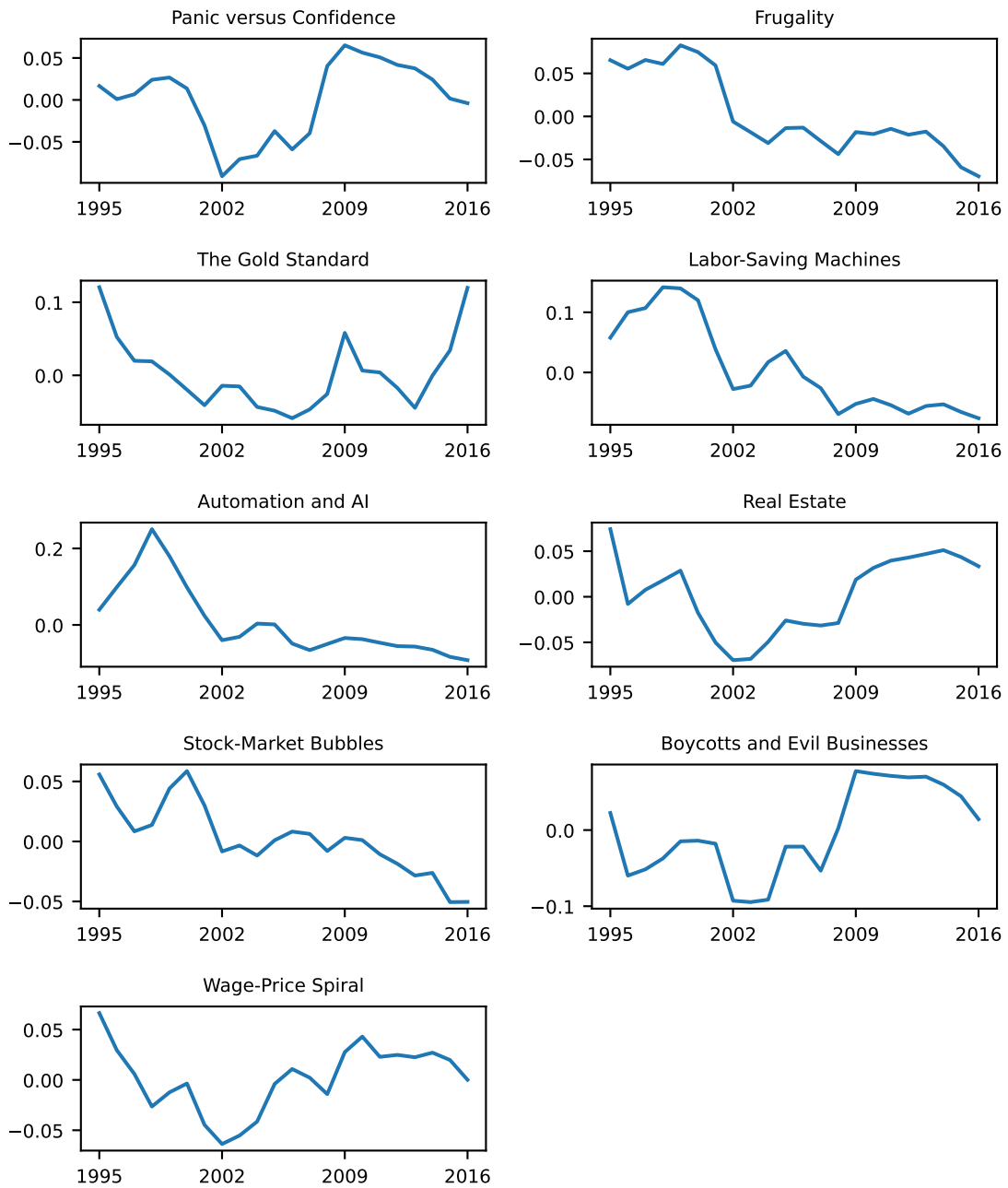
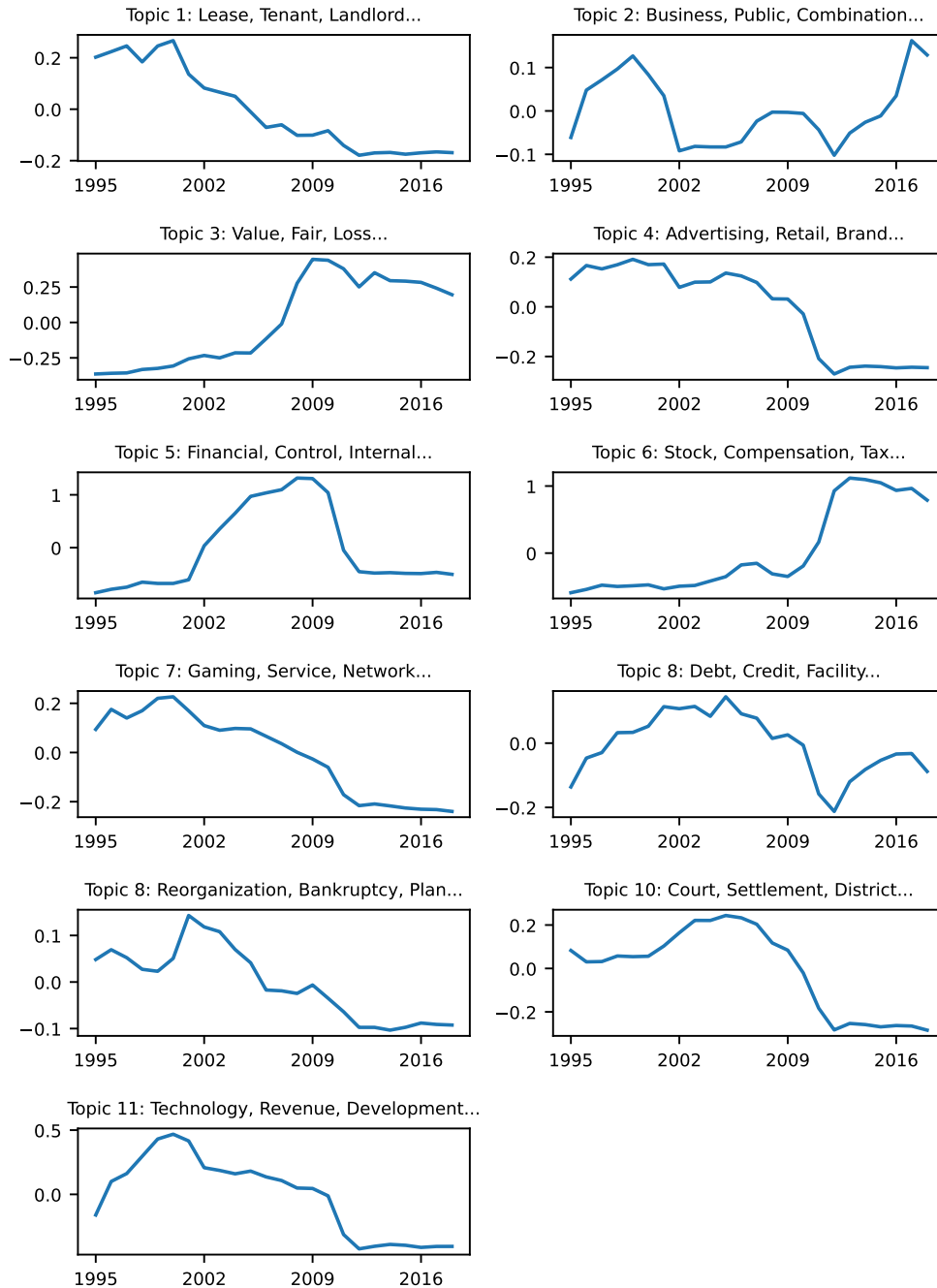


Figure A3: Time-Series for the Selected LDA Topics



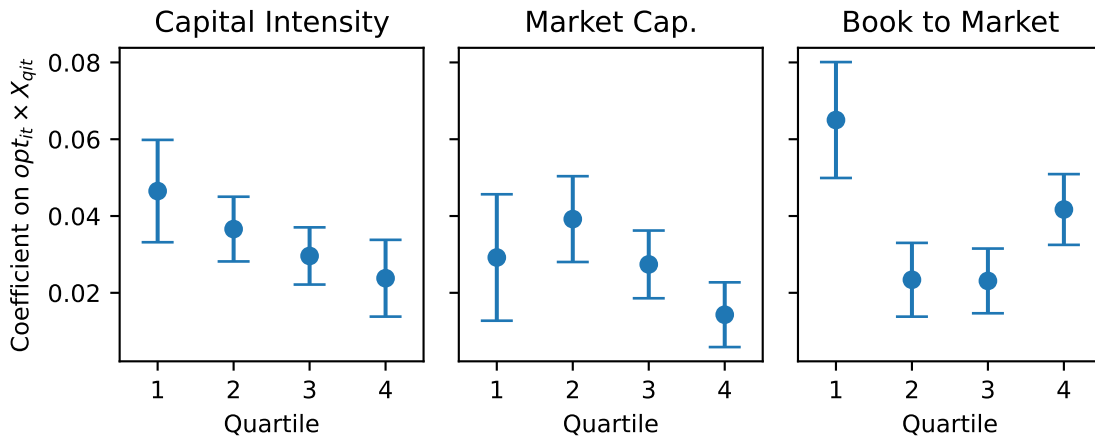
Notes: “Selection” is based on LASSO estimation of the model described in Section 4.1, and estimates of which are reported in Table A9.

Table A5: Variance Decomposition of Narratives

Narrative N_{it}	Fraction Variance Explained By Means				
	t	Ind.	Ind. x t	Firm	All
Net Sentiment	0.014	0.053	0.082	0.511	0.530
Optimism	0.011	0.041	0.067	0.427	0.444
Topic Narratives (25th)	0.010	0.003	0.049	0.252	0.306
Topic Narratives (median)	0.035	0.014	0.099	0.420	0.575
Topic Narratives (75th)	0.087	0.071	0.237	0.645	0.735
Shiller Narratives (25th)	0.002	0.050	0.062	0.758	0.761
Shiller Narratives (median)	0.002	0.071	0.087	0.763	0.770
Shiller Narratives (75th)	0.003	0.095	0.109	0.793	0.794

Notes: Each cell is $1 - \text{Var}[N_{it}^2]/\text{Var}[N_{it}]$, where N_{it} is the narrative intensity and N_{it}^2 is the same after projecting out means at the indicated level. The last column (“All”) partials out industry-by-time means and firm means.

Figure A4: Heterogeneous Effects of Optimism on Hiring



Notes: In each panel, we show estimates from the regression $\Delta \log L_{it} = \sum_{q=1}^4 \beta_q (\text{opt}_{it} - X_{qit}) + \gamma_i + \chi_{j(\cdot):t} + \epsilon_{it}$, where X_{qit} indicates quartile q of the studied variable: one minus the variable cost share of sales, market capitalization, or book-to-market ratio. Error bars are 95% confidence intervals.

Table A6: The Effect of Optimism on Hiring, Conference-Call Measurement

	(1)	(2)	(3)	(4)	(5)
	Outcome is				
	$\Delta \log L_{it}$				$\Delta \log L_{i;t+1}$
optCC _{it}	0.0277 (0.0038)	0.0173 (0.0040)	0.0116 (0.0038)	0.0237 (0.0038)	0.0123 (0.0044)
Industry-by-time FE	×	×	×	×	×
Firm FE	×	×	×		×
Lag labor		×	×	×	×
Current and lag TFP		×	×	×	×
Log Book to Market			×		
Stock Return			×		
Leverage			×		
<i>N</i>	19,625	11,565	10,826	11,919	11,416
<i>R</i> ²	0.300	0.461	0.467	0.172	0.429

Notes: The regression models are identical to those reported in Table 1, but using the measurement of optimism from sales and earnings conference calls. Standard errors are two-way clustered by firm ID and industry-year. In column 5, control variables are dated $t + 1$.

Table A7: The Effect of Optimism on All Inputs

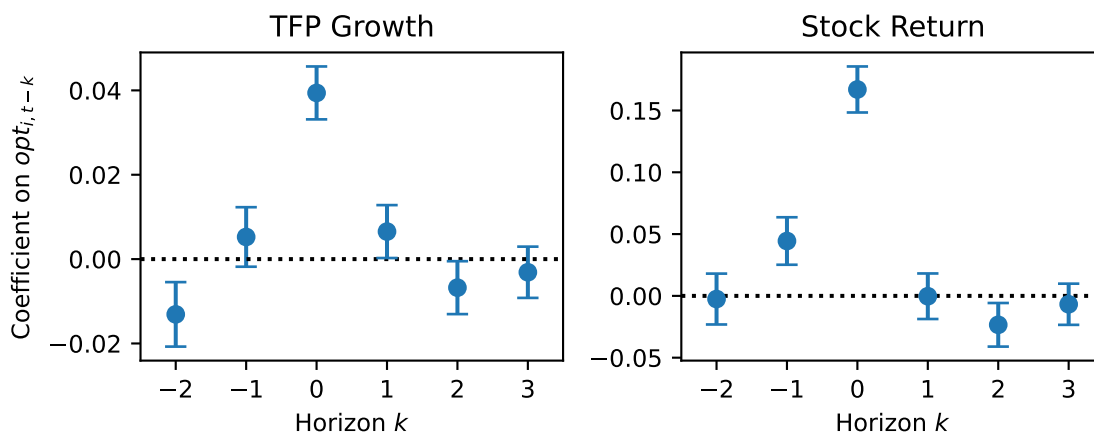
	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome is					
	$\Delta \log L_{it}$		$\Delta \log M_{it}$		$\Delta \log K_{it}$	
opt _{it}	0.0355 (0.0030)	0.0305 (0.0030)	0.0397 (0.0034)	0.0193 (0.0033)	0.0353 (0.0033)	0.0188 (0.0035)
Industry-by-time FE	×	×	×	×	×	×
Firm FE	×	×	×	×	×	×
Lag input	×	×	×	×	×	×
Current and lag TFP	×	×	×	×	×	×
<i>N</i>	71,161	39,298	66,574	39,366	70,215	39,258
<i>R</i> ²	0.259	0.401	0.298	0.418	0.282	0.409

Notes: Standard errors are two-way clustered by firm ID and industry-year. $\delta \log M_t$ is the log difference of all variable cost expenditures (“materials”), the sum of cost of goods sold (COGS) and sales, general, and administrative expenses (SGA). $\delta \log K_t$ is the value of the capital stock is the log difference level of net plant, property, and equipment (PPE) between balance-sheet years $t - 2$ and $t - 1$.

Table A8: The Effect of Optimism on Stock Prices: High-Frequency Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome is stock return on					
	Filing Day		Prior Four Days		Next Four Days	
opt_{it}	0.000145 (0.0007)	-0.000142 (0.0007)	0.00106 (0.0011)	0.000963 (0.0014)	0.00173 (0.0012)	0.00249 (0.0016)
Firm FE	×	×	×	×	×	×
Industry-by-FY FE	×	×	×	×	×	×
Industry-FF3 inter.		×		×		×
N	39,457	39,457	39,396	17,710	39,346	19,708
R^2	0.189	0.246	0.190	0.345	0.206	0.317

Notes: The regression equation for columns (1), (3), and (5) is $R_{i:w(t)} = \beta opt_{it} + \gamma_i + \chi_{j(i):y(i;t)} + \varepsilon_{it}$ where i indexes firms, t is the 10K filing day, $w(t)$ is a window around the day (the same day, the prior four days, or the next four days), and $y(i, t)$ is the fiscal year associated with the specific 10-K. In columns (2), (4), and (6), we add interactions of industry codes with the filing day's (i) the market minus risk free rate, (ii) high-minus-low return, and (iii) small-minus-big return. Standard errors are two-way clustered by firm ID and industry-year.

Figure A5: Dynamic Relationship between Optimism and Firm Fundamentals, Conference-Call Measurement

Notes: The regression models are identical to those reported in Figure 2, but using the measurement of optimism from sales and earnings conference calls. Each coefficient is estimated from a separate projection regression. Error bars are 95% confidence intervals.

Table A9: The Effect of Narratives on Hiring: All Narratives

	(1)	(2)	(3)	(4)
	OLS	OLS	OLS	IV
Shiller: Labor-Saving Machines	0.0106 (0.0028)			
Shiller: Stock Bubbles	0.00968 (0.0031)			
Topic 1: Lease, Tenant, Landlord...		0.0109 (0.0017)		
Topic 2: Business, Public, Combination...		0.0266 (0.0045)		
Topic 3: Value, Fair, Loss...		-0.00383 (0.0016)		
Topic 4: Advertising, Retail, Brand...		0.00864 (0.0024)		
Topic 5: Financial, Control, Internal...		-0.000655 (0.0025)		
Topic 6: Stock, Compensation, Tax...		0.0135 (0.0019)		
Topic 7: Gaming, Service, Network...		0.0146 (0.0040)		
Topic 8: Debt, Credit, Facility...		-0.00584 (0.0022)		
Topic 8: Reorganization, Bankruptcy, Plan...		-0.00842 (0.0018)		
Topic 10: Court, Settlement, District...		-0.00749 (0.0019)		
Topic 11: Technology, Revenue, Development...		0.0259 (0.0040)		
opt_{it}			0.0305 (0.0030)	0.0597 (0.0099)
Industry-by-time FE	×	×	×	×
Firm FE	×	×	×	×
Lag labor	×	×	×	×
Current and lag TFP	×	×	×	×
N	37,462	39,298	39,298	34,106
R^2	0.413	0.405	0.401	0.130
First-stage F	—	—	—	189.0

Notes: Standard errors are two-way clustered by firm ID and industry-year. The first-stage equation for column 4 is described in the main text.

Table A10: The Virality of Optimism, CEO Change Strategy

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
			Outcome is opt_{it}	
Industry lag, $\overline{\text{opt}}_{j(\cdot):t-1}$	0.275 (0.0407)	0.260 (0.2035)	0.195 (0.0760)	0.272 (0.5270)
Peer lag, $\overline{\text{opt}}_{p(\cdot):t-1}$			0.0437 (0.0236)	0.129 (0.1677)
Firm FE	×	×	×	×
Time FE	×	×	×	×
Industry output growth, $\Delta \log Y_{j(\cdot):t-1}$	×	×	×	×
N	50,604	50,604	7,873	7,873
R^2	0.503	0.051	0.508	0.020
First-stage F	—	29.7	—	36.8

Notes: Standard errors are two-way clustered by firm ID and industry-year. The IV strategies instrument the industry and/or peer lag with the CEO-change variation in those averages.

Table A11: The Virality and Associativeness of Optimism, Instrumented With Other Narratives

	(1)	(2)	(3)	(4)
		Outcome is opt_{it}		
	OLS	IV	OLS	IV
Own lag, $\text{opt}_{i:t-1}$	0.209 (0.0071)	0.207 (0.0072)	0.214 (0.0080)	0.200 (0.0084)
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290 (0.0578)	0.393 (0.0597)		
Real GDP growth, $\Delta \log Y_{t-1}$	0.804 (0.2204)	0.672 (0.2153)		
Ind. lag, $\overline{\text{opt}}_{j(i):t-1}$			0.276 (0.0396)	0.437 (0.0748)
Ind. output growth, $\Delta \log Y_{j(i):t-1}$			0.0560 (0.0309)	0.0390 (0.0342)
Firm FE?	×	×	×	×
Time FE?		×	×	
N	64,948	64,569	52,258	47,536
R^2	0.481	0.050	0.501	0.047
First-stage F	—	795.3	—	19.8

Notes: Standard errors are two-way clustered by firm ID and industry-year. In column 2, the endogenous variable is $\overline{\text{opt}}_{t-1}$ and the first-stage equation is

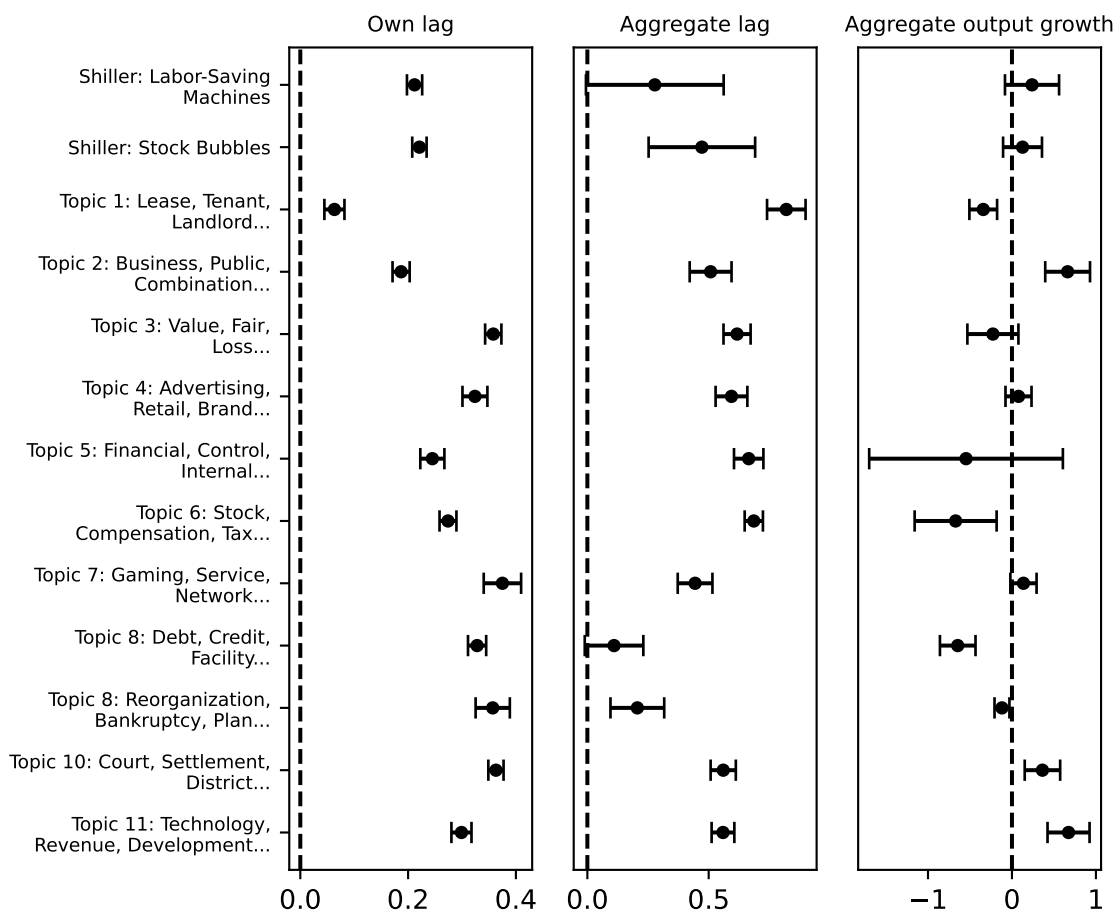
$$\overline{\text{opt}}_{t-1} = \sum_{k=1}^{K_S} \delta_{Sk} \overline{\text{Shiller}}_t^k + \sum_{k=1}^{K_T} \delta_{tk} \overline{\text{topic}}_t^k + \tilde{u} \text{opt}_{i:t-1} + \tilde{r} \Delta \log Y_{t-1} + \tilde{\gamma}_i + u_{it}$$

where the sums are over the LASSO-selected narratives (see Table 4). In column 4, the endogenous variable is $\overline{\text{opt}}_{j(i):t-1}$ and the first-stage equation is

$$\overline{\text{opt}}_{j(i):t-1} = \sum_{k=1}^{K_S} \delta_{Sk} \overline{\text{Shiller}}_{j(i):t}^k + \sum_{k=1}^{K_T} \delta_{tk} \overline{\text{topic}}_{j(i):t}^k + \tilde{u} \text{opt}_{i:t-1} + \tilde{\chi}_t + \tilde{\gamma}_i + u_{it}$$

where the industry means are leave-one-out.

Figure A6: The Virality and Associativeness of Other Identified Narratives



Notes: For each narrative, we estimate the analogue of Equation 20. The three panels respectively show the coefficients on $\hat{\lambda}_{i:t-1}$, $\bar{\lambda}_{j(t):t-1}$ and $\Delta \log Y_{t-1}$, with 95% confidence intervals based on double-clustered (firm and industry-year) standard errors.

Table A12: Sensitivity Analysis for Quantitative Analysis

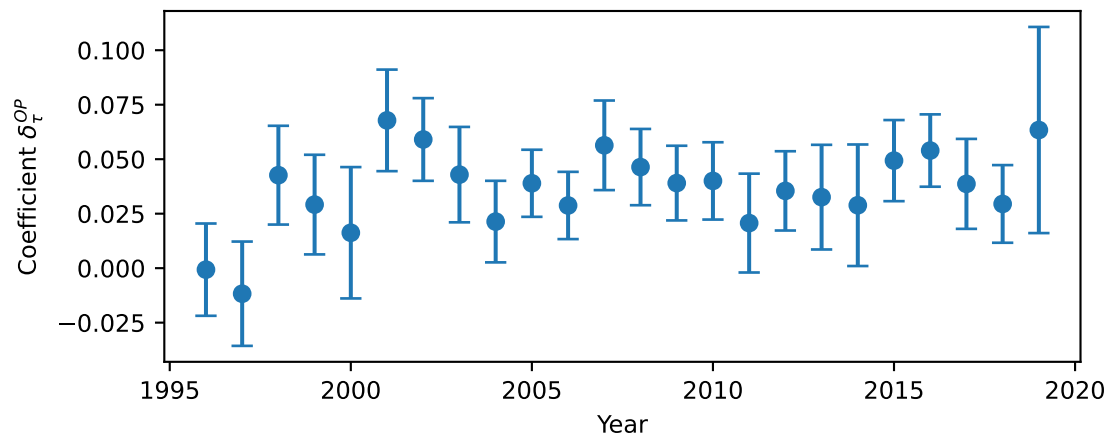
Scenario	Calibrated				Implied		Results	
	α	γ	ψ	ϵ	ω	$m = (1 - \omega)^{-1}$	$\hat{c}_Q(0)$	$\hat{c}_Q(1)$
Baseline	1.0	0.0	0.4	2.6	0.490	1.962	0.048	0.170
High ϵ	1.0	0.0	0.4	5.0	0.333	1.500	0.032	0.150
Low ψ	1.0	0.0	0.2	2.6	0.658	2.923	0.084	0.211
High γ	1.0	1.0	0.4	2.6	-0.784	0.560	0.006	0.091
Decreasing RtS	0.75	0.0	0.4	2.6	0.307	1.444	0.028	0.145
Empirical Multiplier	1.0	0.0	0.6	5.0	0.250	1.333	0.025	0.142

Notes: Sensitivity of percentage output variance explained statically and at an eight-year horizon by optimism under alternative calibrations of the macroeconomic parameters. Baseline corresponds to our main calibration. High ϵ increases the elasticity of substitution from 2.6 to 5.0. Low ψ decreases the Frisch elasticity from 0.4 to 0.2. High γ increases the curvature of consumption utility (indexing income effects in labor supply) from 0.0 to 1.0. Decreasing RtS reduces the returns to scale parameter α from 1.0 to 0.75. Empirical Multiplier adjusts ψ and ϵ so that the output multiplier is 1.33, in line with estimates from [Flynn, Patterson, and Sturm \(2022\)](#).

Table A13: Data Definitions in Compustat

	Quantity	Expenditure
Production, x_{it}	—	sale
Employment, L_{it}	emp	emp industry wage
Materials, M_{it}	—	cogs + xsga dp
Capital, K_{it}	ppegt plus net investment	—

Figure A7: Time-Varying Effects of Optimism on Hiring



Notes: Each dot is a coefficient δ^{OP} estimated from Equation 231. Error bars are 95% confidence intervals, based on standard errors clustered by firm and industry-time.