The Macroeconomics of Narratives

Joel P. Flynn* Karthik A. Sastry†
MIT Harvard

October 16, 2022

Abstract

We study the macroeconomic implications of viral, belief-altering narratives. Empirically, we use natural-language-processing methods to measure narratives in the text of all US public firms' end-of-year reports (Forms 10-K). We find that: (i) firms' hiring decisions respond strongly to narratives, (ii) narratives spread virally among firms, and (iii) this spread is responsive to macroeconomic conditions. To understand the macroeconomic implications of these forces, we embed viral narratives in a Neoclassical business-cycle model. We characterize, in terms of the decision-relevance and virality of narratives, when the unique equilibrium features: (i) non-fundamentally driven business-cycle fluctuations, (ii) multiple, self-fulfilling steady states (hysteresis), and (iii) the coexistence of hump-shaped responses to small shocks with regime-shifting behavior in response to large shocks. Conditional on calibrating standard preference and technological parameters, our empirical estimates identify both the static, general equilibrium effect of narratives on output and their dynamics. Statically, we find that aggregate fluctuations in narratives account for approximately 32% of the output reduction during the Early 2000s Recession and 18% during the Great Recession. Dynamically, we reject the possibility of narratively driven hysteresis for most, but not all, narratives.

*MIT Department of Economics, 50 Memorial Drive, Cambridge, MA 02142. Email: jpflynn@mit.edu.
†Harvard Department of Economics, 1805 Cambridge Street, Cambridge, MA 02138. Email: ksastry@fas.harvard.edu.

We are grateful to Daron Acemoglu, George-Marios Angeletos, Ricardo Caballero, Andrew Caplin, Oğuzhan Çelebi, Hui Chen, John Coglianese, Roberto Corrao, Xavier Gabaix, Luca Gagliardone, Ying Gao, Adam Guren, Basil Halperin, Tarek Hassan, Leonid Kogan, Antoine Levy, Chen Lian, Andrea Manera, Jackson Mejia, Stephen Morris, Georgios Nikolakoudis, Jonathan Parker, Christina Patterson, Charlie Rafkin, Eitan Sapiro-Gheiler, Antoinette Schoar, Robert Shiller (Discussant), Andrei Shleifer, David Thesmar, Adrien Verdelhan, Emil Verner, Jiang Wang, Christian Wolf, and participants in the NBER Summer Institute Behavioral Macroeconomics Meeting, MIT Macro Lunch, MIT Theory Lunch, and MIT Sloan Finance Lunch for helpful comments. First posted version: June 19, 2022.

1 Introduction

In *The General Theory of Employment, Interest, and Money*, Keynes (1936) argues that *animal spirits* that induce "spontaneous optimism" 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), which hypothesizes that viral stories and worldviews are a primary driver of economic fluctuations. However, the importance of narratives for business cycles has not yet been established.

In this paper, we study joint fluctuations in narratives and the macroeconomy. To do this, we develop a framework in which narratives' decision relevance and virality can be formalized, theoretically studied, and empirically evaluated. In our framework, narratives describe subjective models of the macroeconomy and form building blocks of beliefs. Based upon the narratives in which they believe, agents take actions under uncertainty about fundamentals and the behavior of others. The spread of narratives depends both on their prevalence in the population (virality) and how well they describe economic reality (associativeness).

We operationalize this framework to measure narratives and estimate their importance in the data. We apply several natural-language-processing methods to measure narratives in the universe of 10-K regulatory filings, in which all US public firms discuss "perspectives on [their] business results and what is driving them" (SEC, 2011). We find that measured narratives are relevant for decisions, as many narratives predict firms' hiring decisions, and viral, as past narratives of other firms influence firms' narratives in the present.

To understand these findings' macroeconomic implications, we embed viral narratives in a business-cycle model. We show that viral narratives can lead to non-fundamentally driven boom-bust cycles and hysteresis. Conditional on calibrating standard preference and technological parameters, our empirical estimates identify both the static, general equilibrium effect of narratives on output and their dynamics. We find that aggregate fluctuations in narratives generate significant movements in aggregate output, while we empirically reject the possibility of narratively driven hysteresis for most of our estimated narratives.

Measuring Narratives. Our first goal is to empirically evaluate the two premises of our framework for narrative macroeconomics: narratives' *decision-relevance* and *virality*. To this end, we construct a dataset combining information on US public firms' adoption of textual narratives, using data from firms' regulatory filings (Forms 10-K) and earnings call transcripts, and their decisions, using data from Compustat.

We employ three different techniques to measure textual narratives at different levels of granularity, from coarse measurement of narrative sentiment to fine measurements of topic-level discussion. The first technique measures the intensity of positive and negative sentiment using the 10-K-specific dictionary introduced by Loughran and McDonald (2011). We interpret this measure as capturing optimism. The second technique measures 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, Hollander, Van Lent, and Tahoun (2019) and Flynn and Sastry (2022), 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 estimates 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. Since we recover these "topic narratives" via an unsupervised method, they allow the data to speak flexibly about what textual narratives feature most prominently in firm communications.

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. However, it is difficult to ascertain the relationship between narrative and macroeconomic dynamics from the time series alone. This is because narratives serve a dual role of describing true economic fundamentals and encoding non-fundamental beliefs. This challenge for time-series analysis motivates our main empirical strategies, which leverage the considerable cross-sectional variation in narratives to measure the relationship between textual narratives and economic outcomes.

Using these data, we first study the *relevance* of our measured narratives for economic decisions. We focus initially on optimism. We find that optimistic firms, defined as firms with above-median sentiment, hire 3.6 percentage points more than pessimistic firms in a given year, net of firm and sector-time fixed effects. This finding is robust to accounting for firm-level productivity and financial conditions. Moreover, we find that optimism is uncorrelated with future productivity growth and *negatively* correlated with future stock returns and profitability. These findings are inconsistent with the possibility that optimism predicts hiring only because it captures positive firm-level fundamentals (or news thereof). Further, we show using managerial guidance data from IBES that optimism predicts negative errors in sales forecasts (*i.e.,* forecasts greater than the realization). This result verifies that optimistic language manifests in over-optimistic beliefs.

We therefore interpret the association of textual optimism with hiring as arising from non-fundamental, narratively driven, and optimistic beliefs. 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. Because these narratives are high-dimensional and may not be relevant for firm decisions, we use the Rigorous LASSO method of Belloni, Chernozhukov, Hansen, and Kozbur (2016) for estimating their effects on hiring. 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. First, focusing on optimism, we find 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 for aggregate conditions with time fixed effects. Moreover, both the aggregate and industry-level patterns are robust to controlling for future idiosyncratic and aggregate economic conditions. This is inconsistent with the explanation that aggregate optimism drives future optimism through its correlation with omitted positive news about measured 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. This analysis relies on aggregate variation to estimate cross-agent spillovers. By using past aggregate optimism in a panel setting, our estimates are not threatened by the reflection problem of Manski (1993). Nevertheless, common shocks that are not spanned by measured aggregate and industry-level conditions may generate omitted variables bias. To address this concern, we employ strategies based on using idiosyncratic shocks to large firms (the granular IV approach of Gabaix and Koijen, 2020) and the aforementioned plausibly exogenous changes in CEOs as instruments for aggregate and industry-level optimism. We find qualitatively consistent effects. Finally, we perform similar analyses for the other decision-relevant narratively-identified and topic narratives. We find that almost all of them are viral and that many are associative.

These results, taken together, provide strong evidence of the decision-relevance, virality, and associativeness of a large set of textual narratives.

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 real variant of the standard Neoclassical 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 remain and 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 a unique equilibrium in which aggregate output is log-linear in aggregate productivity and a non-linear function of the fraction of optimists in the population, a sufficient statistic for narratives. 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. In this way, movements in aggregate optimism lead to non-fundamental fluctuations in aggregate output.

We next describe the dynamics of narratives and output. For a fixed level of aggregate productivity, while there always exists a steady-state level of optimism and equilibrium is unique, there may nevertheless be multiple steady-state levels of optimism. We provide a necessary and sufficient condition for a particularly extreme type of this steady-state multiplicity: if the decision-relevance, virality, and associativeness of narratives are sufficiently strong, then unanimous optimism and unanimous pessimism are both stable steady states. Moreover, depending on the initial fraction of optimists, the economy is (almost everywhere) globally attracted to one of these extreme steady states. In this way, narratives can generate hysteresis: fixing aggregate productivity, depending on how many optimists there are initially, optimism can either catch on forever ("go viral") or die out entirely.

We next study how the economy evolves in response to shocks. We first study unanticipated, deterministic "MIT shocks." Responses can fall into three qualitative regimes, depending on the economy's initial condition and the size of the shock. If a shock is small, it can have a fully transitory impact on aggregate output, because it fails to even seed a new narrative. If a shock is medium-sized, it can have persistent and hump-shaped effects on aggregate output, because it seeds a new narrative that briefly persists before dying out. If a shock is large, it can have a permanent effect on aggregate output, because it makes a narrative go viral. We next study stochastic behavior. We show that the economy can regularly oscillate between extreme optimism and pessimism and provide analytical upper bounds on the expected period of these oscillations. Both the possibility of these effects and their quantitative magnitudes depend positively on key measurable parameters: the decision relevance, associativeness, and virality of narratives.

Quantification. In the final part of the paper, we calibrate our model to quantify the extent to which fluctuations in narrative optimism explain historical business cycle fluctuations and understand the extent to which narrative dynamics generate hysteresis. We leverage the fact that our empirical estimates identify both the partial equilibrium effects of narratives on hiring and the nature of narrative diffusion. Thus, conditional on using a standard external calibration of preference and production parameters and using the time series of US GDP to discipline the driving process for fundamental shocks, our earlier estimates point identify the parameters of our model.

Statically, we find that aggregate output would be 7% higher in an economy in which all firms are optimistic than one in which all firms are pessimistic. Thus, our measured aggregate peak-to-trough movement in optimism accounts for 32% of output loss during the Early 2000s Recession (following the burst of the dot-com bubble) and 18% during the Great Recession. Dynamically, we find that optimism accounts for 17% of the short-run (one-year) and 66% of the medium-run (two-year) autocovariance in output. For optimism, we quantitatively reject the theoretical condition required for hysteresis in both optimism and output dynamics. But we find this condition is satisfied for other narratives, implying that it is natural for these narratives either to die out or "go viral" depending on initial conditions.[1] Taken together, our results suggest that time-varying, endogenous narratives are a significant driver of macroeconomic fluctuations.

Related Literature. We relate to an empirical literature that measures narratives following Carroll (2001) and Shiller (2017).[2] Of most relevance, Andre, Haaland, Roth, and Wohlfart (2022) use surveys to understand narratives underlying inflation, Goetzmann, Kim, and Shiller (2022) analyze narratives about financial crashes in news media, Macaulay and Song (2022) study how news coverage of specific narratives affects sentiment on social media, and Bybee, Kelly, Manela, and Xiu (2021) 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 literature studying the relationship between firm-level outcomes and their language (Loughran and McDonald, 2011; Hassan, Hollander, Van Lent, and Tahoun, 2019; Hassan, Schreger, Schwedeler, and Tahoun, 2021; Handley and Li, 2020). In contrast to these papers, we calibrate a model to match our firm-level findings and study their general-equilibrium consequences.[3]

Our work 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

[1] Normatively, we show that viral optimism can be welfare-improving even if it is unfounded. Quantitatively, we find that optimism is welfare-improving and welfare-equivalent to a 1.3% production subsidy.

[2] See Carroll and Wang (2022) for a review of the literature on "epidemiological" models of expectations formation that are centered around social interactions.

[3] Flynn and Sastry (2022) and Song and Stern (2021) share this approach of contextualizing the effect of language-based variables on firm-level outcomes in a model.

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.[4]

Second, our work relates 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 micro-founds such shocks via the endogenous evolution of narratives and corresponding degree of optimism.[5] Our work contrasts with that of Maxted (2020), where agents extrapolate from recent changes in fundamentals, in that both equilibrium outcomes and social dynamics determine agents' models.[6] 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.[7]

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; Brock and Hommes, 1997; 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.

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

[4] Thus, 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).

[5] In 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.

[6] Bordalo, Gennaioli, Shleifer, and Terry (2021) and Bordalo, Gennaioli, Kwon, and Shleifer (2021) respectively study how a similar mechanism can generate credit cycles and speculative bubbles.

[7] This 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).

2 Narratives: A Conceptual Framework

We first describe a conceptual framework that formalizes the two premises of the macroeconomics of narratives: that narratives are decision-relevant and that narratives spread virally and associatively. We embed these two premises in an abstract game in which narratives form the building blocks of agents' beliefs and agents care about their own actions, fundamentals, and aggregates of other agents' actions. This game nests our later macroeconomic model in Section 5. We then derive two regression equations that allow us to test the decision-relevance, virality, and associativeness of narratives. We bring these regressions to the data in Section 4 and show that they obtain exactly in our macroeconomic model in Section 6.

2.1 Premise I: Narratives Are Relevant for Decisions

The first premise is that narratives are decision-relevant.[8] 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 are random 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 $k \in \mathcal{K}$, as a probability distribution $N_k \in \Delta(\Theta)$ within the set of narratives $\mathcal{N} = \{N_k\}_{k\in\mathcal{K}}$. For example, if the fundamental θ describes the strength of productivity, then a pessimistic narrative N_P might correspond to the view that "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 $\lambda = \{\lambda_k\}_{k\in\mathcal{K}} \in \Lambda \subseteq \Delta(\mathcal{K})$ 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 $\mathcal{N}$:

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

[8] The management and organizational literature also views narratives as forming a common set of stories that underpin beliefs (see, *e.g.*, Isabella, 1990; Maitlis, 2005; Loewenstein, Ocasio, and Jones, 2012; Vaara, Sonenshein, and Boje, 2016). Relatedly, Acemoglu and Robinson (2021) postulate that culture arises from the combination of latent cultural attributes. By microfounding the process by which cultural attributes combine, our analysis could be applied to this context.

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 $\pi(\theta) = \frac{1}{2}N_O(\theta) + \frac{1}{2}N_P(\theta)$.

We now map the narratives in which agents believe to their decisions. Time is discrete and infinite, indexed by $t \in \mathbb{N}$. Agents care about their own actions $x_{it} \in \mathcal{X}$, aggregate outcomes $Y_t \in \mathcal{Y}$, the fundamental state $\theta_t \in \Theta$, and an idiosyncratic preference shifter $\omega_i \in \Omega$. They have utility functions $u : \mathcal{X} \times \mathcal{Y} \times \Theta \times \Omega \to \mathbb{R}$ and information sets $\mathcal{I}_{it}$. Given their information sets, agents update each narrative belief by applying Bayes' rule and then form their posterior by placing their narrative weights on the updated narrative beliefs. Given a conjecture about the mapping from fundamental states to aggregates $\hat{Y}_t : \Theta \to \mathcal{Y}$, the agents maximize their expected utility given narrative weights λ_{it} and information $\mathcal{I}_{it}$:

$$\max_{x_{it} \in \mathcal{X}} \mathbb{E}_{\pi_{\lambda_{it}}} \left[u(x_{it}, \hat{Y}_t(\theta_t), \theta_t, \omega_i) \mid \mathcal{I}_{it} \right] \tag{2}$$

We linearize these best replies to obtain the following regression equation that allows us to test for the decision-relevance of narratives (see Proposition 6 in Appendix A.1 for the formal arguments):[9]

$$x_{it} = \gamma_i + \chi_t + \sum_{k \in \mathcal{K}} \delta_k \lambda_{k,it} + \varepsilon_{it} \tag{3}$$

In this equation, γ_i captures time-invariant factors (preference shifters), χ_t captures time-varying aggregate variables such as fundamentals or the overall prevalence of narratives, the 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. The hypothesis that narratives are decision-relevant is that $\delta_k \neq 0$ for some $k \in \mathcal{K}$. To test this hypothesis in panel data on firms, we will use firm hiring as the relevant outcome and textual measures of narrative adoption as proxies for $\lambda_{k,it}$, the belief weights in the model. An analog of this equation will hold in equilibrium without approximation in our theoretical model in Section 5, facilitating the use of our estimates for model calibration.

[9]There 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 OLS.

2.2 Premise II: Narrative Spread Is Viral and Associative

We now formalize the second premise: that narratives spread virally and associatively. The extent of narrative penetration 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.

The evolution of the distribution of narratives over time is described by an updating rule $P : \Lambda \times \mathcal{Y} \times \Delta(\Lambda) \to \Delta(\Lambda)$, which returns the probabilities $\{P_{\lambda'}(\lambda, Y, Q)\}_{\lambda' \in \Lambda}$ that an agent with narrative weights λ changes their weights to λ' when the endogenous state is Y and the distribution of narratives in the population is Q.[10] Hence, conditional on a distribution of narratives at time t given by Q_t and realized endogenous outcomes given by Y_t, the next period's distribution of narratives is:

$$Q_{t+1,\lambda'} = \sum_{\lambda \in \Lambda} Q_{t,\lambda} P_{\lambda'}(\lambda, Y_t, Q_t) \tag{4}$$

At this level of generality, the updating function can capture 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.

We linearize the narrative updating equations to obtain the following system of linear probability models (see Proposition 7 in Appendix A.1 for the formal arguments):

$$\mathbb{P}[\lambda_{it} = \lambda \mid \lambda_{i,t-1}, Y_{t-1}, Q_{t-1}] = \zeta_\lambda + \sum_{\lambda' \in \Lambda} u_{\lambda',\lambda} \mathbb{I}[\lambda_{i,t-1} = \lambda'] + r_\lambda' Y_{t-1} + s_\lambda' Q_{t-1} \tag{5}$$

In this equation, u captures the agents' stubbornness in updating (*i.e.,* their proclivity to not update), r captures the associativeness in updating, and s captures virality in updating. Therefore, the hypotheses that narratives are viral and associative correspond, respectively, to $s_\lambda \neq 0$ and $r_\lambda \neq 0$ for some $\lambda \in \Lambda$. Again, we will use panel data on textual narrative adoption to test these hypotheses. Equation 5 is generated without approximation in the special case of our theoretical model in Section 5 that we study quantitatively.

[10] In Appendix B.6, we extend this setting to allow for idiosyncratic fundamentals and updating that depends on their realizations.

3 Data, Measurement, and Descriptive Statistics

We now describe how we develop a panel dataset on firms' narrative loadings and decisions. We measure textual proxies for narratives by applying several natural-language-processing techniques to two corpora of language: the universe of public firms' SEC Forms 10-K and a large sample of earnings calls. 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

Text. Our main source of firm-level textual data is SEC Form 10-K. Each publicly traded firm in the US submits an annual 10-K 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 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.[11] 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 Appendix C.2.

[11] Other machine-learning approaches use, as input, the entire document instead of its bag-of-words representation. Examples include Doc2Vec, as recently employed by Goetzmann, Kim, and Shiller (2022) to study crash narratives, and RELATIO, which was recently developed by Ash, Gauthier, and Widmer (2021). We view integrating these methods into our analysis as an interesting avenue for future study.

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, capital, and other financial variables. 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 ignore firms in the financial and utilities sectors due to their markedly different production and/or market structure. More details about our sample selection are in 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.[12]

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 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.[13]

3.2 Measurement: Recovering Narratives from Language

We employ three techniques to measure textual narratives at different levels of granularity.

[12] Firm 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.

[13] This 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.

Sentiment Narratives. We first measure firm narrative sentiment. We categorize individual words as either positive or negative using the dictionaries constructed by Loughran and McDonald (2011). These dictionaries adjust standard tools for sentiment analysis to more precisely score financial communications, in which certain words (*e.g.*, the leading example "liability") have specific definitions.[14] 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:

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

where $\text{tf}(w)_{it}$ is the term frequency of all bigrams including 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:

$$\text{opt}_{it} = \mathbb{I}\left[\text{sentiment}_{it} \geq \text{med}\left(\text{sentiment}_{it}\right)\right] \tag{7}$$

This variable has a simple interpretation in capturing optimistic narratives, but necessarily collapses more fine-grained discussion of specific topics.

Narrative Identification of Narratives. To measure more specific narratives entertained by firms, we next consider a supervised strategy based on 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 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*. We measure narrative adoption by computing the 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 (2022). For each narrative k, we first compute the term-frequency-inverse-document-frequency (tf-idf) score to obtain

[14] Loughran 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 document scoring methodology in Appendix C.3.

a set of words most indicative of that narrative:

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

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:

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

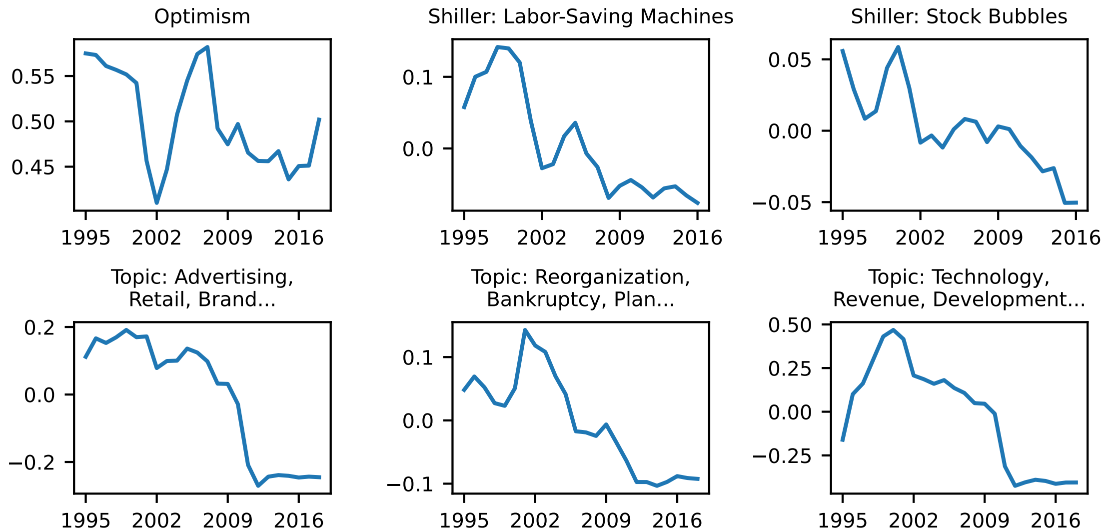
We compute loadings on each narrative, Shiller_{it}^{k}, by taking the z-score. These variables measure a set of more specific topics, but rely on Shiller's specific wording of narratives.

Unsupervised Recovery of Narratives. Finally, to identify narratives without relying on any external references, we use Latent Dirichlet Allocation (LDA), a hierarchical Bayesian model in which documents are constructed by combining a latent set of topic narratives (Blei, Ng, and Jordan, 2003). More specifically, 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 $\vartheta = (\vartheta_1, \ldots, \vartheta_M)$, over which we impose a Dirichlet prior with parameter $\alpha = \{\alpha_k\}_{k \in \mathcal{K}}$, where α_k represents the prior weight that topic k is in any document. Third, the distribution of words across topics is given by $\phi = (\phi_1, \ldots, \phi_K)$, over which we impose a Dirichlet prior with parameter $\beta = \{\beta_{jk}\}_{k \in \mathcal{K}}$, 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 distribution 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, when standard Markov Chain Monte Carlo methods would be slow.[15] Given

[15]For computational reasons, we estimate the model using all available documents from a randomly sampled 10,000 of our 37,684 unique possible firms. We score all documents with this estimated model.

Figure 1: Aggregate Time Series for Six Selected Narratives



Notes: Optimism is measured as the fraction of optimistic firms. The other five time series are cross-sectional averages of z-score transformed variables (zero mean, unit standard deviation).

the estimated LDA, we construct the document-level narrative score as the posterior probability of that topic in the estimated document-specific topic distribution $\hat{p}$:

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

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.

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, and some have notable trends and breaks.[16] 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 cyclicality 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, our framework implies that it is challenging to interpret these basic time-series facts for two reasons.[17] First, 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.

4 Empirical Results

Moving beyond descriptive evidence, we now use our dataset of firm-level outcomes and narrative loadings to test the two premises that narratives are decision-relevant and that narrative spread is viral and associative.

[16] In 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 identifies as relevant for hiring (Figure A3).

[17] In Table A27 we show that failing to control for aggregate time-series variation leads to an upward bias in the impact of narratives on firm decisions.

4.1 Testing Premise I: Narratives Are Decision-Relevant

Empirical Strategy. From the conceptual framework in Section 2 (see Equation 3 and Proposition 6 in Appendix A.1), we have that firm hiring $\Delta \log L_{it}$ can be described to first order by the following regression equation:

$$\Delta \log L_{it} = \sum_{k \in \mathcal{K}} \delta_k \lambda_{k,it} + \gamma_i + \chi_t + \varepsilon_{it} \tag{11}$$

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:

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \tag{12}$$

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 adjustment costs in labor.[18] 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 log stock return (inclusive of dividends), and leverage (total debt over total assets). These controls proxy for both Tobin's q and firm-level financial frictions. These controls are conservative in that they may absorb variation in both omitted firm fundamentals and optimism itself. The point estimate

[18] In Appendix B.9, we show that the controls capture the impact of adjustment costs, to first order, for a forward-looking firm that observes current productivity. This notwithstanding, to evaluate robustness to the presence of richer adjustment dynamics, in Table A6, we control for up to three lags of productivity and labor and our financial controls and continue to find a significant impact of optimism on hiring.

Table 1: Narrative Optimism Predicts Hiring

	(1)	(2)	(3)	(4)	(5)
	Outcome is				
	$\Delta \log L_{it}$				$\Delta \log L_{i,t+1}$
opt_{it}	0.0355	0.0305	0.0250	0.0322	0.0216
	(0.0030)	(0.0030)	(0.0032)	(0.0028)	(0.0037)
Firm FE	✓	✓	✓		✓
Industry-by-time FE	✓	✓	✓	✓	✓
Lag labor		✓	✓	✓	✓
Current and lag TFP		✓	✓	✓	✓
Log Book to Market			✓		
Stock Return			✓		
Leverage			✓		
N	71,161	39,298	33,589	40,580	38,402
R^2	0.259	0.401	0.419	0.142	0.398

Notes: For columns 1-4, the regression model is Equation 12 and the outcome is the log change in firms' employment from year $t-1$ to t. The main regressor is a binary indicator for the optimistic narrative, defined in Section 3.2. In all specifications, we trim the 1% and 99% tails of the outcome variable. In column 5, the regression model is Equation 13, the outcome is the log change in firms' employment from year t to $t+1$, and control variables are dated $t+1$. Standard errors are two-way clustered by firm ID and industry-year.

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 from strict exogeneity violations (Nickell, 1981) and find similar results.[19]

To test if optimism predicts (and does not merely describe) hiring, we finally estimate a specification in which the outcome and control variables are time-shifted one year in advance:

$$\Delta \log L_{i,t+1} = \delta_{-1}^{OP} \text{opt}_{it} + \tau' X_{i,t+1} + \gamma_i + \chi_{j(i),t+1} + \varepsilon_{i,t+1} \tag{13}$$

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. In this specification, 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).[20]

[19] In Table A7, we report standard errors for the estimates of Table 1 under alternative schemes for clustering standard errors.

[20] In Table A8, we report results from our baseline regression Equation 12, using $\text{opt}_{i,t-1}$ as an instrument for opt_{it}. This is robust to any identification concern arising from the simultaneous determination of opt_{it} and $\Delta \log L_{it}$, but estimates the original parameter δ^{OP} (rather than δ_{-1}^{OP}). Our estimates are positive, statistically significant, and larger than our baseline estimates.

Robustness. As an alternative strategy to isolate plausibly exogenous variation in the narratives considered by firms, we study the effects on hiring of changes in narratives induced by plausibly exogenous CEO turnover. We provide the details in Appendix E.1. Specifically, we estimate a variant of Equation 12 over firm-year observations corresponding to the death, illness, or voluntary retirement of a CEO, as measured by Gentry, Harrison, Quigley, and Boivie (2021). We find quantitatively similar effects of narrative optimism on hiring as those reported in Table 1.

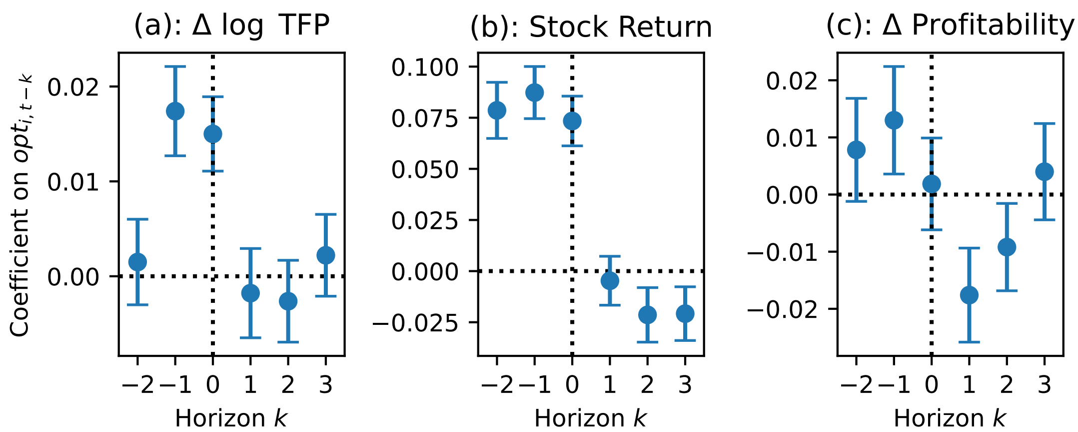
In the Appendix, we also report several additional results that probe the robustness of our main specification. We summarize them briefly here. First, Figure A4 shows estimates of a variant of our baseline regression interacting optimism with quartiles of firm characteristics. We find that the effect 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). Second, Table A9 repeats the analysis of Table 1 with our conference-call-based optimism measure, and finds similar results. Third, Table A10 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. Finally, we have so far studied the impact of binary optimism on hiring. To check if this construction drives our results, in Figure A5 we re-create the regression models of the first three columns of Table 1 with indicators for each decile of the continuous sentiment measure. We find monotonically increasing associations of hiring with higher bins of sentiment, implying that our binary construction is not masking non-monotone effects of the continuous measure.[21]

Inspecting the Mechanism: Narrative Optimism Does Not Predict Future Productivity Growth, Predicts Negative Stock Returns and Profitability. The coefficient of interest, δ^{OP}, measures the impact of optimism on hiring if optimism is uncorrelated with any omitted fundamental factors 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},

[21] In Appendix E.3, we also check whether the effects of narrative optimism depend on the past level of narrative optimism. We find that there are larger marginal effects on average for recently pessimistic firms, but that this heterogeneity is quantitatively small.

Figure 2: Dynamic Relationship of Optimism with Firm Fundamentals



Notes: The regression model is Equation 14, and each coefficient estimate is from a different regression. The outcomes are (a) the log change in TFP, calculated as described in Appendix D.2, (b) the log stock return inclusive of dividends, and (c) changes in profitability, defined as earnings before interest and taxes (EBIT) as a fraction of the previous fiscal year's variable costs. In all specifications, we trim the 1% and 99% tails of the outcome variables. Error bars are 95% confidence intervals, based on standard errors clustered at the firm and industry-year level.

either TFP growth $\Delta \log \hat{\theta}_{it}$, log stock returns R_{it}, or changes in profitability $\Delta\pi_{it}$, on optimism at leads and lags k:[22]

$$Z_{it} = \beta_k \text{opt}_{i,t-k} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{14}$$

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. For $k < 0$, and all three outcome variables, we find evidence of $\beta_k > 0$ – that is, a firm doing well today in terms of TFP growth, stock-market returns, and/or profitability is likely to become optimistic in the future. However, for $k > 0$, and all three 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 it is inconsistent with a story in which optimism is driven by

[22] We measure profitability as earnings before interest and taxes (EBIT) divided by the previous fiscal year's total variable costs (cost of goods sold (COGS) plus selling, general, and administrative expense (SGA), minus depreciation).

Table 2: Narrative Optimism Predicts Over-Optimistic Forecasts

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

Notes: The regression model is Equation 15. The outcomes are binary indicators for whether sales guidance was high relative to realized sales (columns 1 and 2) or high relative to contemporaneous analyst forecasts (columns 3 and 4), as defined in Section 3.1. Standard errors are two-way clustered by firm ID and industry-year.

news about fundamentals.[23] We find, in sharp contrast, that optimistic firms have *negative* stock returns and decreasing profitability in the future. This is consistent with our finding that optimistic firms persistently increase input expenditure (column 5 of Table 1), but see no increase in productivity (panel (a) of Figure 2). Figures A6 and A7 replicate this analysis with conference-call-based optimism and the continuous measure of net sentiment, respectively, and find similar results.

This analysis focuses on real, rather than financial, fundamentals. In Figure A8, we investigate the relationship between narrative optimism, leverage, the capital structure, and payout policy. We find that narrative optimism predicts higher leverage and higher borrowing and has no effect on both equity issuance and payouts. Taken together, this provides evidence that narrative optimism is associated with tighter future financial conditions. This is again inconsistent with a view that narrative optimism drives increases in hiring because it is correlated with positive news about future firm-level financial conditions.

Inspecting the Mechanism: Optimistic Language Predicts Over-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

[23] To 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 A11. 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.

forecasts. As described in Section 3.1, we define variables $\text{GuidanceOptExPost}_{i,t+1}$ and $\text{GuidanceOptExAnte}_{i,t+1}$ 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 $\text{GuidanceOpt}_{i,t+1}$, we estimate the following regression model:

$$\text{GuidanceOpt}_{i,t+1} = \beta \text{opt}_{it} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it} \tag{15}$$

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.[24] Given that we have found that guidance correlates with narrative optimism, it is natural to ask if narrative optimism affects firm decisions conditional on guidance (and *vice versa*). In Appendix E.2, we find that narrative optimism and measured expectations each have predictive power conditional on the other for explaining hiring and capital investment. These results suggest that textual optimism captures aspects of managers' latent beliefs that are not represented in traditional measurement of expectations (here, in guidance data).

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 narratives, we estimate the regression equation implied by our theoretical framework:

$$\Delta \log L_{it} = \sum_{k \in \mathcal{K}} \delta_k \hat{\lambda}_{k,it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \tag{16}$$

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), while constructing standard errors using the approach Belloni, Chernozhukov, Hansen, and Kozbur (2016), to estimate the subset of hiring-relevant narra-

[24] In Table A12, we re-estimate this relationship with alternative measurement schemes. We find a positive relationship between *ex post* optimism and the continuous measure of narrative optimism, and an insignificant relationship between binary or continuous narrative optimism and the continuous difference between guidance and realized sales.

Table 3: Narratives Selected as Relevant for Hiring 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

Notes: Each column lists the narrative variables chosen as relevant regressors in a Rigorous Square-Root LASSO (Belloni, Chernozhukov, Hansen, and Kozbur, 2016) estimation of Equation 16, with the baseline controls as unpenalized regressors. The Shiller (2020) chapters are named for the title of the corresponding book chapter. The topics are named after the three highest-weight bigrams. The corresponding post-LASSO estimates are reported in Table A13.

tives. In Table 3, we list the selected Shiller (2020) and topic narratives. In Appendix Table A13, we report the post-LASSO OLS estimates of Equation 16 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 A13, 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).[25] *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.[26]

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

[25] Appendix Table A3 prints the top ten words per topic and their numerical weights.

[26] We 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 $\delta = 0.0402$ (SE: 0.0044).

equations in which we treat optimism as an endogenous variable, and the (LASSO-selected) Shiller and topic narratives (in sets $\mathcal{K}_S^*$ and $\mathcal{K}_T^*$) as excluded instruments:

$$\begin{aligned}\Delta \log L_{it} &= \delta^{OP}\text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \\ \text{opt}_{it} &= \sum_{k\in\mathcal{K}_S^*} \delta_{Sk}\text{Shiller}_{it}^k + \sum_{k\in\mathcal{K}_T^*} \delta_{Tk}\text{topic}_{it}^k + \tilde{\gamma}_i + \tilde{\chi}_{j(i),t} + \tilde{\tau}'\tilde{X}_{it} + \tilde{\varepsilon}_{it}\end{aligned} \tag{17}$$

where X_{it} are, again, our baseline controls. We provide coefficient estimates for Equation 17 in column 4 of Appendix Table A13. 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 Testing Premise II: Narrative Spread is Viral and Associative

Empirical Strategy. From the conceptual framework in Section 2 (see Equation 5 and Proposition 7 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:

$$\text{opt}_{it} = u \ \text{opt}_{i,t-1} + s \ \overline{\text{opt}}_{t-1} + r \ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it} \tag{18}$$

where $\overline{\text{opt}}_{t-1}$ is average optimism in the previous period, $\Delta \log Y_{t-1}$ is US 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 4, 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 18, 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:

$$\text{opt}_{it} = u_{\text{ind}} \ \text{opt}_{i,t-1} + s_{\text{ind}} \ \overline{\text{opt}}_{j(i),t-1} + s_{\text{peer}} \ \overline{\text{opt}}_{p(i),t-1} + r_{\text{ind}} \ \Delta \log Y_{j(i),t-1} + \gamma_i + \chi_t + \varepsilon_{it} \tag{19}$$

Table 4: Narrative Optimism is Viral and Associative

	(1)	(2)	(3)
	Outcome is opt_{it}		
Own lag, $\text{opt}_{i,t-1}$	0.209	0.214	0.135
	(0.0071)	(0.0080)	(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	0.207
		(0.0396)	(0.0733)
Industry output growth, $\Delta \log Y_{j(i),t-1}$		0.0560	0.0549
		(0.0309)	(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: The regression model is Equation 18 for column 1, and Equation 19 for columns 2 and 3. Aggregate, industry, and peer average optimism are averages of the narrative optimism variable over the respective sets of firms. Industry output growth is the log difference in sectoral value-added calculated from BEA data, linked to two-digit NAICS industries. Standard errors are two-way clustered by firm ID and industry-year.

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 industry 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.[27] The time fixed effect χ_t absorbs aggregate virality and associativeness.

We show the results in columns 2 and 3 of Table 4. 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). In Table A15, we report evidence of stubbornness, virality, and associativeness with the continuous measure of sentiment.[28]

[27] These data are available only from 1997.

[28] In Table A14, we report standard errors for Table 4 under alternative clustering.

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 idiosyncratic optimism, aggregate optimism and GDP are unrelated to other factors that affect changes in optimistic sentiment at the firm level. By using past aggregate optimism in a panel setting, our estimates are not threatened by the reflection problem of Manski (1993). Nevertheless, our estimates may be contaminated by omitted variables bias because aggregate optimism is correlated with common shocks to the economy that are in the error term.

To test for this 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 18, we estimate

$$\begin{aligned}\mathrm{opt}_{it} = u\ \mathrm{opt}_{i,t-1} + s\ \overline{\mathrm{opt}}_{t-1} + \gamma_i + \\ + \sum_{k=-2}^{2}\left(\eta_k^{\mathrm{agg}}\Delta\log Y_{t+k} + \eta_k^{\mathrm{ind}}\Delta\log Y_{j(i),t+k} + \eta_k^{\mathrm{firm}}\Delta\log\theta_{i,t+k}\right) + \varepsilon_{it}\end{aligned} \tag{20}$$

We estimate an analogous specification at the industry level, but with the aggregate leads and lags absorbed. If 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 5, 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). In Table A16, we report analogous estimates with the continuous sentiment variable and find similar results. Taken together, these estimates build confidence that our baseline virality results are not driven by omitted aggregate shocks.

To further test whether our measure of virality captures spillovers, and not omitted common shocks, we pursue two additional instrumental variables strategies. First, in Appendix E.1, we use spillovers from the same plausibly exogenous CEO changes to construct instru-

Table 5: Narrative Optimism is Viral, 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	0.339	0.235	0.222			
	(0.0578)	(0.0763)	(0.1278)	(0.2044)			
Ind. lag, $\overline{\text{opt}}_{j(i),t-1}$					0.276	0.241	0.262
					(0.0396)	(0.0434)	(0.0705)
Firm FE	✓	✓	✓	✓	✓	✓	✓
Time 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: The regression model is Equation 20 for columns 1-4, and an analogous industry-level specification for columns 5-7 (*i.e.,* Equation 19 with past and future controls). Columns 1 and 5 are "baseline estimates" corresponding, respectively, with columns 1 and 3 of Table 4. The added control variables are two leads, two lags, and the contemporaneous value of: real GDP growth (columns 2-4), industry-level output growth (columns 3-4 and 6-7), and firm-level TFP growth (columns 4 and 7). Standard errors are two-way clustered by firm ID and industry-year.

ments for industry and peer-set optimism. We find similar point estimates as in our main analysis. Second, in Appendix E.4, we use size-weighted idiosyncratic shocks to firm-level optimism as an instrument for aggregate size-weighted optimism (a granular IV à la Gabaix and Koijen, 2020). While not comparable to our main estimates as the measure of spillovers is different, we recover a statistically significant virality effect.

The Spread of Hiring-Relevant Narratives. We repeat the estimation of our equation measuring aggregate associativeness and virality, Equation 18, 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 A9. 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, Comination" (relating to firm origination), and "Technology, Revenue, Development" (relating to R&D) having $r > 0$. In Appendix Table A17, we instrument for optimism with the other 13 hiring-relevant narratives in the estimation of Equations 18 and 19. We 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 decision-relevance, virality, and associativeness 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 described in Woodford (2003) and Galí (2008). Time is discrete and infinite, indexed by $t \in \mathbb{N}$. 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:

$$x_{it} = \theta_{it} L_{it}^{\alpha} \tag{21}$$

where $\alpha \in (0, 1]$ 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

$$Y_t = \left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}} \, \mathrm{d}i \right)^{\frac{\epsilon}{\epsilon-1}} \tag{22}$$

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

A representative household consumes final goods C_t and supplies labor $\{L_{it}\}_{i \in [0,1]}$ to the intermediate goods firms with isoelastic, separable, expected discounted utility preferences:

$$\mathcal{U}\left(\{C_t, \{L_{it}\}_{i\in[0,1]}\}_{t\in\mathbb{N}}\right) = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\gamma}}{1-\gamma} - \int_{[0,1]} \frac{L_{it}^{1+\psi}}{1+\psi} \, \mathrm{d}i \right) \right] \tag{23}$$

where $\gamma \in \mathbb{R}_+$ indexes the size of income effects in the household's supply of labor and $\psi \in \mathbb{R}_+$ indexes their inverse Frisch labor supply elasticity to each firm.

Finally, we define the composite parameter:

$$\omega = \frac{\frac{1}{\epsilon} - \gamma}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \tag{24}$$

which indexes the strength of strategic complementarity. So that complementarity is positive but not so extreme that the model features multiple equilibria, we assume that $\omega \in [0, 1)$. This requires that income effects in labor supply do not overwhelm aggregate demand externalities and that these externalities are not too large.

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 $\tilde{\theta}_{it}$:

$$\theta_{it} = \tilde{\theta}_{it}\gamma_i\theta_t \tag{25}$$

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

$$s_{it} = \log\theta_t + \varepsilon_{it} \tag{26}$$

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:[29]

$$\log\theta_t \sim N(\mu, \sigma^2) \tag{27}$$

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

[29]We show in Online Appendix B.4 that our analysis is unchanged if narratives instead pertain to beliefs about idiosyncratic productivity.

firm's prior belief regarding the fundamental can be described as:

$$\pi_{it}(\lambda_{it}) = N\left(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P, \sigma^2\right) \tag{28}$$

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 $Q_t = \int_{[0,1]} \lambda_{it}\,\mathrm{d}i$ 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:

$$Q_{t+1} = Q_t P_O(\log Y_t, Q_t) + (1-Q_t)P_P(\log Y_t, Q_t) \tag{29}$$

This aggregates the behavior of individual firms whose updating of *models* depends on aggregate conditions $(\log Y_t, Q_t)$, and whose beliefs about economic variables are determined via Bayes' rule, conditional on those models, as described above. In view of our evidence that firms update associatively and virally (Section 4.2), we assume that P_O and P_P are both increasing functions. In view of our evidence that firms are stubborn, or that optimism is persistent at the firm level, we assume that $P_O \geq P_P$. For technical reasons, we assume that P_O and P_P are continuous and almost everywhere differentiable.

These conditions, motivated by the data, rule out some models for how individual firms update their beliefs, among the class of models that aggregates to Equation 29. An important such model that is ruled out is one in which firms observe aggregate variables $\log Y_t$ and Q_t and use Bayes' rule to update their beliefs over models. As we formalize in Appendix B.1, this "Bayesian benchmark" rules out dependence on Q_t, conditional on $\log Y_t$, in firms' updating ("virality"). Moreover, this "Bayesian benchmark" predicts that agents converge to holding the better-fitting empirical model exponentially quickly, which is at odds with our finding of cyclical dynamics for aggregate optimism (Figure 1). However, richer Bayesian models that are consistent with our empirical results can be nested by our reduced-form updating probabilities.

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:

$$P_O(\log Y, Q) = \left[\frac{u}{2} + r\log Y + sQ\right]_0^1 \quad \text{and} \quad P_P(\log Y, Q) = \left[-\frac{u}{2} + r\log Y + sQ\right]_0^1 \tag{30}$$

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

5.4 Equilibrium

An equilibrium is a path for all variables:

$$\mathcal{E} = \left\{Y_t, C_t, Q_t, \theta_t, \{L_{it}, x_{it}, p_{it}, w_{it}, \lambda_{it}, s_{it}, \tilde{\theta}_{it}\}_{i\in[0,1]}\right\}_{t\in\mathbb{N}} \tag{31}$$

such that (i) narrative weights λ_{it} follow a Markov process consistent with Equation 29 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 $\mathcal{E}$, (iii) L_{it} is consistent with production technology (Equation 21) 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 22, (vii) C_t satisfies goods market clearing, $C_t = Y_t$, and (viii) Q_t evolves according to Equation 29.

6 Theoretical Results

We now study the equilibrium dynamics of narratives and output. We find that narratives induce non-fundamental fluctuations in the economy and have the potential to generate hysteresis. Moreover, we show that our empirical estimates identify the model. We use this mapping to the data to quantify and test the model's predictions in Section 7.

6.1 Characterizing Equilibrium Dynamics

To solve for equilibrium production, it suffices to solve for intermediates goods production. These firms maximize expected profits, as priced by the representative household:

$$\Pi = \mathbb{E}_{it}[C_t^{-\gamma}\left(p_{it}x_{it} - w_{it}L_{it}\right)] \tag{32}$$

where $C_t^{-\gamma}$ 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:

$$p_{it} = Y_t^{\frac{1}{\epsilon}} x_{it}^{-\frac{1}{\epsilon}} \tag{33}$$

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:

$$L_{it}^{\psi} = w_{it} C_t^{-\gamma} \tag{34}$$

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

$$L_{it} = \theta_{it}^{-\frac{1}{\alpha}} x_{it}^{\frac{1}{\alpha}} \tag{35}$$

Finally, by imposing goods market clearing $C_t = Y_t$ and substituting Equations 33, 34, and 35 into Equation 32, we obtain that the intermediates goods firms solve the following profit maximization problems:

$$\max_{x_{it}} \mathbb{E}_{it} \left[Y_t^{-\gamma} \left(Y_t^{\frac{1}{\epsilon}} x_{it}^{1-\frac{1}{\epsilon}} - Y_t^{\gamma} \theta_{it}^{-\frac{1+\psi}{\alpha}} x_{it}^{\frac{1+\psi}{\alpha}} \right) \right] \tag{36}$$

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

$$\left(1 - \frac{1}{\epsilon}\right) \mathbb{E}_{it} \left[Y_t^{\frac{1}{\epsilon} - \gamma} \right] x_{it}^{-\frac{1}{\epsilon}} = \frac{1+\psi}{\alpha} \mathbb{E}_{it} \left[\theta_{it}^{-\frac{1+\psi}{\alpha}} \right] x_{it}^{\frac{1+\psi-\alpha}{\alpha}} \tag{37}$$

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. In this equation, a given firm's narrative affects their expected marginal costs of production, via the expectation of idiosyncratic productivity, and their expected marginal benefits of production, via the expectation of aggregate output (which encompasses aggregate demand externalities, asset pricing forces, and wage pressure). Moreover, in equilibrium, the distribution of narratives in the population affects the level of aggregate output and agents' expectations thereof.

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:

$$\begin{aligned}\log Y_t = \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\Bigg[\exp\Bigg\{\frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg(\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\ -\log\mathbb{E}_{it}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right]+\log\mathbb{E}_{it}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)\log Y_t\right\}\right]\Bigg)\Bigg\}\Bigg]\end{aligned} \tag{38}$$

where the outer expectation operator integrates over the realizations of productivity shocks $(\tilde{\theta}_{it}, \gamma_i)$, 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:*

$$\log Y(\log\theta_t, Q_t) = a_0 + a_1\log\theta_t + f(Q_t) \tag{39}$$

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 38 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 $\varepsilon-$equilibrium for any $\varepsilon > 0$ for some sufficiently large support for fundamentals (Lemma 3). Hence, the quasi-linear equilibrium we study is the limit of the unique equilibrium with bounded fundamentals as the bound becomes large. This justifies our restriction in analyzing this class of equilibrium. △

Narratives Drive Non-Fundamental Fluctuations in Aggregate Output. The coefficient a_1 and function f respectively 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

$$f(1) = \frac{\alpha\delta^{OP}}{1-\omega} \tag{40}$$

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.

Identification of Model Parameters. We now show how our empirical strategy identifies the aggregate behavior of output and optimism.

Corollary 1 (Identification of Model Parameters)**.** *In equilibrium, firms' hiring decisions obey the following equation:*

$$\Delta \log L_{it} = c_{0,i} + c_1 \log \theta_t + c_2 f(Q_t) + c_3 \log \theta_{it} + c_4 \log L_{i,t-1} + \delta^{OP} \lambda_{it} + \zeta_{it} \tag{41}$$

where ζ_{it} is an IID normal random variable with zero mean. Thus, conditional on $(\alpha, \epsilon, \gamma, \psi)$, δ^{OP} uniquely identifies f, the equilibrium effect of optimism on aggregate output.

Proof. See Appendix A.4. □

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 time fixed-effect. As we previously discussed, aggregate fundamentals also appear in the time fixed-effect of the regression. These facts highlight formally the necessity of combining cross-sectional variation and some structural model for general-equilibrium interaction to identify the effect of optimism on economic outcomes.

The Dynamics of Optimism. Finally, using Proposition 1, we express the dynamics of the economy as a first-order nonlinear stochastic difference equation for aggregate optimism:

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*

$$T(Q_t, \log \theta_t) = Q_t P_O(a_0 + a_1 \log \theta_t + f(Q_t), Q_t) + (1 - Q_t) P_P(a_0 + a_1 \log \theta_t + f(Q_t), Q_t) \tag{42}$$

Proof. See Appendix A.5 □

6.2 Dynamics: Steady-State Multiplicity and Hysteresis

We next characterize the steady states of optimism and their stability, for fixed aggregate fundamentals. This analysis highlights how associative, viral optimism affects dynamics even in the absence of shocks.

Steady-State Characterization. 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_θ^* is a deterministic steady state for level of productivity θ if it is a fixed point of the corresponding map, $T_\theta(Q_\theta^*) = Q_\theta^*$. The following result establishes that a deterministic steady state always exists and provides necessary and sufficient conditions for extreme optimism and pessimism to be (stable) 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* $\theta \in \Theta$
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:*

$$\theta_P = \exp\left\{\frac{P_P^{-1}(0;0) - a_0}{a_1}\right\} \quad \text{and} \quad \theta_O = \exp\left\{\frac{P_O^{-1}(1;1) - a_0 - f(1)}{a_1}\right\} \tag{43}$$

   where $P_P^{-1}(x;Q) = \sup\{Y : P_P(Y,Q) = x\}$ *and* $P_O^{-1}(x;Q) = \inf\{Y : P_O(Y,Q) = x\}$*.*[30]
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. □

This result 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 $P_O^{-1}(1;1)$.

Hysteresis. Proposition 2 demonstrates the possibility for hysteresis: multiple steady states of optimism that are entirely self-fulfilling. 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:

[30] With the convention that the infimum of an empty set is $+\infty$ and the supremum of an empty set is $-\infty$.

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

$$P_O^{-1}(1;1) - P_P^{-1}(0;0) \leq f(1) \tag{44}$$

Proof. See Appendix A.7 □

To gain intuition for these results, and to derive a parametric condition for hysteresis that we will later empirically assess, we compute these conditions 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:

$$1 \leq u + s + rf(1) \tag{45}$$

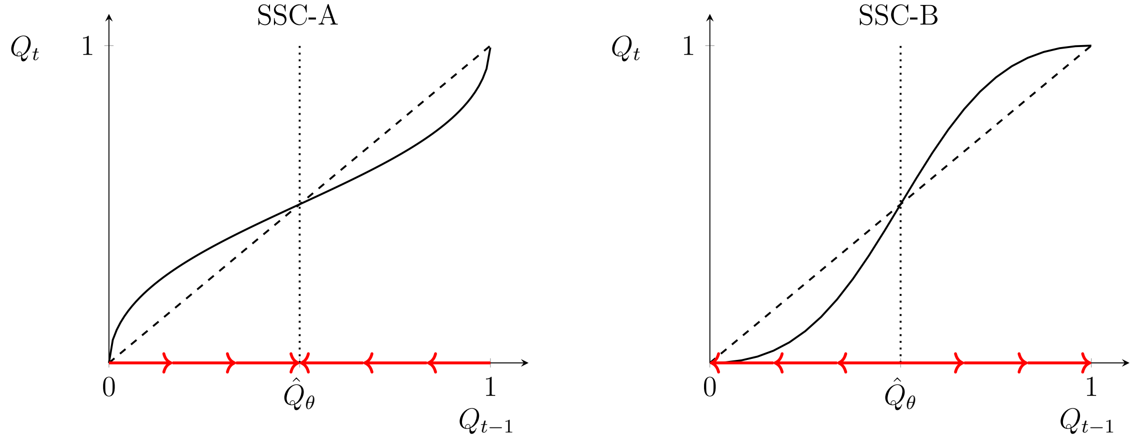
which is to say that stubbornness, associativeness, virality, and the equilibrium impact of optimism on output are sufficiently large. In Section 7.3, we will empirically assess this condition and its quantitative implications in our calibration. △

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 $\theta \in \Theta$ there exists $\hat{Q}_\theta \in [0,1]$ such that $T_\theta(Q) > Q$ for all $Q \in (0, \hat{Q}_\theta)$ and $T_\theta(Q) < Q$ for all $Q \in (\hat{Q}_\theta, 1)$. We say that T is *strictly single-crossing from below* (SSC-B) if for all $\theta \in \Theta$ there exists $\hat{Q}_\theta \in [0,1]$ such that $T_\theta(Q) > Q$ for all $Q \in (\hat{Q}_\theta, 1)$ and $T_\theta(Q) < Q$ for all $Q \in (0, \hat{Q}_\theta)$. 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.

Lemma 1 (Steady States under the SSC Property)**.** *If T_θ 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 $Q = \hat{Q}_\theta$. Moreover, when T_θ 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 $\hat{Q}_\theta$, they are unstable with respective basins of attraction $\{0\}$ and $\{1\}$. When T_θ is SSC-B: whenever extreme optimism is a steady state, it is stable with basin of attraction $(\hat{Q}_\theta, 1]$; whenever extreme pessimism is a steady state it is stable with basin of attraction $[0, \hat{Q}_\theta)$; and intermediate optimism is always unstable with basin of attraction $\{\hat{Q}_\theta\}$.*

Proof. See Appendix A.8 □

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



Notes: In each subfigure, the solid line is an example transition map T_θ, the dashed line is the 45-degree line, the dotted vertical line indicates the interior steady state $\hat{Q}_\theta$, 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.

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 $\hat{Q}_\theta \in [0, 1]$, 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 understand the determinants of this criticality threshold, we study our LAV example:

Example 1 (continuing from p. 29)**.** The LAV model satisfies SSC-B if u, r, and s are sufficiently large and $\theta \in (\theta_O, \theta_P)$. Moreover, in the SSC-B case, the criticality threshold is given by the unique solution to the equation:

$$\hat{Q}_\theta = u(2\hat{Q}_\theta - 1) + s\hat{Q}_\theta + r(a_0 + a_1 \log \theta + f(\hat{Q}_\theta)) \tag{46}$$

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

$$\hat{Q}_\theta \approx \frac{\frac{u}{2} - r(a_0 + a_1 \log \theta)}{u + s + rk - 1} \tag{47}$$

Hence, greater virality, associativeness and decision relevance make the criticality threshold lower and therefore make it easier for an epidemic of extreme optimism to take hold. △

6.3 Impulse Responses and Stochastic Fluctuations

Having characterized narrative dynamics with fixed fundamentals, we now study how the economy responds to deterministic and stochastic fundamental shocks. For this analysis, we restrict attention to the SSC class, noting that this is an assumption solely on primitives.[31]

Hump-Shaped and Discontinuous Impulse Responses. 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 $\theta = 1$:

$$\theta_t = \begin{cases} 1, & t = 0, \\ \hat{\theta}, & t = 1, \\ 1, & t \geq 2. \end{cases} \tag{48}$$

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.

Proposition 3 (SSC-A Impulse Response Functions). *In the SSC-A case, suppose that $Q_0 = \hat{Q}_1 \in (0, 1)$. The impulse response of the economy is given by:*

$$\log Y_t = \begin{cases} a_0 + f(\hat{Q}_1), & t = 0, \\ a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \qquad Q_t = \begin{cases} \hat{Q}_1, & t \leq 1, \\ Q_2, & t = 2, \\ T_1(Q_{t-1}), & t \geq 3. \end{cases} \tag{49}$$

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 $Q_t \to \hat{Q}_1$. The IRF is hump-shaped if and only if $\hat{\theta} < \exp\{(f(Q_2) - f(\hat{Q}_1))/a_1\}$.

Proof. See Appendix A.9 □

[31] This 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 $\{I_j\}_{j \in \mathcal{J}}$ such that $\cup_{j \in \mathcal{J}} I_j = [0, 1]$ and the restricted functions $T_{\theta,j} : I_j \to [0, 1]$ 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 $j \in \mathcal{J}$. Thus, applying our results to these restricted functions, we have a complete description of the global dynamics.

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 with the SSC-B case, wherein impulse responses can be discontinuous in the shock size. 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 is given by:*

$$\log Y_t = \begin{cases} a_0, & t = 0, \\ a_0 + a_1 \log \hat{\theta}, & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \qquad Q_t = \begin{cases} 0, & t \leq 1, \\ P_P(a_0 + a_1 \log \hat{\theta}, 0), & t = 2, \\ T_1(Q_{t-1}), & t \geq 3. \end{cases} \tag{50}$$

These impulse responses fall into the following four exhaustive cases:

1. *$\hat{\theta} \leq \theta_P$, No Lift-Off: $Q_t = 0$ for all $t \in \mathbb{N}$.*
2. *$\hat{\theta} \in (\theta_P, \theta^*)$, Transitory Impact: Q_t is monotonically declining for all $t \geq 2$ and $Q_t \to 0$.*
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 $Q_t \to 1$*

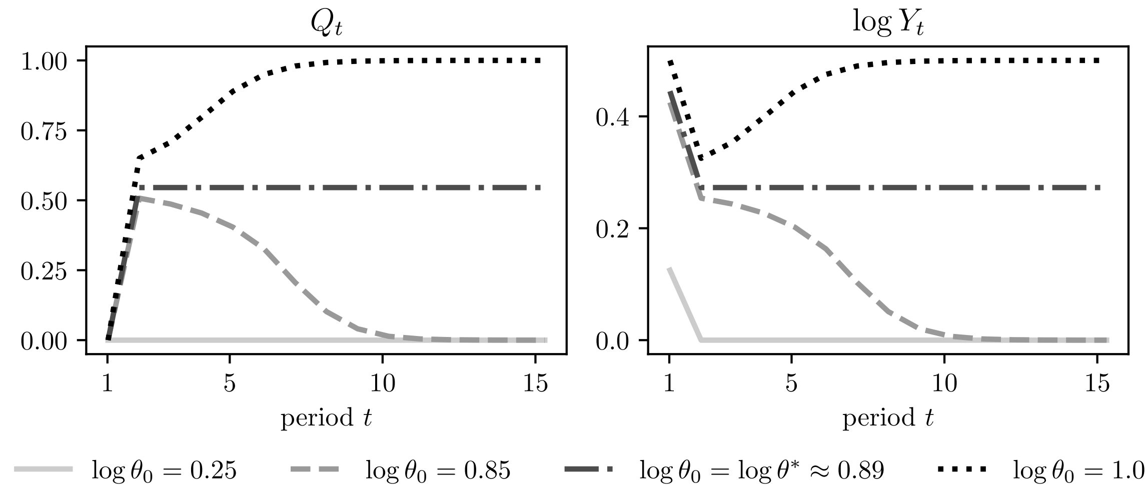
where the critical shock threshold is $\theta^* = \exp\{(P_P^{-1}(\hat{Q}_1; 0) - a_0)/a_1\} > \theta_P$. In the transitory case, the output IRF is hump-shaped if and only if $\hat{\theta} < \exp\{f(P_P(a_0 + a_1 \log \hat{\theta}, 0))/a_1\}$.

Proof. See Appendix A.10. □

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

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



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

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.

Stochastic Boom-Bust Cycles. Having characterized the deterministic impulse propagation mechanisms at work in the economy, we now turn to understand 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:

$$T_{PO} = \mathbb{E}_H \left[\min\{\tau \in \mathbb{N} : Q_\tau = 1\} | Q_0 = 0 \right] , \; T_{OP} = \mathbb{E}_H \left[\min\{\tau \in \mathbb{N} : Q_\tau = 0\} | Q_0 = 1 \right] \quad (51)$$

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:*

$$
\begin{aligned}
T_{PO} &\leq \frac{1}{1 - H\left(\exp\left\{\frac{P_P^{\dagger}(1;0) - a_0}{a_1}\right\}\right)} \\
T_{OP} &\leq \frac{1}{H\left(\exp\left\{\frac{P_O^{\dagger}(0;1) - a_0 - f(1)}{a_1}\right\}\right)}
\end{aligned}
\tag{52}
$$

where $P_P^{\dagger}(x;Q) = \inf\{Y : P_P(Y,Q) = x\}$ *and* $P_O^{\dagger}(x;Q) = \sup\{Y : P_O(Y,Q) = x\}$*. Moreover, when* $P_O^{\dagger}(0;1) - P_P^{\dagger}(1;0) \leq f(1)$*, 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.[32]

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 $H\left(\exp\left\{\frac{P_O^{\dagger}(0;1) - a_0 - f(1)}{a_1}\right\}\right)$ 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 $P_O^{\dagger}(0;1)$, and when co-ordination motives are weak as measured by $f(1)$.

6.4 Additional Results and Extensions

Before proceeding to our quantification of the model, we briefly summarize additional results and extensions contained within the Appendix.

Welfare Implications. So far, we have studied the positive implications of fluctuations in optimism. In Appendix B.2, we study the normative implications of optimism and provide conditions under which its presence is welfare improving, despite it being misspecified.

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

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.

Continuous Optimism. Our baseline model featured, as in our main empirical specifications, only two narratives. In Appendix B.3, we generalize the setting studied in this section to feature a continuum of models. We show that very similar dynamics for both real output and narratives obtain in this specification of the model and the condition for extremal multiplicity is almost identical. Thus, the qualitative and quantitative features of the baseline model carry over to this richer setting.

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.5, 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 qualitatively similar dynamics obtain.

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.6, we extend the multi-dimensional narrative equilibrium characterization when we 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. This matters for equilibrium output only insofar as it induces a time-varying covariance between and optimism and productivity. We find no empirical evidence for cyclicality of this covariance. We therefore abstract from this channel in our quantitative analysis.

Contrarianism, Endogenous Cycles, and Chaos. While this model generates narratively driven fluctuations, it cannot generate fully endogenous cycles and chaotic dynamics. In Appendix B.7, we extend this model to allow for contrarianism and the possibility that pessimists may be more likely to become optimists than optimists are to remain optimists. Allowing for these features generates the possibility of endogenous cycles of arbitrary period and topological chaos (sensitivity to arbitrarily small changes in initial conditions). This model also admits a structural test for the presence of cycles and chaos that we bring directly to the data and reject at the 95% confidence level that either cycles or chaos obtain.

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.8, 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 and their qualitative properties. 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. We reject the condition for extremal multiplicity and hysteresis for optimism, but fail to reject it for other decision-relevant narratives. Taken together, we therefore find that viral narratives may explain a significant fraction of the US business cycle.

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 6 and additional details in Appendix F.

First, we have shown in Section 6 that, to identify the static relationship between output and optimism, we need to estimate f. In turn, f requires knowledge of: δ^{OP}, the partial-equilibrium effect of optimism on hiring; α, the returns-to-scale parameter; ϵ, the elasticity of substitution between varieties; and ω, the extent of complementarity (which itself depends on γ, indexing income effects in labor supply, and ψ, the inverse Frisch elasticity of labor supply). In our main analysis, we combine our baseline regression estimate of $\hat{\delta}^{OP} = 0.0355$ (see Table 1) with an external calibration of α, ϵ, γ, and ψ, which together also pin down ω. In Section 7.4, we study the sensitivity of our results to this external calibration, and we introduce two other calibration strategies for complementarity: using estimates of demand multipliers from the literature and inferring a demand multiplier for optimism using our own firm-level regressions.

Table 6: Model Calibration

Fixed	ϵ	Elasticity of substitution	2.6
	γ	Income effects in labor supply	0
	ψ	Inverse Frisch elasticity	0.4
	α	Returns-to-scale	1
Calibrated	$\mu_O - \mu_P$	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	0.208
	r	Associativeness	0.804
	s	Virality	0.290

Notes: "Fixed" parameters are externally set. "Calibrated" parameters are chosen to hit various moments. Our specific calibration methods are described in Section 7.1.

For the external calibration, 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 US 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 inverse Frisch elasticity of labor supply $\psi = 0.4$, which is within the range of standard macroeconomic estimates (Peterman, 2016). 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 $\omega = 0.49$. Finally, we observe that the calibration of δ^{OP} puts only one 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.

Second, we calibrate the process for updating probabilities. We use our regression estimates of the LAV form (see Equation 30) which corresponds to the linear probability model (see Table 4).[33] This yields values of $u = 0.208$ for stubbornness, $r = 0.804$ for associativeness, and $s = 0.290$ for virality.

Third, 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

[33] While the linear probability model does not necessarily yield probabilities between zero and one, our estimates of u, r and s imply updating probabilities that are always between zero and one so long as output does not deviate by more than 30%.

persistent fundamentals based on the analysis in Appendix B.5. Concretely, we suppose that $\log \theta_t$ is a Gaussian AR(1) process with persistence ρ and IID innovations $u_t \sim N(0, \sigma^2)$:

$$\log \theta_t = \rho \log \theta_{t-1} + u_t \tag{53}$$

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

$$\log Y_t^f = \log Y_t - f(Q_t) \tag{54}$$

In Appendix F, 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 US 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 Static Effects of Optimism on US GDP

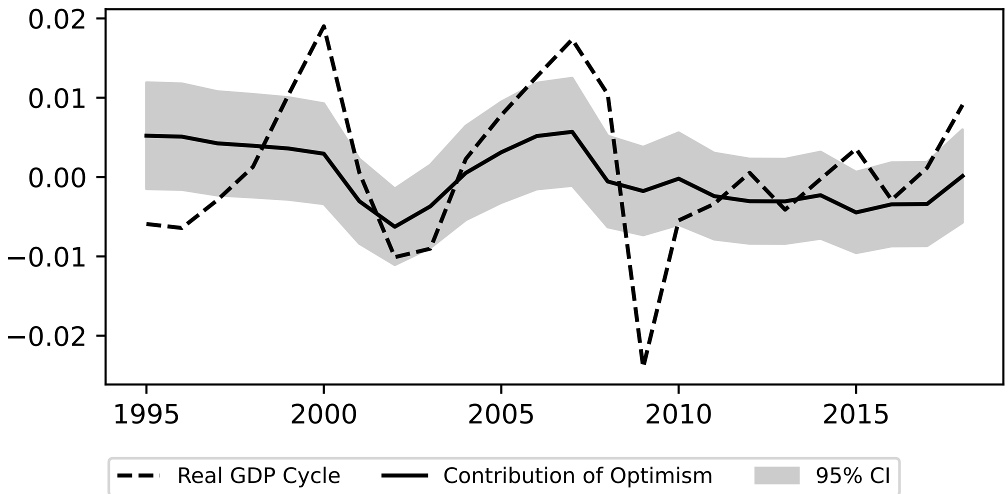
We first use our calibration to estimate the contribution of measured aggregate optimism to business cycles. To do this, we take our measured aggregate optimism time series and compute its output effects via our calibrated f function. These estimates rely solely on our estimated partial equilibrium response 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). The estimates do not rely on the estimated parameters governing narrative updating.

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

$$\% \text{ Explained}(t_0, t_1) = 100 \cdot \frac{f(Q_{t_1}) - f(Q_{t_0})}{\log Y_{t_1} - \log Y_{t_0}} \tag{55}$$

[34]We apply the Baxter and King (1999) band-pass filter to post-war quarterly US 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.

Figure 5: The Effect of Optimism on Historical US GDP



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 log output. The 95% confidence interval incorporates uncertainty from the calibration of δ^{OP} using the delta method.

where Q is measured optimism, $\log Y$ is the measured cyclical component of log real GDP, and (t_0, t_1) are the endpoints. We report these results in Table 7. 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 narrative optimism contributed significantly to recent economic downturns.

7.3 Viral Narratives and Economic Fluctuations

Our theoretical analysis delimited two qualitatively different regimes for macroeconomic dynamics with viral optimism: one with stochastic fluctuations around a stable steady state, and one with hysteresis and (almost) global convergence to extreme steady states. For the LAV case which we have taken to the data, the condition for the second (SSC-B) case is given by Equation 45. We compute the empirical analog of this condition as:

$$M = \hat{u} + \hat{s} + \hat{r}\hat{f}(1) - 1 \tag{56}$$

If $M > 0$, the calibrated model features hysteresis in the dynamics of optimism and output; if $M < 0$, the model features oscillations around a stable steady state. We find $M = -0.44 < 0$ with a standard error of 0.052, implying stable oscillations (the SSC-A case). This reflects

Table 7: The Effect of Optimism on US Recessions

	Change (%)		
Period	Detrended GDP	Optimism Component $f(Q_t)$	% Explained
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: The first column gives the change in detrended, annualized real GDP over the stated periods. The second column gives the component of this change attributed to the change in aggregate optimism by the model. The third column is the fraction of the real GDP change explained by optimism, defined in Equation 55. Standard errors for columns 2 and 3 incorporate uncertainty from estimating δ^{OP} and are calculated using the delta method.

the fact that decision-relevance, stubbornness, virality, and associativeness are sufficiently small.

We now investigate how these fluctuations in viral optimism contribute to the 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:

$$\hat{c}_Q(\ell) := 100 \cdot \frac{\text{Cov}[\log Y_t, \log Y_{t-\ell}] - \text{Cov}[\log Y_t^f, \log Y_{t-\ell}^f]}{\text{Cov}[\log Y_t, \log Y_{t-\ell}]} \tag{57}$$

We report our findings in Table 8. 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.

While we found that the model of narrative optimism was consistent with stable fluctuations, this conclusion could have been overturned were the narrative more persistent via higher stubbornness or higher virality. In Figure 6, we plot our point estimate of virality and stubbornness as a plus and its 95% confidence interval as a dotted ellipse. We also plot, as a dashed line, the condition for $M = 0$, fixing r and $f(1)$; to the left of this line, $M < 0$, and to the right of this line, $M > 0$. Even though our estimates of virality and stubbornness were somewhat imprecise, they nevertheless rule out the possibility of hysteresis dynamics for optimism at the 5% level. In the same plot, we also indicate via colors $\hat{c}_Q(0)$, the fraction of variance explained by optimism. We find that this is numerically stable over the 95% confidence set.

Table 8: 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: The first row gives the autocovariance (multiplied by 100^2 for ease of reading) of log output from the model simulation at the indicated lags. The second row gives the fraction of autocovariance explained by endogenous optimism, as defined in Equation 57.

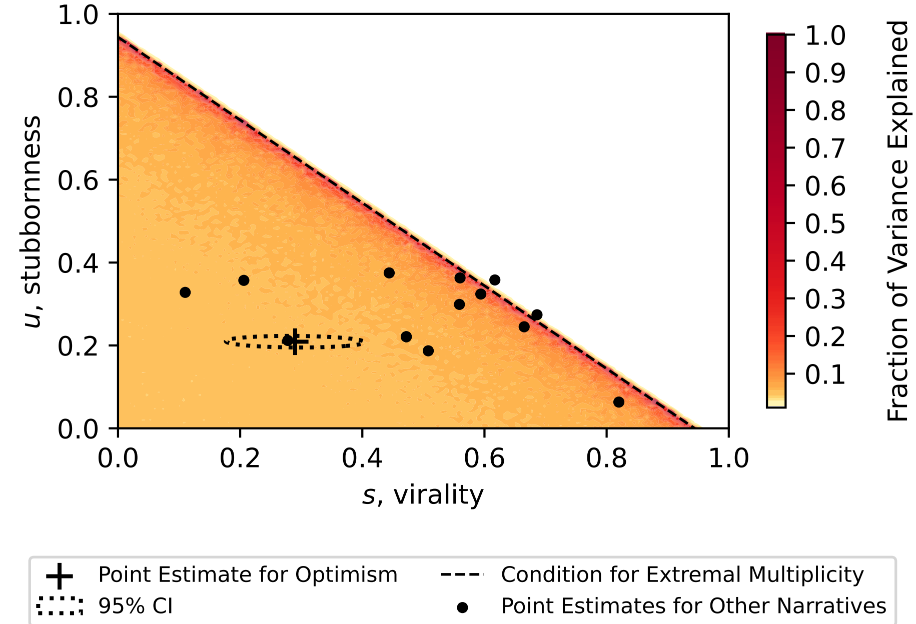
To the immediate left of the boundary line for $M = 0$, $\hat{c}_Q(0)$ becomes very large. This is because, when this condition is close to holding or failing, all steady states are close to being unstable, so optimism is volatile. To the right of this boundary, optimism explains essentially zero output 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? As one benchmark for the possible values of stubbornness and virality, we consider the estimated values for our other decision-relevant narratives. We plot the associated point estimates of (s, u) for the thirteen decision-relevant narratives as black dots in Figure 6. Several of the thirteen plotted points are close to the condition for extremal multiplicity. Two are across the threshold. Thus, if optimism were to counterfactually have the stubbornness and virality of either of these two narratives, the joint dynamics of optimism and output would feature hysteresis.[35]

Other Narratives and Hysteresis. We can similarly perform tests for the possibility of hysteresis for the other decision-relevant narratives. To do this, we compute M using the estimated stubbornness, associativeness, virality, and partial equilibrium effects on hiring of the other 13 narratives. We plot the results of this in Figure A10. We find that many of our narratives have values for M very close to zero, while one narrative's 95% confidence interval for M contains zero. Thus, while we can reject that optimism features hysteresis dynamics, we cannot do so for all of the narratives that we consider. Heuristically, this finding matches up with the corresponding time series (see Figures 1, A2, A3): while optimism appears to fluctuate in a stable fashion, other narratives have time series that appear to undergo regime shifts, with many notably happening around the Great Recession.

[35]The narratives across the line are Topic 3 (Value, Fair, Loss) and Topic 6 (Stock, Compensation, Tax).

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 shading corresponds to the fraction of variance explained by optimism, or $\hat{c}_Q(0)$ defined in Equation 57. The plus is our calibrated value from Table 6, and the dotted line is the boundary of a 95% confidence set. The dots are calibrated values for other narratives from Figure A9. The dashed line is the condition of extremal multiplicity from Corollary 3 and Equation 45.

7.4 Parametric Sensitivity Analysis

We finally consider how our calibration of macroeconomic parameters matters for our findings. Our findings are summarized in Table A18, which reports, under the alternative parameterizations detailed below, the fraction of output variance explained by optimism, the fraction of output autocovariance explained by optimism, and the fractions of the 2000-02 and 2007-09 downturns explained by optimism.

We first focus on the calibration of macroeconomic complementarity and, by extension, the demand multiplier. Recall that $f(Q) \approx \frac{\alpha\delta^{OP}}{1-\omega}Q$, where $\frac{1}{1-\omega}$ 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 $\frac{1}{1-\omega} = 1.96$. Our numerical results from adjusting the multiplier, holding fixed $(\delta^{OP}, \alpha, \epsilon)$, convey that the contribution of optimism is monotone in this number; moreover, for our explanation of observed recessions, this relationship is almost exactly one-to-one due to the near-linearity of f in Q.

In row 1 of Table A18, we report results with $\psi = 2.5$ or a Frisch elasticity of $1/\psi = 0.4$, which, while implying macroeconomically counterfactual behavior of wages and employment, is more consistent with microeconomic estimates (Peterman, 2016). This reduces the multiplier to $\frac{1}{1-\omega} = 1.15$. In row 2, we consider $\gamma = 1.0$ which generates a multiplier less than one, $\frac{1}{1-\omega} = 0.56$. We next consider a strategy of calibrating the multiplier from structural studies in the literature. In row 3, we target the estimated multiplier of 1.33 from Flynn, Patterson, and Sturm (2022), which is based on a structural model of network linkages in the modern US economy and, consistent with our application, is for an *unfinanced* demand shock.[36] In row 4, we consider a calibration of the demand multiplier using our own data. The exact method, explained in Appendix F.3, boils down to re-estimating the effect of optimism on hiring (Equation 12), *without* a time fixed effect to soak up the "missing intercept," and instead adding a parametric control for aggregate productivity (using the data of Fernald, 2014). Our model gives a formula for the bias in the estimate of δ^{OP} (*i.e.*, the contamination with GE effects), from which we can back out the demand multiplier, conditional on having correctly measured aggregate TFP. This procedure yields an estimate of $\frac{1}{1-\omega} = 1.46$, close to our calibration from Flynn, Patterson, and Sturm (2022).

We finally consider sensitivity to the calibrations of the elasticity of substitution ϵ (row 5 of Table A18) and the returns-to-scale α (row 6 of Table A18) holding fixed the multiplier (via adjustment in ψ). Changing ϵ has close to no effect on our results, due to the aforementioned near-linearity of f. Reducing α, or assuming decreasing returns to scale, dampens the effect of optimism on output because it implies a smaller production effect of our estimated effect of optimism on labor.

8 Conclusion

This paper studies the macroeconomic implications of viral, belief-altering narratives. We develop a conceptual framework in which narratives form building blocks of agents' models of exogenous and, in equilibrium, endogenous variables and spread virally and associatively 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 non-fundamentally driven boom-bust cycles, hysteresis, and impulse responses that are hump-shaped over time and discontinuous in the sizes of shocks. When we calibrate the model to match the data, we find that the business-cycle implications of narratives are quantitatively

[36] By contrast, empirical estimates based on the economy's response to fiscal shocks (Ramey, 2011; Chodorow-Reich, 2019) do not map to the required multiplier as they include the effects of financing.

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.

Acemoglu, D., and J. A. Robinson (2021): "Culture, institutions and social equilibria: A framework," Discussion Paper 28832, National Bureau of Economic Research.

Adam, K., A. Marcet, and J. Beutel (2017): "Stock price booms and expected capital gains," *American Economic Review*, 107(8), 2352–2408.

Allen, F., S. Morris, and H. S. Shin (2006): "Beauty contests and iterated expectations in asset markets," *The Review of Financial Studies*, 19(3), 719–752.

Andre, P., I. Haaland, C. Roth, and J. Wohlfart (2022): "Narratives about the macroeconomy," Discussion Paper DP17305, CEPR.

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.

Ash, E., G. Gauthier, and P. Widmer (2021): "RELATIO: Text semantics capture political and economic narratives," Discussion Paper 2108.01720, arXiv.

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 28347, 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.

Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2016): "Inference in high-dimensional panel models with an application to gun control," *Journal of Business & Economic Statistics*, 34(4), 590–605.

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 2763942, SSRN.

Blanchard, O. J., and N. Kiyotaki (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.

Bordalo, P., N. Gennaioli, A. Shleifer, and S. J. Terry (2021): "Real credit cycles," Working Paper 28416, National Bureau of Economic Research.

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.

Brock, W. A., and C. H. Hommes (1997): "A rational route to randomness," *Econometrica*, 65(5), 1059–1095.

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," Working Paper 29344, 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.

Carroll, C. D. (2001): "The epidemiology of macroeconomic expectations," Working Paper 8695, National Bureau of Economic Research.

Carroll, C. D., and T. Wang (2022): "Epidemiological Expectations," in *Handbook of Economic Expectations*, vol. 1, p. Forthcoming. Elsevier.

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.

Chodorow-Reich, G. (2019): "Geographic cross-sectional fiscal spending multipliers: What have we learned?," *American Economic Journal: Economic Policy*, 11(2), 1–34.

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.

Fernald, J. (2014): "A quarterly, utilization-adjusted series on total factor productivity," Accessed from `https://www.johnfernald.net/`.

Flynn, J. P., C. Patterson, and J. Sturm (2022): "Fiscal policy in a networked economy," Working Paper 29619, National Bureau of Economic Research.

Flynn, J. P., and K. A. Sastry (2022): "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. Iijima, and Y. Ishii (2020): "Stability and robustness in misspecified learning models," Discussion paper, Cowles Foundation.

Fudenberg, D., G. Lanzani, and P. Strack (2021): "Limit points of endogenous misspecified learning," *Econometrica*, 89(3), 1065–1098.

Gabaix, X., and R. S. Koijen (2020): "Granular instrumental variables," Working Paper 28204, National Bureau of Economic Research.

Galí, J. (2008): *Monetary policy, Inflation, and the Business Cycle: an Introduction to the New Keynesian Framework and its Applications*. Princeton University Press.

Gennaioli, N., Y. Ma, and A. Shleifer (2016): "Expectations and investment," *NBER Macroeconomics Annual*, 30(1), 379–431.

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.

Goetzmann, W. N., D. Kim, and R. J. Shiller (2022): "Crash Narratives," Working Paper 30195, National Bureau of Economic Research.

Handley, K., and J. F. Li (2020): "Measuring the effects of firm uncertainty on economic activity: New evidence from one million documents," Working Paper 27896, National Bureau of Economic Research.

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.

Hassan, T. A., J. Schreger, M. Schwedeler, and A. Tahoun (2021): "Sources and transmission of country risk," Discussion paper, National Bureau of Economic Research.

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.

Liberti, J. M., and M. A. Petersen (2019): "Information: Hard and soft," *Review of Corporate Finance Studies*, 8(1), 1–41.

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.

Macaulay, A., and W. Song (2022): "Narrative-Driven Fluctuations in Sentiment: Evidence Linking Traditional and Social Media," Working paper, University of Oxford Department of Economics.

Maenhout, P. J., A. Vedolin, and H. Xing (2021): "Robustness and dynamic sentiment," Working Paper 3798445, SSRN.

Maitlis, S. (2005): "The social processes of organizational sensemaking," *Academy of management journal*, 48(1), 21–49.

Manski, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, 60(3), 531–542.

Marcet, 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," 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, Northwestern University.

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.

Peterman, W. B. (2016): "Reconciling micro and macro estimates of the Frisch labor supply elasticity," *Economic inquiry*, 54(1), 100–120.

Ramey, V. A. (2011): "Can government purchases stimulate the economy?," *Journal of Economic Literature*, 49(3), 673–85.

Schaal, E., and M. Taschereau-Dumouchel (2020): "Herding Through Booms and Busts," Working paper, CREI.

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, Available at: https://www.sec.gov/files/reada10k.pdf.

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.

Song, W., and S. Stern (2021): "Firm inattention and the transmission of monetary policy: A text-based approach," Working Paper 3900641, SSRN.

Spiegler, R. (2016): "Bayesian networks and boundedly rational expectations," *The Quarterly Journal of Economics*, 131(3), 1243–1290.

Tversky, A., and D. Kahneman (1973): "Availability: A heuristic for judging frequency and probability," *Cognitive psychology*, 5(2), 207–232.

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*.

Werning, I. (2017): "Comment on: Is the Macroeconomy Locally Unstable and Why Should We Care?," in *NBER Macroeconomics Annual 2016*, vol. 31, pp. 540–552. University of Chicago Press.

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 3 and 5

We first provide two assumptions under which Equation 3 holds as a linear approximation. In what follows, we impose the technical requirements that $\mathcal{X}$ is convex, compact subset of $\mathbb{R}$, $\mathcal{Y}$ 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. We moreover assume that all narratives are equally sensitive to information.

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}. Moreover, conditional on the signal, the conditional expectation of θ_t under each narrative k is given by $\mathbb{E}_k[\theta_t|s_{it}] = \alpha s_{it} + c_k + O(||s_{it}||^2)$.*

A sufficient condition for this assumption is that all fundamentals and signals are Gaussian and the signal-to-noise ratio of the signal is constant across all narratives. 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 6. *Under Assumptions 1 and 2, we have that:*

$$x_{it} = \gamma_i + \chi_t + \sum_{k\in\mathcal{K}} \delta_k \lambda_{k,it} + \varepsilon_{it} + O(||(x_{it}, Y_t, \theta_t, Q_t, \omega_i, \nu_{it}, \lambda_{it})||^2) \tag{58}$$

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:

$$\mathbb{E}[x_{it}|i, t, \lambda_{it}] = \gamma_i + \chi_t + \sum_{k\in\mathcal{K}} \delta_k \lambda_{k,it} \tag{59}$$

Proof. By Assumption 1, from the agents' problems (Equation 2), their best replies must solve the following first-order condition (where we suppress all individual and time subscripts):

$$\mathbb{E}_{\pi_\lambda}\left[u_x(x, \hat{Y}(\theta), \theta, \omega)|s\right] = 0 \tag{60}$$

We linearize this first-order conditions in (x, Y, θ, ω) around values $(\bar{x}, \bar{Y}, \bar{\theta}, \bar{\omega})$ which satisfy $\mathbb{E}_{\pi_\lambda}\left[u_x(\bar{x}, \bar{Y}(\theta), \bar{\theta}, \bar{\omega})|s\right] = 0$. This gives

$$\mathbb{E}_{\pi_\lambda}\left[u_{xx}(x-\bar{x}) + u'_{xY}(Y-\bar{Y}) + u'_{x\theta}(\theta-\bar{\theta}) + u'_{x\omega}(\omega-\bar{\omega})|s\right] + R = 0 \tag{61}$$

where $(u_{xx}, u_{xY}, u_{x,\theta}, u_{x\omega})$ are constants equal to the corresponding derivatives evaluated at $(\bar{x}, \bar{Y}, \bar{\theta}, \bar{\omega})$, the remainder R is $O(||(x, Y, \theta, \omega, \nu, \lambda)||^2)$. We can rearrange the above, and use the fact that ω is known to the agent, to write:

$$x = \bar{x} + \frac{1}{|u_{xx}|}u'_{x\omega}(\omega-\bar{\omega}) + \frac{1}{|u_{xx}|}\mathbb{E}_{\pi_\lambda}\left[u'_{xY}(Y-\bar{Y}) + u'_{x\theta}(\theta-\bar{\theta})|s\right] + \frac{1}{|u_{xx}|}R \tag{62}$$

Moreover, we know that $\hat{Y} = \hat{Y}(Q, \theta)$. Thus, assuming that $\hat{Y}$ is continuously differentiable, we may linearize $Y = \bar{Y} + Y'_Q(Q-\bar{Q}) + Y'_\theta(\theta-\bar{\theta}) + \hat{R}$, where $\bar{Y} = \hat{Y}(\bar{Q}, \bar{\theta})$ and $\hat{R}$ is the error induced by the approximation of $\hat{Y}$, which is $O(||Y, Q, \theta||^2)$. Substituting this approximation into Equation 62 gives

$$\begin{aligned} x &= \bar{x} + \frac{1}{|u_{xx}|}u'_{x\omega}(\omega-\bar{\omega}) + \frac{1}{|u_{xx}|}\mathbb{E}_{\pi_\lambda}\left[u'_{xY}(Y'_Q(Q-\bar{Q}) + Y'_\theta(\theta-\bar{\theta})) + u'_{x\theta}(\theta-\bar{\theta})|s\right] + \tilde{R} \\ &= \gamma + \tilde{\chi} + \mathbb{E}_{\pi_\lambda}\left[\tilde{\theta}|s\right] + \tilde{R} \end{aligned} \tag{63}$$

where $\gamma = \bar{x} + \frac{1}{|u_{xx}|}u'_{x\omega}(\omega-\bar{\omega}) - \frac{1}{|u_{xx}|}u'_{xY}(Y'_Q\bar{Q} + Y'_\theta\bar{\theta}) - \frac{1}{|u_{xx}|}u'_{x\theta}\bar{\theta}$, $\tilde{\chi} = \frac{1}{|u_{xx}|}u'_{xY}Y'_QQ$, and $\tilde{\theta} = (u'_{xY}Y'_\theta + u'_{x\theta})\,\theta$, $\tilde{R} = \frac{1}{|u_{xx}|}R + \hat{R}$.

We next re-write the conditional expectation of $\tilde{\theta}$ as linear in two arguments, the signal s and prior mean $\mathbb{E}_{\pi_\lambda}[\tilde{\theta}]$. Using first the linearity of the expectation operator and the linearity of forming beliefs from narratives, and second Assumption 2 to re-write each narrative-specific conditional expectation in terms of the signal and the prior, we write

$$\mathbb{E}_{\pi_\lambda}\left[\tilde{\theta}|s\right] = \sum_{k\in\mathcal{K}}\lambda_k\mathbb{E}_k[\tilde{\theta}|s] = \alpha\tilde{s} + E_\pi\mathbb{E}_{\pi_\lambda}[\tilde{\theta}] + \check{R} \tag{64}$$

where E_π is a constant, $\mathbb{E}_{\pi_\lambda}[\theta] = \sum_{k\in\mathcal{K}}\lambda_k\mathbb{E}_k[\theta]$ is the average prior mean across narratives, $\check{R}$ is the error induced by the approximation and is $O(||(\theta, \nu, \lambda)||^2)$, and the transformed signal is $\tilde{s} = (u'_{xY} Y'_\theta + u'_{x\theta})\, s = \tilde{\theta} + \tilde{\nu}$, with $\tilde{\nu} = (u'_{xY} Y'_\theta + u'_{x\theta})\, \nu$ independent of $\tilde{\theta}$ and of mean zero by Assumption 2. Defining $\chi = \tilde{\chi} + \kappa\tilde{\theta}$, $\varepsilon = E_s \hat{\nu}$ and $\delta_k = E_\pi \mathbb{E}_k[\tilde{\theta}]$, we may write:

$$x = \gamma + \chi + \sum_{k \in \mathcal{K}} \delta_k \lambda_k + \varepsilon + \bar{R} \tag{65}$$

where $\bar{R} = \tilde{R} + \check{R} = O(||(x, Y, \theta, Q, \omega, \nu, \lambda)||^2)$. 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:

$$x_{it} = \gamma_i + \chi_t + \sum_{k \in \mathcal{K}} \delta_k \lambda_{k,it} + \varepsilon_{it} + O(||(x_{it}, Y_t, \theta_t, Q_t, \omega_i, \nu_{it}, \lambda_{it})||^2) \tag{66}$$

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. □

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 7. *Under Assumption 3, we have that:*

$$\mathbb{P}[\lambda_{it} = \lambda | \lambda_{i,t-1}, Y_{t-1}, Q_{t-1}] = \zeta_\lambda + u'_\lambda \lambda_{i,t-1} + r'_\lambda Y_{t-1} + s'_\lambda Q_{t-1} + O(||(\lambda_{i,t-1}, Y_{t-1}, Q_{t-1})||^2) \tag{67}$$

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

$$\mathbb{P}[\lambda_{it} = \lambda | \lambda_{i,t-1} = \lambda', Y_{t-1}, Q_{t-1}] = \zeta_\lambda + u_{\lambda,\lambda'} + r'_\lambda Y_{t-1} + s'_\lambda Q_{t-1} + O(||(\lambda_{i,t-1}, Y_{t-1}, Q_{t-1})||^2) \tag{68}$$

Completing the proof. □

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 $f : [0, 1] \to \mathbb{R}$:

$$\log Y(\theta, Q) = a_0 + a_1 \log \theta + f(Q) \tag{69}$$

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

$$\log \theta_t | s_{it}, \lambda_{it} \sim N\left(\kappa s_{it} + (1-\kappa)(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P), \sigma^2_{\theta|s}\right) \tag{70}$$

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

$$\kappa = \frac{1}{1+\frac{\sigma_\varepsilon^2}{\sigma_\theta^2}} \quad \text{and} \quad \sigma^2_{\theta|s} = \frac{1}{\frac{1}{\sigma_\theta^2}+\frac{1}{\sigma_\epsilon^2}} \tag{71}$$

with κ being the signal-to-noise ratio and $\sigma^2_{\theta|s}$ the variance of fundamentals conditional on the signal. Thus, idiosyncratic productivity has conditional distribution given by:

$$\log \theta_{it} | s_{it}, \lambda_{it} \sim N\left(\log \gamma_i + \kappa s_{it} + (1-\kappa)(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P), \sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) \tag{72}$$

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:

$$\begin{aligned} \log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right] &= \log \mathbb{E}_{it}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right] \\ &= -\frac{1+\psi}{\alpha}\mu_{it} + \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \eta^2 \end{aligned} \tag{73}$$

Under our conjecture (Equation 69), we can moreover compute:

$$\begin{aligned} \log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right] &= \log \mathbb{E}_{it}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)(a_0 + a_1 \log\theta_t + f(Q_t)\right\}\right] \\ &= \left(\frac{1}{\epsilon}-\gamma\right)[a_0 + a_1(\mu_{it} - \log\gamma_i) + f(Q_t)] + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2 [\eta^2 - \sigma^2_{\tilde{\theta}}] \end{aligned} \tag{74}$$

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

$$\begin{aligned} \log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\mu_{it} - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\eta^2 \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)[a_0 + a_1(\mu_{it}-\log\gamma_i) + f(Q_t)] + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2[\eta^2 - \sigma^2_{\tilde{\theta}}] \Bigg] \end{aligned} \tag{75}$$

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:

$$\begin{aligned}
\log Y_t &= \log\left[\left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}\right] \\
&= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}\right] \\
&= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right]
\end{aligned} \tag{76}$$

Moreover, expanding the terms in Equation 75, we have that:

$$\begin{aligned}
\log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\
&+\frac{1+\psi}{\alpha}\left[\log\gamma_i+\kappa s_{it}+(1-\kappa)\left[\lambda_{it}\mu_O+(1-\lambda_{it})\mu_P\right]\right] \\
&-\frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}}\right) \\
&+\left(\frac{1}{\epsilon}-\gamma\right)\left[a_0+a_1\left(\kappa s_{it}+(1-\kappa)\left[\lambda_{it}\mu_O+(1-\lambda_{it})\mu_P\right]\right)+f(Q_t)\right] \\
&+\frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}\Bigg]
\end{aligned} \tag{77}$$

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:

$$\begin{aligned}
\delta_t(\lambda_{it}) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\
&+\frac{1+\psi}{\alpha}\left[\mu_\gamma+\kappa\log\theta_t+(1-\kappa)\left[\lambda_{it}\mu_O+(1-\lambda_{it})\mu_P\right]\right] \\
&-\frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}}\right) \\
&+\left(\frac{1}{\epsilon}-\gamma\right)\left[a_0+a_1\left(\kappa\log\theta_t+(1-\kappa)\left[\lambda_{it}\mu_O+(1-\lambda_{it})\mu_P\right]\right)+f(Q_t)\right] \\
&+\frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}\Bigg]
\end{aligned} \tag{78}$$

and:

$$\hat{\sigma}^2 = \left(\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\right)^2 \left[\left(\frac{1+\psi}{\alpha}\right)^2 \sigma_\gamma^2 + \kappa^2 \left[\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon}-\gamma\right)\right]^2 \sigma_\varepsilon^2\right] \tag{79}$$

Thus, we have that:

$$\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right] = \exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(\lambda_{it}) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} \tag{80}$$

and so:

$$\begin{aligned}\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right] &= Q_t\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(1) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} \\ &+ (1-Q_t)\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(0) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} \\ &= \left[Q_t\exp\left\{\frac{\epsilon-1}{\epsilon}(\delta_t(1)-\delta_t(0))\right\} + (1-Q_t)\right]\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_t(0) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}\end{aligned} \tag{81}$$

Yielding:

$$\log Y_t = \delta_t(0) + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 + \frac{\epsilon}{\epsilon-1}\log\left(Q_t\exp\left\{\frac{\epsilon-1}{\epsilon}(\delta_t(1)-\delta_t(0))\right\} + (1-Q_t)\right) \tag{82}$$

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

$$\delta_t(1) - \delta_t(0) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\left(\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon}-\gamma\right)\right)(1-\kappa)(\mu_O-\mu_P) = \alpha\delta^{OP} \tag{83}$$

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

$$\begin{aligned}\delta_t(0) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}(\mu_\gamma + (1-\kappa)\mu_P) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2(\sigma_{\theta|s}^2 + \sigma_{\tilde{\theta}}^2) \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)(a_0 + a_1(1-\kappa)\mu_P) + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma_{\theta|s}^2 \\ &+ \left[\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon}-\gamma\right)\right]\kappa\log\theta_t + \left(\frac{1}{\epsilon}-\gamma\right)f(Q_t)\Bigg]\end{aligned} \tag{84}$$

By matching coefficients between Equations 82 and Equation 69, we obtain a_0, a_1, and f.

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

$$a_1 = \frac{\left[\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon} - \gamma\right)\right]\kappa}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \tag{85}$$

Under our maintained assumption that $\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \in [0,1)$, as $\kappa \in [0,1]$, we have that this has a unique solution:

$$a_1 = \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \tag{86}$$

It is moreover positive.

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

$$f(Q) = \frac{\frac{1}{\epsilon} - \gamma}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} f(Q) + \frac{\epsilon}{\epsilon - 1}\log\left(1 + Q\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\right\} - 1\right]\right) \tag{87}$$

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

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(1 + Q\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\right\} - 1\right]\right) \tag{88}$$

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:

$$\begin{aligned} a_0 = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}&\left[\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left(\mu_\gamma + (1-\kappa)\mu_P\right) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right)\right. \\ &\left. + \left(\frac{1}{\epsilon} - \gamma\right)\left(a_0 + a_1(1-\kappa)\mu_P\right) + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon} - \gamma\right)^2\sigma^2_{\theta|s}\right] + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 \end{aligned} \tag{89}$$

Solving this equation yields:

$$\begin{aligned} a_0 = \frac{1}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}&\left[\frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}\left[\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left(\mu_\gamma + (1-\kappa)\mu_P\right) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right)\right.\right. \\ &\left.\left. + \left(\frac{1}{\epsilon} - \gamma\right)a_1(1-\kappa)\mu_P + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon} - \gamma\right)^2\sigma^2_{\theta|s}\right] + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2\right] \end{aligned} \tag{90}$$

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 77 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 38. 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^∞-norm and consider the operator $V_M : \mathcal{C} \to \mathcal{C}$ given by:

$$\begin{aligned} V_M(g)(\theta, Q) = \frac{\epsilon}{\epsilon - 1} \log \mathbb{E}_{(\theta,Q)} \Bigg[\exp \Bigg\{ \frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \Bigg(\log \left(\frac{1 - \frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}} \right) \\ - \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ -\frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] + \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \Bigg) \Bigg\} \Bigg] \end{aligned} \tag{91}$$

The following two conditions are sufficient for this operator to be a contraction: (i) monotonicity: for all $g, h \in \mathcal{C}$ such that $g \geq h$, we have that $V_M(g) \geq V_M(h)$ (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^∞-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:

$$\log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \geq \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) h \right\} \right] \tag{92}$$

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

$$\log \mathbb{E}_{(s,Q)}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)(g+a)\right\}\right] = \log \mathbb{E}_{(s,Q)}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)g\right\}\right] + \left(\frac{1}{\epsilon}-\gamma\right)a \quad (93)$$

And so:

$$\begin{aligned} V_M(g+a)(\theta,Q) &= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_{(\theta,Q)}\Bigg[\exp\Bigg\{\frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg(\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\ &- \log \mathbb{E}_{(s,Q)}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right] + \log \mathbb{E}_{(s,Q)}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)g\right\}\right] + \left(\frac{1}{\epsilon}-\gamma\right)a\Bigg)\Bigg\}\Bigg] \\ &= V_M(g)(\theta,Q) + \omega a \end{aligned} \quad (94)$$

where $\omega \in [0,1)$ by assumption. Note that the modulus of contraction ω is precisely the claimed strategic complementarity parameter in Equation 24. This verifies equilibrium uniqueness. □

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 $p \in (1,\infty)$ 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 $\varepsilon > 0$. Let the unique quasi-linear equilibrium we have guessed and verified be $\log Y^*$. We say that g is a ε-equilibrium if

$$||g - V_M(g)||_p < \varepsilon \quad (95)$$

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 Y^* is a ε-equilibrium for bounded fundamentals for any $\varepsilon > 0$ for some bound M:

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

Proof. Now extend from $\mathcal{C}$, $V_M : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ as in Equation 91. We observe that V_M is continuous in the limit in M in the sense that $V_M(g) \to V(g)$ as $M \to \infty$ for all $g \in L^p(\mathbb{R})$. This observation follows from noting that both $\log \mathbb{E}_{(s,Q)}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right]$ and $\log \mathbb{E}_{(s,Q)}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)g\right\}\right]$ are convergent pointwise for $M \to \infty$ for all (s,Q). In Proposition 1, we showed that $V(\log Y^*) = \log Y^*$. Thus, we have that: $V_M(\log Y^*) \to$

$V(\log Y^*) = \log Y^*$, which implies that:

$$\lim_{M\to\infty} || \log Y^* - V_M(\log Y^*)||_p = 0 \tag{96}$$

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

$$|| \log Y^* - V_M(\log Y^*)||_p < \varepsilon \quad \forall M \in \mathbb{N} : M > \bar{M} \tag{97}$$

Completing the proof. □

A.4 Proof of Corollary 1

Proof. From Equation 77, we may express:

$$\begin{aligned} \log x_{it} = \text{cons} + b_3 f(Q_t) + \frac{1+\psi}{\alpha}(\log\gamma_i + \kappa s_{it}) + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \kappa s_{it} \\ \left[\frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\right](1-\kappa)(\mu_O - \mu_P)\lambda_{it} \end{aligned} \tag{98}$$

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

$$\begin{aligned} \log L_{it} = & -\frac{1}{\alpha}\log\theta_{it} + \text{cons} + c_4 f(Q_t) + \text{cons}_i \\ & + \frac{1}{\alpha}\left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\right]\kappa\log\theta_t \\ & + \frac{1}{\alpha}\left[\frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\right](1-\kappa)(\mu_O - \mu_P)\lambda_{it} \\ & + \xi'_{it} \end{aligned} \tag{99}$$

where $\xi'_{it} \sim N(0, \sigma^2_\xi)$ and IID. Comparing the above with the definition of $\alpha\delta^{OP}$ in Equation 83, we see that the coefficient on $\log\theta_t$ in the above expression is δ^{OP}. We finally observe from Equation 88 that $f(Q)$ depends on $(\epsilon, \gamma, \psi, \alpha)$ and δ^{OP}. Hence, given $(\epsilon, \gamma, \psi, \alpha)$, f is identified uniquely from the studied regression estimate. □

A.5 Proof of Corollary 2

Proof. This is immediate by substituting Equation 159 into Equation 29. □

A.6 Proof of Proposition 2

Proof. We prove the three claims in sequence.

(1) The map $T_\theta : [0,1] \to [0,1]$ is continuous for all $\theta \in \Theta$ as f, P_O and P_P are continuous functions. Moreover, it maps a compact set to a compact set. Thus, by Brouwer's fixed point theorem, there exists a Q^*_θ such that $Q^*_\theta = T_\theta(Q^*_\theta)$ for all $\theta \in \Theta$.

(2) To characterize the existence of extremal steady states, observe that $Q = 1$ is a steady state for θ if and only if $T_\theta(1) = P_O(a_o + a_1 \log\theta + f(1), 1) = 1$ and $Q = 0$ is a steady state for θ if and only if $T_\theta(0) = P_P(a_0 + a_1 \log\theta, 0) = 0$. Thus, $Q = 1$ is a steady state if and only if $P_O^{-1}(1;1) \leq a_0 + a_1 \log\theta + f(1)$ and $Q = 0$ is a steady state if and only if $P_P^{-1}(0;0) \geq a_0 + a_1 \log\theta$. To obtain the result as stated, we re-arrange these inequalities in terms of $\log\theta$ and exponentiate.

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

$$\begin{aligned} T'_\theta(Q) &= P_O(a_0 + a_1 \log\theta + f(Q), Q) - P_P(a_0 + a_1 \log\theta + f(Q), Q) \\ &\quad + Q \frac{\mathrm{d}}{\mathrm{d}Q} P_O(a_0 + a_1 \log\theta + f(Q), Q) + (1-Q) \frac{\mathrm{d}}{\mathrm{d}Q} P_P(a_0 + a_1 \log\theta + f(Q), Q) \end{aligned} \tag{100}$$

Thus, for $\theta < \theta_P$ and $Q = 1$:

$$\begin{aligned} T'_\theta(0) &= P_O(a_0 + a_1 \log\theta, 0) - P_P(a_0 + a_1 \log\theta, 0) \\ &\quad + \frac{\mathrm{d}}{\mathrm{d}Q} P_P(a_0 + a_1 \log\theta + f(Q), Q) \mid_{Q=0} \\ &= P_O(a_0 + a_1 \log\theta, 0) \end{aligned} \tag{101}$$

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 $T'_\theta(0) < 1$ when $P_O(a_0 + a_1 \log\theta, 0) < 1$. Moreover, for $\theta < \theta_P$, we have that: $P_O(a_0 + a_1 \log\theta, 0) \leq P_O(a_0 + a_1 \log\theta_P, 0) = P_O(P_P^{-1}(0;0), 0)$. Thus, a sufficient condition for $T'_\theta(0) < 1$ for $\theta < \theta_P$ is that $P_O(P_P^{-1}(0;0), 0) < 1$.

For $\theta > \theta_O$ and $Q = 0$, we have that:

$$
\begin{aligned}
T'_\theta(1) &= P_O(a_0 + a_1 \log\theta + f(1), 1) - P_P(a_0 + a_1 \log\theta + f(1), 1) \\
&\quad + \frac{\mathrm{d}}{\mathrm{d}Q} P_O(a_0 + a_1 \log\theta + f(1), 1) \mid_{Q=1} \\
&= P_P(a_0 + a_1 \log\theta + f(1), 1)
\end{aligned}
\tag{102}
$$

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 $P_P(a_0 + a_1 \log\theta + f(1), 1) \geq P_P(a_0 + a_1 \log\theta_O + f(1), 1) = P_P(P_O^{-1}(1, 1), 1)$. Thus, a sufficient condition for $T'_\theta(1) < 1$ for $\theta > \theta_O$ is that $P_P(P_O^{-1}(1, 1), 1) > 0$. □

A.7 Proof of Corollary 3

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

A.8 Proof of Lemma 1

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

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

$$
\mathcal{B}_\theta(Q^*) = \left\{ Q_0 \in [0, 1] : \lim_{n\to\infty} T^n_\theta(Q_0) = Q^* \right\} \tag{103}
$$

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

$$
T^{n+1}_\theta(Q_0) = T_\theta \circ T^n_\theta(Q_0) \geq T_\theta \circ T^{n-1}_\theta(Q_0) = T^n_\theta(Q_0) \tag{104}
$$

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

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

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

$$T_\theta^{n+1}(Q_0) = T_\theta \circ T_\theta^n(Q_0) \leq T_\theta \circ T_\theta^{n-1}(Q_0) = T_\theta^n(Q_0) \tag{105}$$

by monotonicity of T_θ, which proves the inductive hypothesis. Observe moreover that the sequence $\{T_\theta^n(Q_0)\}_{n\in\mathbb{N}}$ is bounded as $T_\theta^n(Q_0) \in [0,1]$ for all $n \in \mathbb{N}$. Hence, by the monotone convergence theorem, $\lim_{n\to\infty} T_\theta^n(Q_0)$ exists. Finally, toward a contradiction, suppose that $Q_0^\infty = \lim_{n\to\infty} T_\theta^n(Q_0) > 0$. By SSC-B of T we have that $T_\theta(Q_0^\infty) < Q_0^\infty$, but this contradicts that $Q_0^\infty > 0$. Thus, we have that $Q_0^\infty = 0$. Hence, $[0, \hat{Q}_\theta) \subseteq \mathcal{B}_\theta(0)$. Second, consider $Q_0 = \hat{Q}$. We have that $T_\theta(\hat{Q}_\theta) = \hat{Q}$. Thus, $Q_0^\infty = \hat{Q}$. Hence $\hat{Q}_\theta \in \mathcal{B}_\theta(\hat{Q}_\theta)$. Third, consider $Q_0 \in (\hat{Q}_\theta, 1]$. By the exact arguments of the first part, we have that $(\hat{Q}_\theta, 1] \subseteq \mathcal{B}_\theta(1)$. Observing $\mathcal{B}_\theta(0)$, $\mathcal{B}_\theta(\hat{Q}_\theta)$, and $\mathcal{B}_\theta(1)$ 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 48, we obtain the formula for the output IRF. For the fraction of optimists, we see that:

$$\begin{aligned} Q_2 &= \hat{Q}_1 P_O(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1) + (1 - \hat{Q}_1) P_P(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1) \\ &> \hat{Q}_1 P_O(a_0 + f(\hat{Q}_1), \hat{Q}_1) + (1 - \hat{Q}_1) P_P(a_0 + f(\hat{Q}_1), \hat{Q}_1) = \hat{Q}_1 \end{aligned} \tag{106}$$

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 $\log Y_1 \leq \log Y_2$. This corresponds to

$$\log Y_1 = a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1) \leq a_0 + f(\hat{Q}_2) = \log Y_2 \tag{107}$$

which rearranges to the desired expression. □

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 $P_P^{-1}(0; 0) \geq a_0 + a_1 \log \hat{\theta}$ 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 $[0, \hat{Q}_1)$, $(\hat{Q}_1, 1]$, $\{\hat{Q}_1\}$. 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 $t \in \mathbb{N}$. If $Q_2 > \hat{Q}_1$, we have monotone convergence of Q_t to 1. Moreover, the threshold for $\hat{\theta}$ such that $Q_2 = \hat{Q}^*$ 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. □

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 $\{\theta_t\}_{t \in \mathbb{N}}$ 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 $\tau_{PO} = \min\{t \in \mathbb{N} : Q_\tau = 1, Q_0 = 0\}$. In the model, we have that $Q_{t+1} = T(Q_t, \theta_t)$. Fix a path of fundamentals $\{\theta_t\}_{t \in \mathbb{N}}$ and define the fictitious $\overline{Q}$ process as:

$$\overline{Q}_{t+1} = \mathbb{I}[T(\overline{Q}_t, \theta_t) = 1] \tag{108}$$

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

$$\overline{Q}_1 = \mathbb{I}[T(0, \theta_0) = 1] \leq T(0, \theta_0) = Q_1 \tag{109}$$

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

$$\overline{Q}_t = \mathbb{I}[T(\overline{Q}_{t-1}, \theta_{t-1}) = 1] \leq \mathbb{I}[T(Q_{t-1}, \theta_{t-1}) = 1] \leq T(Q_{t-1}, \theta_{t-1}) = Q_t \tag{110}$$

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

As $\overline{Q}_t \leq Q_t$ for all $t \in \mathbb{N}$, we have that:

$$\overline{\tau}_{PO} = \min\{t \in \mathbb{N} : \overline{Q}_\tau = 1, \overline{Q}_0 = 0\} \geq \min\{t \in \mathbb{N} : Q_\tau = 1, Q_0 = 0\} = \tau_{PO} \tag{111}$$

Else, we would have at $\overline{\tau}_{PO}$ that $Q_{\overline{\tau}_{PO}} < \overline{Q}_{\overline{\tau}_{PO}}$, 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 $\{\overline{Q}_t\}_{t\in\mathbb{N}}$ until stopping are given by the set:

$$\mathcal{G}_{PO} = \{(0^{(n-1)}, 1)\} : n \geq 1\} \tag{112}$$

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

$$\begin{aligned}
\mathbb{P}[Q_{t+1} = 1 | Q_t = 0] &= \mathbb{P}\left[P_P(a_0 + a_1 \log \theta_t, 0) = 1\right] \\
&= \mathbb{P}\left[\theta_t \geq \exp\left\{\frac{P_P^\dagger(1;0) - a_0}{a_1}\right\}\right] \\
&= 1 - H\left(\exp\left\{\frac{P_P^\dagger(1;0) - a_0}{a_1}\right\}\right)
\end{aligned} \tag{113}$$

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

$$\tau_{PO} \prec_{FOSD} \overline{\tau}_{PO} \sim \text{Geo}\left(1 - H\left(\exp\left\{\frac{P_P^\dagger(1;0) - a_0}{a_1}\right\}\right)\right) \tag{114}$$

An immediate corollary is that:

$$T_{PO} = \mathbb{E}[\tau_{PO}] \leq \mathbb{E}[\overline{\tau}_{PO}] = \frac{1}{1 - H\left(\exp\left\{\frac{P_P^\dagger(1;0) - a_0}{a_1}\right\}\right)} \tag{115}$$

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

$$\underline{Q}_{t+1} = \mathbb{I}[T(\underline{Q}_t, \theta_t) \neq 0] \tag{116}$$

with $\underline{Q}_0 = 1$. In this case, by an analogous induction have that $\underline{Q}_t \geq Q_t$ for all $t \in \mathbb{N}$ for all sequences $\{\theta_t\}_{t\in\mathbb{N}}$. And so, we have that if $\underline{Q}_t$ has reached 0 then so too has Q_t. Thus, $T_{OP}^* \geq T_{OP}$. The possible sample paths in this case are:

$$\mathcal{G}_{OP} = \{(1^{(n-1)}, 0)\} : n \geq 1\} \tag{117}$$

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

$$\begin{aligned} \mathbb{P}[Q_{t+1} = 0 | Q_t = 1] &= \mathbb{P}\left[\theta_t \leq \exp\left\{\frac{P_O^\dagger(0;1) - a_0 - f(1)}{a_1}\right\}\right] \\ &= H\left(\exp\left\{\frac{P_O^\dagger(0;1) - a_0 - f(1)}{a_1}\right\}\right) \end{aligned} \tag{118}$$

And so we have:

$$T_{OP} \leq \frac{1}{H\left(\exp\left\{\frac{P_O^\dagger(0;1) - a_0 - f(1)}{a_1}\right\}\right)} \tag{119}$$

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

$$\Theta^* = \left(-\infty, \exp\left\{\frac{P_O^\dagger(0;1) - a_0 - f(1)}{a_1}\right\}\right] \cup \left[\exp\left\{\frac{P_P^\dagger(1;0) - a_0}{a_1}\right\}, \infty\right) \tag{120}$$

and suppose that θ takes values only in this set, where the two sub-intervals are disjoint as $P_O^\dagger(0;1) - P_P^\dagger(1;0) \leq f(1)$. 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, $Q_t = \overline{Q}_t = \underline{Q}_t$ 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 $\mu_A < \exp\left\{\frac{P_O^\dagger(0;1)-a_0-f(1)}{a_1}\right\}$ and $\mu_B > \exp\left\{\frac{P_P^\dagger(1;0)-a_0}{a_1}\right\}$. Taking the limit as $\sigma \to 0$, the support of H converges to being contained within Θ^*. $\square$

Online Appendix

for "The Macroeconomics of Narratives" by Flynn and Sastry

Contents

B Model Extensions **74**

B.1 Comparison to the Bayesian Benchmark 74
B.2 Welfare Implications 76
B.3 Continuous Narratives 80
B.4 Narratives About Idiosyncratic Fundamentals 83
B.5 Multi-Dimensional Narratives and Persistent States 84
B.6 Persistent Idiosyncratic Shocks and Belief Updating 88
B.7 Contrarianism, Endogenous Cycles, and Chaos 91
B.8 Narratives in Games and the Role of Higher-Order Beliefs 96
B.9 Model with Firm Dynamics 100

C Additional Details on Textual Data **102**

C.1 Obtaining and Processing 10-Ks 102
C.2 Obtaining and Processing Conference Call Text 103
C.3 Measuring Positive and Negative Words 103

D Additional Details on Firm Fundamentals Data **105**

D.1 Compustat: Data Selection 105
D.2 Compustat: Calculation of TFP 105

E Additional Empirical Results **109**

E.1 Alternative Empirical Strategy: CEO Change Event Studies 109
E.2 Narrative Optimism, Beliefs, and Hiring 110
E.3 State-Dependent Effects of Sentiment 112
E.4 Measuring Virality via Granular Instrumental Variables 113

F Additional Details on Model Quantification **115**

F.1 Solution of Model With Persistent Fundamentals 115
F.2 Calibration Methodology 117
F.3 Estimating a Demand Multiplier in Our Empirical Setting 118
F.4 Simulation Methodology 120

G Our Analysis and Shiller's *Narrative Economics* **121**

G.1 The Modelling of Narratives 121
G.2 Our Work and Shiller's Seven Propositions 123
G.3 The Perennial Economic Narratives: Our Empirical Findings 126

H Additional Figures and Tables **127**

B Model Extensions

This appendix covers several model extensions. First, we study equilibrium dynamics under a benchmark model of Bayesian model updating and contrast these predictions with those obtained in our main analysis (B.1). Second, we theoretically characterize and quantify the normative implications of narrative fluctuations (B.2). Third, fourth, fifth, and sixth we extend the baseline model to respectively incorporate a continuum of different levels of optimism (B.3), narratives about idiosyncratic fundamentals (B.4), multi-dimensional narratives and persistent fundamentals (B.5), and narrative updating that depends on idiosyncratic fundamentals (B.6). In each case, we characterize equilibrium dynamics and show how our main theoretical insights extend. Seventh, we show how endogenous cycles and chaotic dynamics can obtain when agents are contrarian and implement an empirical test for their presence (B.7). Eighth, 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 à la Morris and Shin (2002) (B.8). Finally, we sketch an extension of our abstract framework to allow for persistent idiosyncratic states and adjustment costs and discuss the implications for our measurement (B.9).

B.1 Comparison to 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 aggregate optimism and applying Bayes' rule under rational expectations. Formally, this corresponds to the following law of motion for Q_t:

$$Q_{t+1} = \int_{[0,1]} \mathbb{P}_i[\mu = \mu_O | \{\log Y_j, Q_j\}_{j=0}^t] \, \mathrm{d}i \tag{121}$$

where $\mathbb{P}_i[\mu = \mu_0 | \emptyset] = \lambda_{i0}$ for some $\lambda_{i0} \in (0,1)$ for all $i \in [0,1]$, and conditional probabilities are computed under rational expectations with knowledge of $\{\lambda_{i0}\}_{i\in[0,1]}$. We define the log-odds ratio of an agent's belief as $\Omega_{it} = \log \frac{\lambda_{it}}{1-\lambda_{it}}$. The following Proposition characterizes the dynamics of agents' subjective models under the Bayesian benchmark:

Proposition 8 (Dynamics under the Bayesian Benchmark)**.** *Each agent's log-odds ratio follows a random walk with drift, or $\Omega_{i,t+1} = \Omega_{it} + a + \xi_t$, where $a = \mathbb{E}_H\left[\frac{(\log\theta_t - \mu_P)^2 - (\log\theta_t - \mu_O)^2}{\sigma^2}\right]$ 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:*

$$\log Y_t = \begin{cases} a_0 + a_1 \log \theta_t & \text{if } a < 0, \\ a_0 + a_1 \log \theta_t + f(1) & \text{if } a > 0. \end{cases} \tag{122}$$

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

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

$$\theta_0 = \frac{\log Y_0 - a_0 - f(Q_0)}{a_1} \tag{123}$$

Thus, we have that $\lambda_{i1} = \mathbb{P}[\mu = \mu_O | \theta_0, \lambda_{i0}]$. Moreover, all agents know that $Q_1 = \int_{[0,1]} \lambda_{i1} \, \mathrm{d}i$. Thus, agents can sequentially identify θ_t by observing only $\{Y_j\}_{j \leq t}$ (and not $\{Q_j\}_{j \leq t}$) by computing:

$$\theta_t = \frac{\log Y_t - a_0 - f(Q_t)}{a_1} \tag{124}$$

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

$$\lambda_{i,t+1} = \mathbb{P}_i[\mu = \mu_O | \{\theta_j\}_{j=1}^t] = \lambda_{i,t+1} = \mathbb{P}_i[\mu = \mu_O | \{Y_j\}_{j=1}^t] \tag{125}$$

By application of Bayes rule, we obtain:

$$\lambda_{i,t+1} = \mathbb{P}[\mu = \mu_O | \theta_t, \lambda_{i,t}] = \frac{f_O(\theta_t)\lambda_{i,t}}{f_O(\theta_t)\lambda_{i,t} + f_P(\theta_t)(1 - \lambda_{i,t})} \tag{126}$$

which implies that:

$$\begin{aligned} \frac{\lambda_{i,t+1}}{1 - \lambda_{i,t+1}} &= \frac{f(\log \theta_t | \mu = \mu_O)}{f(\log \theta_t | \mu = \mu_P)} \frac{\lambda_{i,t}}{1 - \lambda_{i,t}} \\ &= \exp\left\{ \frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{\sigma^2} \right\} \frac{\lambda_{i,t}}{1 - \lambda_{i,t}} \end{aligned} \tag{127}$$

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

$$\begin{aligned} \Omega_{i,t+1} &= \Omega_{i,t} + \frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{\sigma^2} \\ &= \Omega_{it} + a + \xi_t \end{aligned} \tag{128}$$

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

$$\frac{\Omega_{i,t}}{t} = \frac{1}{t}\Omega_{i,0} + \frac{t-1}{t}a + \frac{1}{t}\sum_{i=1}^{t}\xi_i \to^{a.s.} a \tag{129}$$

Hence, almost surely, we have that $Q_t \to 1$ if $a > 0$ and $Q_t \to 0$ 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 from combining this characterization for Q_t with the characterization in our main analysis of equilibrium output conditional on optimism and fundamentals (Proposition 1), which continues to hold in the model of this appendix. $\square$

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. Moreover, the log-odds ratio converges linearly and so the odds ratio in favor of the better fitting model converges exponentially quickly. Thus the Bayesian benchmark model makes a prediction that is at odds with our finding of cyclical dynamics for aggregate optimism (Figure 1), and moreover, in the long run, rules out the features of macroeconomic dynamics that we derive in Section 6 as consequences of the endogenous evolution of narrative optimism.

B.2 Welfare Implications

In this appendix, we derive the normative implications of narratives for the economy.

Theory. 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 9 (Narratives and Welfare). *For any path of aggregate optimism $\mathbf{Q} = \{Q_t\}_{t=0}^{\infty}$, aggregate welfare is given by*

$$\begin{aligned}\mathcal{U}(\mathbf{Q}) = U_C^* \sum_{t=0}^{\infty} \beta^t \exp\{(1-\gamma)f(Q_t)\} \\ - U_L^* \sum_{t=0}^{\infty} \beta^t \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \exp\{(1+\psi)d_3 f(Q_t)\}\end{aligned} \tag{130}$$

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 in a pessimistic steady state *if and only if*

$$\frac{U_C^*}{U_L^*} \times \frac{\exp\{(1-\gamma)f(1)\} - 1}{\exp\{(1+\psi)(d_2 + d_3 f(1))\} - 1} > 1 \tag{131}$$

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

$$\tau^* = \exp\left\{(1-\omega)\left(\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}\right) f(1)\right\} - 1 \tag{132}$$

Proof. We have that welfare for any path of optimism $\mathbf{Q} = \{Q_t\}_{t\in\mathbb{N}}$ is given by:

$$\mathcal{U}(\mathbf{Q}) = \sum_{t=0}^{\infty} \beta^t \left(\mathbb{E}_H\left[\frac{C_t(Q_t,\theta_t)^{1-\gamma}}{1-\gamma}\right] - \mathbb{E}_H\left[\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1+\psi} \mathrm{d}i\right]\right) \tag{133}$$

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:

$$\begin{aligned}
\mathbb{E}_H\left[\frac{C_t^{1-\gamma}(Q_t,\theta_t)}{1-\gamma}\right] &= \mathbb{E}_H\left[\frac{1}{1-\gamma}\exp\{(1-\gamma)\log Y(Q_t,\theta)\}\right] \\
&= \mathbb{E}_H\left[\frac{1}{1-\gamma}\exp\{(1-\gamma)\left(a_0 + a_1\log\theta + f(Q_t)\right)\}\right] \\
&= \frac{1}{1-\gamma}\exp\left\{(1-\gamma)\left(a_0 + a_1\mu_H + f(Q_t)\right) + \frac{1}{2}a_1^2\sigma_H^2\right\} \\
&= \frac{1}{1-\gamma}\exp\left\{(1-\gamma)\left(a_0 + a_1\mu_H\right) + \frac{1}{2}a_1^2\sigma_H^2\right\}\exp\{(1-\gamma)f(Q_t)\} \\
&= U_C^*\exp\{(1-\gamma)f(Q_t)\}
\end{aligned} \tag{134}$$

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

$$L_{it} = d_1\log\theta_t + d_2\lambda_{it} + d_3 f(Q_t) + v_{it} \tag{135}$$

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

$$\begin{aligned}
&\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1+\psi}\mathrm{d}i \\
&= \frac{1}{1+\psi}\left(Q_t\exp\{(1+\psi)d_2\} + (1-Q_t)\right) \\
&\quad \times \exp\left\{(1+\psi)(d_1\log\theta + \mu_v + d_3 f(Q_t)) + \frac{1}{2}(1+\psi)^2\sigma_v^2\right\}
\end{aligned} \tag{136}$$

Hence, the expectation over θ is given by:

$$
\begin{aligned}
&\mathbb{E}_H\left[\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1+\psi} \mathrm{d}i\right] \\
&= \frac{1}{1+\psi}\left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \\
&\quad \times \exp\left\{(1+\psi)d_3 f(Q_t)\right\} \exp\left\{(1+\psi)(d_1\mu_H + \mu_v) + \frac{1}{2}(1+\psi)^2(\sigma_v^2 + d_1^2\sigma_H^2)\right\} \\
&= U_L^*\left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \exp\left\{(1+\psi)d_3 f(Q_t)\right\}
\end{aligned}
\tag{137}
$$

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

$$
\begin{aligned}
\mathcal{U}(\mathbf{Q}) = U_C^* &\sum_{t=0}^{\infty} \beta^t \exp\left\{(1-\gamma) f(Q_t)\right\} \\
&- U_L^* \sum_{t=0}^{\infty} \beta^t \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \exp\left\{(1+\psi)d_3 f(Q_t)\right\}
\end{aligned}
\tag{138}
$$

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 $p_{it}^C = Y_t^{\frac{1}{\varepsilon}} x_{it}^{-\frac{1}{\varepsilon}}$, 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:

$$
\begin{aligned}
\log x_{it} = &\frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) - \log \mathbb{E}_{it}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right] \right. \\
&\left. + \log \mathbb{E}_{it}\left[\exp\left\{\left(\frac{1}{\epsilon} - \gamma\right)\log Y_t\right\}\right] \right) + \Xi(\tau)
\end{aligned}
\tag{139}
$$

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

$$
\log Y(\theta, \tau) = a_0 + a_1 \log\theta + \frac{1}{1-\omega}\Xi(\tau) \tag{140}
$$

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

$$
\log x_{it}(\tau) = \log x_{it}(0) + \frac{1}{1-\omega}\Xi(\tau) \tag{141}
$$

Which implies that:

$$\log L_{it}(\tau) = \log L_{it}(0) + \frac{1}{\alpha}\frac{1}{1-\omega}\Xi(\tau) \tag{142}$$

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

$$\begin{aligned}\mathcal{U}(\tau) = U_C^* \sum_{t=0}^{\infty} \beta^t \exp\left\{(1-\gamma)\frac{1}{1-\omega}\Xi(\tau)\right\} \\ - U_L^* \sum_{t=0}^{\infty} \beta^t \exp\left\{(1+\psi)d_3\frac{1}{1-\omega}\Xi(\tau)\right\}\end{aligned} \tag{143}$$

as $d_3 = \frac{1}{\alpha}$. Hence:

$$\mathcal{U}(1) = \mathcal{U}(\tau^*) \tag{144}$$

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

$$\tau^* = \exp\left\{(1-\omega)\left(\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}\right) f(1)\right\} - 1 \tag{145}$$

Completing the proof. □

This result sheds light on the potential for non-fundamental optimism to increase aggregate welfare. 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.

Quantification. Proposition 9 can be directly applied in our numerical calibration from Section 7 to calculate the welfare effects of narrative optimism without approximation. 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.[37] The second is a counterfactual scenario with permanent extreme optimism, or $Q_t \equiv 1$ for all t. The third is a counterfactual scenario with permanent extreme pessimism, or $Q_t \equiv 0$ 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

[37] Relative 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.

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.

B.3 Continuous Narratives

Our main analysis featured two levels of optimism. However, much of our analysis generalizes to a setting with a continuum of levels of optimism. The model is as in Section 5, but now $\mu \in [\mu_P, \mu_O]$ and the distribution of narratives is given by $Q_t \in \Delta([\mu_P, \mu_O])$. The probabilistic transition between models is now given by a Markov kernel $P : [\mu_P, \mu_O] \times \mathcal{Y} \times \Delta^2([\mu_P, \mu_O]) \to \Delta([\mu_P, \mu_O])$ where $P_{\mu'}(\mu, \log Y, Q)$ is the density of agents who have model μ who switch to μ' when aggregate output is Y and the distribution of narratives is Q.

Characterizing Equilibrium Output. By modifying the guess-and-verify arguments that underlie Proposition 1, we can obtain an almost identical representation of equilibrium aggregate output:

Proposition 10 (Equilibrium Characterization with Continuous Narratives)**.** *There exists a quasi-linear equilibrium:*

$$\log Y(\log \theta_t, Q_t) = a_0 + a_1 \log \theta_t + f(Q_t) \tag{146}$$

Moreover, the density of narratives evolves according to the following difference equation:

$$\mathrm{d}Q_{t+1}(\mu') = \int_{\mu_P}^{\mu_O} P_{\mu'}(\mu, a_0 + a_1 \log \theta_t + f(Q_t), Q_t) \mathrm{d}Q_t(\mu) \tag{147}$$

Proof. By appropriately modifying the steps of the proof of Proposition 1, the result follows. Throughout, simply replace $\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P$ with $\tilde{\mu}_{it} \sim Q_t$ and λ_{it} with $\tilde{\mu}_{it}$ as appropriate. The proof follows as written until the aggregation step. At this point, we instead obtain:

$$\log Y_t = \delta_t(\mu_P) + \frac{1}{2}\frac{\epsilon - 1}{\epsilon}\hat{\sigma}^2 + \frac{\epsilon}{\epsilon - 1} \log \left(\int_{\mu_P}^{\mu_O} \exp\left\{ \frac{\epsilon - 1}{\epsilon}(\delta_t(\tilde{\mu}) - \delta_t(\mu_P)) \right\} \mathrm{d}Q_t(\tilde{\mu}) \right) \tag{148}$$

where $\delta_t(\mu_P) = \delta_t(0)$ and $\delta_t(\tilde{\mu}) - \delta_t(\mu_P) = \alpha \delta^{OP} \frac{\tilde{\mu} - \mu_P}{\mu_O - \mu_P}$. Hence, we have that a_0 and a_1 are as in Proposition 1 and f is instead given by:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \log\left(\int_{\mu_P}^{\mu_O} \exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{\tilde{\mu}-\mu_P}{\mu_O-\mu_P}\right\} \mathrm{d}Q(\tilde{\mu})\right) \tag{149}$$

Completing the proof. □

Importantly, observe that we still obtain a marginal representation in terms of the partial equilibrium effect of going from full pessimism to full optimism on hiring δ^{OP}, as we have empirically estimated.

Equilibrium Dynamics. We have seen that a continuum of models poses no difficulty for the static analysis. The challenge for the dynamic analysis is that the state variable, the evolution of which is fully characterized by Proposition 10, is now infinite-dimensional. This notwithstanding, by use of approximation arguments, we can reduce the dynamics to an essentially identical form to that which we have studied in the main text.

To this end, define the cumulant generating function (CGF) of the cross-sectional distribution of narratives as:

$$K_Q(\tau) = \log\left(\mathbb{E}_Q[\exp\{\tau\tilde{\mu}\}]\right) \tag{150}$$

We therefore have that $\log\left(\mathbb{E}_Q[\exp\{\tau(\tilde{\mu}-z)\}]\right) = K_Q(\tau) - \tau z$. It follows by Equation 149 that:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\omega}\left[K_Q\left(\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{1}{\mu_O-\mu_P}\right) - \frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{\mu_P}{\mu_O-\mu_P}\right] \tag{151}$$

By Maclaurin series expansion, we can express the CGF to first-order as:

$$K_Q(\tau) = \mu_Q\tau + O(\tau^2) \tag{152}$$

We therefore have that:

$$f(Q) = \frac{1}{1-\omega}\alpha\delta^{OP}\frac{\mu_Q-\mu_P}{\mu_O-\mu_P} + O\left(\left(\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{1}{\mu_O-\mu_P}\right)^2\right) \tag{153}$$

We now can express the static, general equilibrium effects in terms of mean of the narrative distribution. With some abuse of notation, we now write $f(\mu_Q) = f(Q)$. Of course, this CGF-based approach would allow one to consider higher-order effects through the variance, skewness, kurtosis, and higher cumulants as desired.

In the next steps, we provide conditions on updating that allow us to express the dynamics solely in terms of the mean of the narrative distribution. To do this, we assume that $P_{\mu'}(\mu, \log Y, Q) = P_{\mu'}(\mu'', \log Y, \mu_Q)$ for all $Q \in \Delta^2([\mu_P, \mu_O])$ and all $\mu, \mu', \mu'' \in [\mu_P, \mu_O]$. This is tantamount to assuming no stubbornness (all agents update the same regardless of the model they start with) and that virality only matters via the mean. Under this assumption, we can write $P_{\mu'}(\log Y(\log \theta, \mu_Q), \mu_Q)$ and express the difference equation as:

$$\begin{aligned} \mathrm{d}Q_{t+1}(\mu') &= \int_{\mu_P}^{\mu_O} P_{\mu'}(a_0 + a_1 \log \theta_t + f(\mu_{Q,t}), \mu_{Q,t}) \mathrm{d}Q_t(\mu) \\ &= P_{\mu'}(a_0 + a_1 \log \theta_t + f(\mu_{Q,t}), \mu_{Q,t}) \end{aligned} \tag{154}$$

It then suffices to take the mean of Q_{t+1} to express the system in terms of the one-dimensional state variable $\mu_{Q,t}$:

$$\mu_{Q,t+1} = T(\mu_{Q,t}, \theta_t) = \int_{\mu_P}^{\mu_O} \mu' P_{\mu'}(a_0 + a_1 \log \theta_t + f(\mu_{Q,t}), \mu_{Q,t}) \mathrm{d}\mu' \tag{155}$$

Which is simply a continuous state analog of the difference equation expressed in Corollary 2 expressed in terms of average beliefs.

Steady State Multiplicity. We now obtain the analogous characterization of extremal steady state multiplicity in this setting, *i.e.,* when it is possible that all agents being maximally pessimistic and all agents being maximally optimistic are simultaneously deterministic steady states. To this end, define the following two inverses:

$$\begin{aligned} \hat{P}^{-1}(x; \mu_Q) &= \sup\{Y : P(Y, Q) = \delta_x\} \\ \check{P}^{-1}(x; \mu_Q) &= \inf\{Y : P(Y, Q) = \delta_x\} \end{aligned} \tag{156}$$

where δ_x denotes the Dirac delta function on x. We define analogous objects to the previous θ_O and θ_P:

$$\theta_O = \exp\left\{\frac{\check{P}^{-1}(\mu_O; \mu_O) - a_0 - f(1)}{a_1}\right\}, \ \theta_P = \exp\left\{\frac{\hat{P}^{-1}(\mu_P; \mu_P) - a_0 - f(1)}{a_1}\right\} \tag{157}$$

The following result establishes that these thresholds characterize extremal multiplicity:

Proposition 11 (Steady State Multiplicity with Continuous States)**.** *Extreme optimism and pessimism are simultaneously deterministic steady states for θ if and only if $\theta \in [\theta_O, \theta_P]$, which is non-empty if and only if*

$$\check{P}^{-1}(\mu_O; \mu_O) - \hat{P}^{-1}(\mu_P; \mu_P) \leq f(1) \tag{158}$$

Proof. This follows exactly the same steps as the proofs of Proposition 2 and Corollary 3, replacing the appropriate inverses defined above. □

Thus, the same conditions that give rise to multiplicity with binary narratives obtain with a continuum of levels of optimism. Indeed, observe that restricting to first-order approximations above was unnecessary. We could have considered an arbitrary order, say k, of approximation of the CGF and obtained a system of difference equations for the first k cumulants. Proposition 11 would still hold as written, as under the extremal steady states, all higher cumulants are identically zero and remain so under the provided condition. Naturally, however, the general dynamics only reduce to those resembling the simple model under the first-order approximation. Nevertheless, we observe that this is a first-order approximation to the exact equilibrium dynamics and not simply an approximation of the dynamics of an approximate equilibrium.

B.4 Narratives About Idiosyncratic Fundamentals

In the main analysis, we assumed that narratives described properties of aggregate fundamentals. In this section, we characterize equilibrium dynamics when narratives describe properties of idiosyncratic fundamentals. Concretely, we now instead suppose that all agents believe that $\log \theta_t \sim N(0, \sigma^2)$, or agree about the distribution of aggregate productivity. Moreover, as in the baseline, all agents believe that others' idiosyncratic productivity follows $\log \tilde{\theta}_{jt} \sim N(0, \sigma^2_{\tilde{\theta}})$ for all $j \neq i$. However, agents disagree about the mean of their own idiosyncratic productivity: optimistic agents believe that $\log \tilde{\theta}_{it} \sim N(\mu_O, \sigma^2_{\tilde{\theta}})$ while pessimistic agents believe that $\log \tilde{\theta}_{it} \sim N(\mu_P, \sigma^2_{\tilde{\theta}})$. The rest of the model is identical.

In this context, dynamics are identical conditional on the static relationship between output and narratives. Moreover, the static relationship between output and narratives is now identical (up to a constant) conditional on estimating the partial equilibrium effect of optimism on hiring. This is formalized by the following result:

Proposition 12 (Equilibrium Characterization with Narratives About Idiosyncratic Fundamentals)**.** *There exists a unique equilibrium such that:*

$$\log Y(\log \theta_t, Q_t) = \tilde{a}_0 + a_1 \log \theta_t + \tilde{f}(Q_t) \tag{159}$$

for coefficients $\tilde{a}_0$ and $a_1 > 0$, and a strictly increasing function f, where a_1 is identical to that from Proposition 1 and

$$\tilde{f}(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \log\left(1 + Q\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\tilde{\delta}^{OP}\right\} - 1\right]\right) \tag{160}$$

where $\tilde{\delta}^{OP}$ is defined in Equation 161.

Proof. The proof follows exactly the steps of the proof of Proposition 1 where the aggregate narrative is replaced with an idiosyncratic one. To be concrete, the computation of $\log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right]$ and the method of aggregation are identical to those in the proof of Proposition 1. The only difference is in the computation of $\log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right]$. Now, Equation 74 differs in that $\mu_{it} = \log \gamma_i + \kappa s_{it}$. Tracking this through to Equation 78, lines 1, 2, 3, and 5 are identical and line 4 differs only in that the term $(1-\kappa)[\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P]$ is now set equal to zero. The analysis then follows up to Equation 83, at which point we have that the exact formula for δ^{OP} changes and is now given by:

$$\alpha\tilde{\delta}^{OP} = \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}(1-\kappa)(\mu_O - \mu_P) \tag{161}$$

The formula for $\delta_t(0)$ is identical except for in the second line where the term $a_1(1-\kappa)\mu_P$ is now equal to zero. The formula for a_1 remains the same. Conditional on $\tilde{\delta}^{OP}$, the formula for f remains the same. The formula for a_0 is identical except for the second line where the term $(1/\epsilon - \gamma)a_1(1-\kappa)\mu_P$ is now equal to zero. □

This Proposition makes clear that output differs in this case only up to an intercept and in changing the mapping from structural parameters to the partial-equilibrium effect of optimism on hiring. Nonetheless, interpreted via the model above, our empirical exercise directly identifies the now-relevant parameter $\tilde{\delta}^{OP}$. As a result, neither our theoretical nor quantitative analysis is sensitive to making narratives be about idiosyncratic conditions. The only difference is that the point calibrations for κ and $(\mu_O - \mu_P)$ would change, while the aggregate dynamics would remain identical.

B.5 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:

$$\log \theta_t = (1-\rho)\mu + \rho \log \theta_{t-1} + \sigma \nu_t \tag{162}$$

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 $\{(\mu_k, \rho_k, \sigma_k)\}_{k \in \mathcal{K}}$. We restrict that agents 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.[38] 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 13 (Equilibrium Characterization with Multi-Dimensional Narratives and Persistence)**.** *There exists a quasi-linear equilibrium:*

$$\log Y(\log\theta_t, \log\theta_{t-1}, Q_t) = a_0 + a_1 \log\theta_t + a_2 \log\theta_{t-1} + f(Q_t, \theta_{t-1}) \tag{163}$$

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

$$Q_{t+1,k} = \sum_{k' \in \mathcal{K}} Q_{t,k'} P_{k'}(k, a_0 + a_1 \log\theta_t + a_2 \log\theta_{t-1} + f(Q_t, \theta_{t-1}), Q_t) \tag{164}$$

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:

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

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 agents' posterior beliefs given narrative weights. Let E denote the standard basis for $\mathbb{R}^K$ with k-th basis vector denoted by

$$e_k = \{\underbrace{0, \dots, 0}_{k-1}, 1, \underbrace{0, \dots, 0}_{K-k}\} \tag{166}$$

We have that $\lambda_{it} = e_k$ for some $k \leq K$. Under this narrative loading, we have that agent's

[38] Formally, this means that the variance of the noise in agents' signals satisfies $\sigma^2_{\varepsilon,k} \propto \sigma^2_k$ across narratives.

posteriors are given by:

$$\log \theta_{it}|\lambda_{it}, s_{it} \sim N\left(\log \gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \tag{167}$$

with:

$$\begin{aligned} \mu(e_k, \theta_{t-1}) &= (1-\rho_k)\mu_k + \rho_k \log \theta_{t-1} \\ \sigma^2_{\theta|s}(e_k) &= \frac{1}{\frac{1}{\sigma^2_k} + \frac{1}{\sigma^2_{\epsilon,k}}} \quad \kappa = \frac{1}{1 + \frac{\sigma^2_{\varepsilon,k}}{\sigma^2_k}} \end{aligned} \tag{168}$$

for all $k \leq K$,where κ does not depend on k as $\sigma^2_{\varepsilon,k} \propto \sigma^2_k$. Hence, we can compute agents' best replies by evaluating:

$$\log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right] = -\frac{1+\psi}{\alpha}\left(\log \gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \tag{169}$$

$$\begin{aligned} \log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right] = &\left(\frac{1}{\epsilon}-\gamma\right)\left(a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1})\right) \\ &+ \frac{1}{2}\left(\frac{1}{\epsilon}-\gamma\right)^2 a_1^2 \sigma^2_{\theta|s}(\lambda_{it}) \end{aligned} \tag{170}$$

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

$$\begin{aligned} \log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left[\log \gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right] \\ &- \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1})\right] \\ &+ \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2 \sigma^2_{\theta|s}(\lambda_{it})\Bigg] \end{aligned} \tag{171}$$

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 79 and:

$$\begin{aligned}\delta_t(e_k) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right)\\&+\frac{1+\psi}{\alpha}\left[\log\gamma_i+\kappa\log\theta_t+(1-\kappa)\mu(e_k,\theta_{t-1})\right]\\&-\frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}(e_k)+\sigma^2_{\tilde{\theta}}\right)\\&+\left(\frac{1}{\epsilon}-\gamma\right)\left[a_0+a_1\left(\kappa\log\theta_t+(1-\kappa)\mu(e_k,\theta_{t-1})\right)+a_2\log\theta_{t-1}+f(Q_t,\theta_{t-1})\right]\\&+\frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}(e_k)\Bigg]\end{aligned} \tag{172}$$

for all $k \leq K$. Aggregating these best replies, using Equation 80, we obtain that:

$$\begin{aligned}\log Y_t &= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right]\\&=\frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(e_k)+\frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}\right)\\&=\delta_t(e_1)+\frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2+\frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\left(\delta_t(e_k)-\delta_t(e_1)\right)\right\}\right)\end{aligned} \tag{173}$$

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 $k \leq K$ 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:

$$f(Q,\theta_{t-1}) = \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}f(Q,\theta_{t-1})+\frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1})\right\}\right) \tag{174}$$

and so:

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

Completing the proof. □

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.6 Persistent Idiosyncratic Shocks and Belief Updating

We now extend the analysis from Section B.5 to the case where agents' idiosyncratic states drive narrative updating and are persistent. Concretely, in that setting, we let $P_{k'}$ depend on $(Y_t, Q_t, \tilde{\theta}_{it})$ and idiosyncratic productivity shocks evolve according to an AR(1) process:

$$\log \tilde{\theta}_{it} = \rho_{\tilde{\theta}} \log \tilde{\theta}_{i,t-1} + \zeta_{it} \tag{176}$$

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

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 $\tilde{\theta}_{it-1}$. 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 $F_{\tilde{\theta}}$, and the conditional distribution of narratives given $\tilde{\theta}$ as $\check{Q}_{t,k|\tilde{\theta}} = \frac{\check{Q}_{t,k}(\tilde{\theta})}{f_{\tilde{\theta}}(\tilde{\theta})}$.

Proposition 14 (Equilibrium Characterization with Multi-Dimensional Narratives, Aggregate and Idiosyncratic Persistence, and Idiosyncratic Narrative Updating)**.** *There exists a quasi-linear equilibrium:*

$$\log Y(\log \theta_t, \log \theta_{t-1}, \check{Q}_t) = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(\check{Q}_t, \theta_{t-1}) \tag{177}$$

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

Proof. This proof follows closely that of Proposition 13. Under narrative loading λ_{it}, we have that the agent's posterior regarding $\log \theta_{it}$ is given by:

$$\log \theta_{it} | \tilde{\theta}_{it-1}, \lambda_{it}, s_{it} \sim N\left(\log \gamma_i + \rho_{\tilde{\theta}} \log \tilde{\theta}_{it-1} + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\xi}\right) \tag{178}$$

where $\mu(\lambda_{it}, \theta_{t-1})$, κ, and $\sigma^2_{\theta|s}(\lambda_{it})$ are as in Proposition 13. Then substitute $\log \gamma_i + \rho_{\tilde{\theta}} \tilde{\theta}_{it-1}$ for $\log \gamma_i$ and follow the Proof of Proposition 13 until the aggregation step (Equation 173).

We now instead have that:

$$\begin{aligned}
\log Y_t &= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\tilde{\theta}_{it-1},\lambda_{it}\right]\right] \\
&= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_t(e_k,\tilde{\theta}_{it-1})+\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\right)^2\hat{\sigma}^2\right\}\right] \\
&= \frac{\epsilon}{\epsilon-1}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(e_k,\tilde{\theta})+\frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right) \\
&= \delta_t(e_1,1)+\frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 \\
&\quad+\frac{\epsilon}{\epsilon-1}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\epsilon}\left(\delta_t(e_k,\tilde{\theta})-\delta_t(e_1,1)\right)\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right)
\end{aligned} \tag{179}$$

Again, $\hat{\sigma}^2$ is a constant and $\delta_t(e_1,0)$ depends linearly on $\log\theta_t$ and $\log\theta_{t-1}$ and $\delta_t(e_k,\tilde{\theta})-\delta_t(e_1,1)$ does not depend on $\log\theta_t$ for all $k\le K$. 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 13, we then have that:

$$f(\check{Q},\theta_{t-1})=\frac{\frac{\epsilon}{\epsilon-1}}{1-\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1},\tilde{\theta})\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right) \tag{180}$$

Completing the proof. □

We can use this result to study the additional effects induced by persistent idiosyncratic fundamentals. To do this, we restrict to the case of our main analysis with optimism and pessimism. In this context, we have that:

$$f(\check{Q})=\frac{\frac{\epsilon}{\epsilon-1}}{1-\omega}\log\left(\mathbb{E}_{\tilde{\theta}}\left[\check{Q}_{t|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\}+(1-\check{Q}_{t|\tilde{\theta}})\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\}\right]\right) \tag{181}$$

where:

$$\begin{aligned}
\delta_{OP}(\tilde{\theta}) &= \alpha\delta_{OP}+\frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}}\log\tilde{\theta} \\
\delta_{PP}(\tilde{\theta}) &= \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}}\log\tilde{\theta}
\end{aligned} \tag{182}$$

We define $\xi = \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}}$ and observe that we can write:

$$\begin{aligned}
&\check{Q}_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\} + (1-\check{Q}_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\} \\
&= Q_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\left(\alpha\delta_{OP} + \xi\log\tilde{\theta}\right)\right\} + (1-Q_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\} \\
&= Q_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\}\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\}-1\right] + \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\}
\end{aligned} \tag{183}$$

Taking the expectation of the relevant terms, we obtain:

$$\begin{aligned}
&\mathbb{E}_{\tilde{\theta}}\left[\check{Q}_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\} + (1-\check{Q}_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\}\right] \\
&= \left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\}-1\right]\exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}Q_t \\
&\quad + \mathrm{Cov}_t\left(Q_{t|\tilde{\theta}}, \tilde{\theta}^{\frac{\epsilon-1}{\epsilon}\xi}\right) + \exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}
\end{aligned} \tag{184}$$

Thus, we have that the contribution of optimism to output is given by:

$$\begin{aligned}
f(\check{Q}_t) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\omega}\log\Bigg(&\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\}-1\right]\exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}Q_t \\
&+ \mathrm{Cov}_t\left(Q_{t|\tilde{\theta}}, \tilde{\theta}^{\frac{\epsilon-1}{\epsilon}\xi}\right) + \exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}\Bigg)
\end{aligned} \tag{185}$$

We observe that the first term is almost identical to that in our main analysis. This term is now intermediated by the effect of heterogeneity in previous productivity (to see this, observe that this vanishes when $\rho_{\tilde{\theta}} = 0$). Second, there is a new effect stemming from the covariance of optimism and productivity. Intuitively, when more optimistic firms are also more productive, they increase their production by more and this increases output. Finally, there is a level effect of heterogeneous productivity.

Thus, the sole new qualitative force is the covariance effect. To the extent that this does not vary with time, it can have no effect on dynamics. We investigate this in the data by estimating the regression model

$$\log\hat{\theta}_{it} = \sum_{\tau=1995}^{2019}\beta_\tau\cdot(\mathrm{opt}_{i\tau}\cdot\mathbb{I}[\tau=t]) + \chi_{j(i),t} + \gamma_i + \varepsilon_{it} \tag{186}$$

where $(\chi_{j(i),t}, \gamma_i)$ are industry-by-time and firm fixed effects, and β_s measures the (within-industry, within-firm) difference in mean log TFP for optimistic and pessimistic firms in each year. If the β_s vary systematically with the business cycle, then the shifting productivity composition of optimists over the business cycle is an important component of business-cycle dynamics.

We plot our coefficient estimates β_τ in Figure A12. The estimates are generally positive, but economically small relative to the large observed variation in TFP, $\log \theta_{it}$, which has an in-sample standard deviation of 0.84. Outside of the first two years and last year of the sample, we find very limited evidence of time variation. Moreover, the variation that exists is not obviously correlated with the business cycle. This suggests that the compositional effect for optimists driven by narrative updating in response to idiosyncratic conditions is not, at least in our data, quantitatively significant.

B.7 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 appendix, we relax this assumption and delineate precise, testable conditions under which cyclical and chaotic dynamics occur. We do so in a model with "contrarian" agents whose updating contradicts recent data and/or consensus. Our analysis of endogenous narratives with contrarianism therefore 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.

We begin by defining cycles and chaos. There exists a cycle of period $k \in \mathbb{N}$ 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 $Q, Q' \in S$ such that $Q \neq Q'$, we have that $\limsup_{t\to\infty} |T^t(Q) - T^t(Q')| > 0$ and $\liminf_{t\to\infty} |T^t(Q) - T^t(Q')| = 0$ and (ii) for every $Q \in S$ and periodic point $Q' \in [0, 1]$, $\limsup_{t\to\infty} |T^t(Q) - T^t(Q')| > 0$. 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.

A Variant Model with the Potential for Cycles and Chaos. We will study the issue of cycles and chaos under the simplifying assumption that,[39] in equilibrium, the induced probabilities that optimists and pessimists respectively become optimists are quadratic and given by:[40]

$$\tilde{P}_O(Q) = a_O + b_O Q - cQ^2\, ,\ \tilde{P}_P(Q) = a_P + b_P Q - cQ^2 \tag{187}$$

with parameters $(a_O, a_P, b_O, b_P, c) \in \mathbb{R}^5$ such that $P_O([0,1]), P_P([0,1]) \subseteq [0,1]$. 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:

Proposition 15. *The following statements are true:*

1. *When $\tilde{P}_O \geq \tilde{P}_P$ 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. The dynamics of optimism are characterized by the transition map

$$\begin{aligned} T(Q) &= Q(a_O + b_O Q - cQ^2) + (1-Q)(a_P + b_P Q - cQ^2) \\ &= a_P + (a_O - a_P + b_P)Q - (c + b_P - b_O)Q^2 \end{aligned} \tag{188}$$

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 $\check{T}(x) = \eta x(1-x)$ with

$$\eta = 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)} \tag{189}$$

Two maps $T : [0,1] \to [0,1]$ and $T' : [0,1] \to [0,1]$ are topologically conjugate if there exists a continuous, invertible function $h : [0,1] \to [0,1]$ such that $T' \circ h = h \circ T$. If T is

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

[40] This 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 $f(Q) \approx \frac{\alpha\delta^{OP}}{1-\omega}Q$. In this case:

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

topologically conjugate to T' and we know the orbit of T', we can compute the orbit of T via the formula:

$$T^k(Q) = \left(h^{-1} \circ T'^k \circ h\right)(Q) \tag{190}$$

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

To show the topological conjugacy of T and $\check{T}$, 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:

$$\hat{T}(\hat{h}(Q)) = (k + \hat{\alpha}^2) + 2\hat{\alpha}\hat{\beta}Q + \hat{\beta}^2Q^2 \tag{191}$$

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

$$\hat{h}^{-1}(\hat{T}(\hat{h}(Q))) = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}} + 2\hat{\alpha}Q + \hat{\beta}Q^2 \tag{192}$$

   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:

$$\omega_0 = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}},\ \omega_1 = 2\hat{\alpha},\ \omega_2 = -\hat{\beta} \tag{193}$$

   We therefore have that:

$$k = \hat{\beta}\omega_0 + \hat{\alpha}(1 - \hat{\alpha}) = -\omega_2\omega_0 + \frac{\omega_1}{2}\left(1 - \frac{\omega_1}{2}\right) \tag{194}$$

   with $\hat{h}(Q) = \frac{\omega_1}{2} - \omega_2 Q$.
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:

$$\check{T}(\check{h}(Q)) = \eta\left(\check{\alpha}(1 - \check{\alpha}) + \check{\beta}(1 - 2\check{\alpha})Q - \check{\beta}^2Q^2\right) \tag{195}$$

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

$$\check{h}^{-1}(\check{T}(\check{h}(Q))) = \frac{\eta\check{\alpha}(1 - \check{\alpha}) - \check{\alpha}}{\check{\beta}} + \eta(1 - 2\check{\alpha})Q - \eta\check{\beta}Q^2 \tag{196}$$

Matching coefficients, we find:

$$k = \frac{\eta\check{\alpha}(1-\check{\alpha}) - \check{\alpha}}{\check{\beta}},\ 0 = \eta(1-2\check{\alpha}),\ 1 = -\eta\check{\beta} \tag{197}$$

We therefore obtain that:

$$k = \eta(\check{\alpha} - \eta(1-\check{\alpha})) = \frac{\eta}{2}\left(1 - \frac{\eta}{2}\right) \tag{198}$$

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

$$T = \hat{h}^{-1} \circ \check{h}^{-1} \circ \check{T} \circ \check{h} \circ \hat{h} \tag{199}$$

with

$$\begin{aligned} \eta &= 1 + \sqrt{1 - 4\left(-\omega_2\omega_0 + \frac{\omega_1}{2}\left(1 - \frac{\omega_1}{2}\right)\right)} = 1 + \sqrt{(\omega_1 - 1)^2 + 4\omega_2\omega_0} \\ &= 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)} \end{aligned} \tag{200}$$

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

Having shown 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. $\tilde{P}_O \geq \tilde{P}_P$ 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 $1 + \sqrt{6}$ (see Weisstein, 2001, for reference). In this case, $c = 0$. This is in addition to the requirements that $\max_{Q\in[0,1]} \tilde{P}_i(Q) \leq 1$ and $\min_{Q\in[0,1]} \tilde{P}_i(Q) \geq 0$ for $i \in \{O, P\}$, which can be expressed as:

$$\begin{aligned} &\max_{Q\in[0,1]} \tilde{P}_i(Q) = \max\left\{a_i, a_i + b_i - c, \left(a_i + \frac{b_i^2}{4c}\right)\mathbb{I}[0 \leq b_i \leq 2c]\right\} \leq 1 \\ &\min_{Q\in[0,1]} \tilde{P}_i(Q) = \min\{a_i, a_i + b_i - c\} \geq 0 \end{aligned} \tag{201}$$

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

$$\begin{aligned} &\max_{(a_O,a_P,b_O,b_P)\in\mathbb{R}^4} (a_O - a_P + b_P - 1)^2 + 4a_P(b_P - b_O) \\ &\text{s.t.} \max\{a_O, a_O + b_O\} \leq 1, \max\{a_P, a_P + b_P\} \leq 1 \\ &\min\{a_O, a_O + b_O\} \geq 0, \min\{a_P, a_P + b_P\} \geq 0 \end{aligned} \tag{202}$$

Exact solution of this program via Mathematica yields that the maximum value is 5. This implies that the maximum value of η is $1 + \sqrt{5} \approx 3.23$, which is greater than 3 but less than $1 + \sqrt{6}$. 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 $a_0 = a_P = 0$, $b_O = b_P = 4$, $c = 4$. Thus, cycles of any period $k \in \mathbb{N}$ and chaotic dynamics can occur (see Weisstein, 2001, for reference).

□

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.

An Empirical Test for Cycles and Chaos. Proposition 15 shows how to translate an updating rule of the form of Equation 187 into predictions about the potential for cycles and chaos. We now estimate this updating rule in the data to test these predictions empirically. Concretely, in our panel dataset of firms, we estimate the regression model

$$\begin{aligned} \text{opt}_{it} = \alpha_1 \, \text{opt}_{i,t-1} + \beta_1 \text{opt}_{i,t-1} \cdot \overline{\text{opt}}_{i,t-1} + \\ \beta_2 (1 - \text{opt}_{i,t-1}) \cdot \overline{\text{opt}}_{i,t-1} + \tau \left(\overline{\text{opt}}_{i,t-1}\right)^2 + \gamma_i + \varepsilon_{it} \end{aligned} \tag{203}$$

where γ_i is a firm fixed effect. This model allows the effects of virality to depend on agents' previous state. In the mapping to Equation 187, $\alpha = a_P$, $\alpha_1 = a_O - a_P$, $\beta_1 = b_O$, $\beta_2 = b_P$, and $\tau = c$. With estimates of each regression parameter, denoted by a hat, we also obtain an estimate of the logistic map parameter η defined in Equation 189:

$$\hat{\eta} = 1 + \sqrt{(\hat{\alpha}_1 + \hat{\beta}_2 - 1)^2 + 4\hat{\alpha}_1(\hat{\tau} + \hat{\beta}_2 - \hat{\beta}_1)} \tag{204}$$

Since $\hat{\eta}$ is a nonlinear function of estimated parameters in the regression, we can conduct inference on $\hat{\eta}$ using the delta method. Moreover, this constitutes a test for the possibility of cycles and chaos in the model by the logic of Proposition 15. Specifically, as described in the proof of that result, there are two main cases. First, if $\eta < 3$, then case 1 of the result obtains: there are neither cycles of any period nor chaotic dynamics. Second, if $\eta \geq 3$, there can be cycles of period 2 or more and/or chaos. Moreover, if $\eta > 3.57$, chaotic dynamic obtain.

Our estimates are presented in Table A19. Our point estimate of η is 1.443 and the 95% confidence interval is $(0.076, 2.810)$. This rules out, at the 5% level, the presence of cycles and/or chaos. The 99% confidence interval is $(-0.354, 3.240)$, which does not rule out cycles. The p-value for the chaotic dynamics threshold is 0.001. Thus, our results provide strong evidence against the possibility of chaos due to viral optimism, and marginally weaker evidence against the possibility of cycles. This test complements the literature on endogenous cycles in macroeconomic models (see, *e.g.*, Boldrin and Woodford, 1990; Beaudry, Galizia, and Portier, 2020) by providing a micro-founded test within a structural economic model, which may ameliorate challenges associated with interpreting pure time-series evidence (see, e.g., Werning, 2017).

B.8 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):

$$x_{it} = \alpha \mathbb{E}_{it}[\theta_t] + \beta \mathbb{E}_{it}[Y_t] \tag{205}$$

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 $Y_t = \int_{[0,1]} x_{it} \mathrm{d}i$. 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:

$$\overline{\mathbb{E}}_t[\theta_t] = \int_{[0,1]} \mathbb{E}_{it}\left[\theta_t\right] \mathrm{d}i \tag{206}$$

and the higher-order average expectations operator for $k \in \mathbb{N}$ as:

$$\overline{\mathbb{E}}_t^k[\theta_t] = \int_{[0,1]} \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^{k-1}[\theta_t]\right] \mathrm{d}i \tag{207}$$

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

$$Y_t = \alpha \sum_{k=1}^{\infty} \beta^{k-1} \overline{\mathbb{E}}_t^k[\theta_t] \tag{208}$$

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:

$$\mathbb{E}_t[\theta_t | s_{it}, \lambda_{it}] = \kappa s_{it} + (1-\kappa)\left(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right) \tag{209}$$

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

$$\overline{\mathbb{E}}_t[\theta_t] = \kappa\theta_t + (1-\kappa)(Q_t\mu_O + (1-Q_t)\mu_P) \tag{210}$$

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:

$$\overline{\mathbb{E}}_t^k[\theta_t] = \kappa^k\theta_t + (1-\kappa^k)(Q_t\mu_O + (1-Q_t)\mu_P) \tag{211}$$

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

$$Y_t = \frac{\alpha}{1-\beta}\left(\frac{(1-\beta)\kappa}{1-\beta\kappa}\theta_t + \frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right)\right) \tag{212}$$

Moreover, agents' actions follow:

$$\begin{aligned} x_{it} = \alpha\frac{1}{1-\beta\kappa}\left[\kappa\theta_t + \kappa\varepsilon_{it} + (1-\kappa)\left(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right)\right] \\ + \beta\frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right) \end{aligned} \tag{213}$$

Proof. To substantiate the arguments in the main text, by aggregating Equation 205, we obtain that:

$$Y_t = \alpha\overline{E}_t[\theta_t] + \beta\overline{E}_t[Y_t] \tag{214}$$

Thus, by recursive substitution k times we obtain that:

$$Y_t = \alpha\sum_{j=1}^{k}\beta^{j-1}\overline{E}_t^j[\theta_t] + \beta^k\overline{E}_t^k[Y_t] \tag{215}$$

Moreover, we have that:

$$\overline{\mathbb{E}}_t^j[\theta_t] = \kappa^j\theta_t + (1-\kappa^j)(Q_t\mu_O + (1-Q_t)\mu_P) \tag{216}$$

and thus that:

$$\begin{aligned} \alpha\sum_{j=1}^{k}\beta^{j-1}\overline{E}_t^j[\theta_t] &= \alpha\sum_{j=1}^{k}\beta^{j-1}\left(\kappa^j\theta_t + (1-\kappa^j)(Q_t\mu_O + (1-Q_t)\mu_P)\right) \\ &= \alpha\sum_{j=1}^{k}\beta^{j-1}(Q_t\mu_O + (1-Q_t)\mu_P) + \alpha\beta^{-1}\sum_{j=1}^{k}(\beta\kappa)^j\left[\theta_t - (Q_t\mu_O + (1-Q_t)\mu_P)\right] \end{aligned} \tag{217}$$

Hence:

$$\begin{aligned} \lim_{k\to\infty}\alpha\sum_{j=1}^{k}\beta^{j-1}\overline{E}_t^j[\theta_t] = \frac{\alpha}{1-\beta}(Q_t\mu_O + (1-Q_t)\mu_P) + \\ \frac{\alpha\kappa}{1-\beta\kappa}\left[\theta_t - (Q_t\mu_O + (1-Q_t)\mu_P)\right] \end{aligned} \tag{218}$$

We therefore have that there is a unique equilibrium if $\lim_{k\to\infty} \beta^k \overline{E}_t^k[Y_t] = 0$. 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 $\lim_{k\to\infty} \beta^k \overline{E}_t^k[Y_t] = c$ for some $c \in \mathbb{R}$ for almost all $\theta \in \Theta$. Hence, the equilibrium is given by:

$$\begin{aligned} Y_t &= \frac{\alpha}{1-\beta}(Q_t\mu_O + (1-Q_t)\mu_P) + \frac{\alpha\kappa}{1-\beta\kappa}\left[\theta_t - (Q_t\mu_O + (1-Q_t)\mu_P)\right] + c \\ &= \frac{\alpha}{1-\beta}\left(\frac{(1-\beta)\kappa}{1-\beta\kappa}\theta_t + \frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right)\right) + c \end{aligned} \tag{219}$$

But then we have that $c = 0$ by computing $\lim_{k\to\infty} \beta^k \overline{E}_t^k[Y_t] = 0$ under this equilibrium.

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

$$\begin{aligned} x_{it} &= \alpha\mathbb{E}_{it}[\theta_t] + \beta\mathbb{E}_{it}[Y_t] \\ &= \alpha\mathbb{E}_{it}[\theta_t] + \beta\mathbb{E}_{it}\left[\frac{\alpha}{1-\beta}\left(\frac{(1-\beta)\kappa}{1-\beta\kappa}\theta_t + \frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right)\right)\right] \\ &= \left(\alpha + \beta\frac{\alpha}{1-\beta}\frac{(1-\beta)\kappa}{1-\beta\kappa}\right)\mathbb{E}_{it}[\theta_t] + \beta\frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right) \\ &= \alpha\frac{1}{1-\beta\kappa}\left(\kappa s_{it} + (1-\kappa)\left(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right)\right) \\ &\quad + \beta\frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right) \end{aligned} \tag{220}$$

Completing the proof. □

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:

$$\frac{\partial Y_t}{\partial Q_t} = \frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}(\mu_O - \mu_P) = \alpha\sum_{j=1}^{\infty}\beta^{j-1}(1-\kappa^j)(\mu_O - \mu_P) \tag{221}$$

The first expression is composed of the relative importance of fundamentals $\frac{\alpha}{1-\beta}$, the impact of prior beliefs on the entire hierarchy of higher-order beliefs about exogenous and endogenous outcomes $\frac{1-\kappa}{1-\beta\kappa}$ and the difference between the two narratives $\mu_O - \mu_P$. The second expression re-expands the heirarchy of beliefs, to highlight how fraction

$$\frac{\beta^{j-1}(1-\kappa^j)}{\frac{1}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}} \tag{222}$$

of the total effect is driven by beliefs of order j. These weights decline more slowly if complementarity β or prior weights $1-\kappa$ are high.

Finally, our result shows how the regression equation relating individual actions with narrative weights, estimated in our main analysis, holds in equilibrium in the linearized beauty contest. Thus, our empirical strategy is compatible with the interpretation that the macroeconomy is best described by a linear beauty contest, and moreover can be ported to other settings where this modeling assumption may be appropriate, such as that of financial speculation (see *e.g.,* Allen, Morris, and Shin, 2006).

B.9 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 idiosyncratic 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 $x_{it} \in \mathcal{X}$. Their objective function still takes as an input their action, aggregate outcomes $Y_t \in \mathcal{Y}$, 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 $\Phi : \mathbb{R} \to \mathbb{R}_+$ equal to $\Phi(x - x_{-1})$ when their last action was x_{-1}. Thus, we let their flow utility be $u(x, Y, \theta, \tilde{\theta}) - \Phi(x - x_{-1})$. 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^x_{t-1}, 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 $\hat{Y}(F^x_{t-1}, Q_t, \theta_t, \theta_{t-1})$. 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^x_{t-1}, Q_t).

We can therefore represent any firm policy function as:

$$x_{it} = g(x_{it-1}, \theta_{t-1}, \tilde{\theta}_{it-1}, F^x_{t-1}, Q_t, \lambda_{it}, s_{it}) \tag{223}$$

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

$$x_{it} \approx \gamma_i + \chi_t + \sum_{k=1}^{K} \delta_k \lambda_{k,it} + \gamma \theta_{it-1} + \omega x_{it-1} + \varepsilon_{it} \tag{224}$$

where the aggregate fixed effect now absorbs $(F^x_{t-1}, Q_t, \theta_t, \theta_{t-1})$, 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 `nltk` 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.[41] 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

[41] In 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 (2022). In this Appendix, we review essential points. We refer the reader to the Appendix material of Flynn and Sastry (2022) for certain details.

Our dataset is Compustat Annual Fundamentals. Our main variables of interest are defined in Appendix Table A20. 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 A20. 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`).[42] As in Ottonello and Winberry (2020) and Flynn and Sastry (2022), 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 add 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 a better balance of sector size. More summary information about these industries is provided in Appendix F of Flynn and Sastry (2022).

D.2 Compustat: Calculation of TFP

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

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

[42] A small difference from Flynn and Sastry (2022) 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:

$$\text{Share}_{M,j'} = \frac{\sum_{i:j(i)=j'} \sum_t \text{MaterialExpenditure}_{it}}{\sum_{i:j(i)=j'} \sum_t \text{Sales}_{it}} \tag{225}$$

2. If $\text{Share}_{M,j'} \leq \mu^{-1}$, then set

$$\begin{aligned} \alpha_{M,j'} &= \mu \cdot \text{Share}_{M,j'} \\ \alpha_{K,j'} &= 1 - \alpha_{M,j'} - \alpha_{L,j'} \end{aligned} \tag{226}$$

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

$$\begin{aligned} \alpha_{M,j'} &= 1 \\ \alpha_{K,j'} &= 0 \end{aligned} \tag{227}$$

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

$$\log \tilde{\theta}_{it} = \log \text{Sales}_{it} - \frac{1}{\mu} \left(\alpha_{M,j(i)} \cdot \log \text{MatExp}_{it} + \alpha_{K,j(i)} \cdot \log \text{CapStock}_{it} \right) \tag{228}$$

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

$$\log \hat{\theta}_{it} = \log \tilde{\theta}_{it} - \chi_{j(i),t} \tag{229}$$

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 isoleastic, 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

$$Y_i = \theta_i M_i^{\alpha_j} K_i^{1-\alpha_j} \tag{230}$$

They act as a monopolist facing the demand curve

$$p_i = Y_i^{-\frac{1}{\epsilon}} \tag{231}$$

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

$$\max_{M_i, K_i} \left\{ (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} - q_j M_i - r_j K_i \right\} \tag{232}$$

We first justify our formulas for the input shares (Equation 226). 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

$$\begin{aligned} q_j &= M_i^{-1} \alpha_j \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \\ r_j &= K_i^{-1} (1-\alpha_j) \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \end{aligned} \tag{233}$$

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

$$\begin{aligned} \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{q_j M_i}{p_i Y_i} \\ 1 - \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{r_j K_i}{p_i Y_i} \end{aligned} \tag{234}$$

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

$$Y = \left(\left(\frac{\epsilon-1}{\epsilon}\right)\theta_i\left(\frac{\alpha_j}{q_j}\right)^{\alpha}\left(\frac{1-\alpha_j}{r_j}\right)^{1-\alpha_j}\right)^{\epsilon} \Rightarrow p = \left(\frac{\epsilon}{\epsilon-1}\right)\theta_i^{-1}\left(\frac{q_j}{\alpha_j}\right)^{\alpha_j}\left(\frac{r_j}{1-\alpha_j}\right)^{1-\alpha_j} \tag{235}$$

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:

$$\alpha = \mu\frac{q_j M_i}{p_i Y_i} \tag{236}$$

Finally, we now apply Equations 228 and 229 to calculate productivity. Assume that we observe materials expenditure $q_j M_i$ and capital value $p_{K,j}K_i$, where $p_{K,j}$ is an (unobserved) price of capital. We find

$$\log\tilde{\theta}_i = \left(1-\frac{1}{\epsilon}\right)(\log\theta_i - \alpha\log q_j - (1-\alpha)\log p_{K,j}) \tag{237}$$

We finally observe that the industry-level means are

$$\chi_j = \left(1-\frac{1}{\epsilon}\right)\left(\log\bar{\theta}_j - \alpha\log q_j - (1-\alpha)\log p_{K,j}\right) \tag{238}$$

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

$$\log\hat{\theta}_i = \left(1-\frac{1}{\epsilon}\right)(\log\theta_i) \tag{239}$$

or our measurement captures physical TFP, up to scale.

E Additional Empirical Results

E.1 Alternative Empirical 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.

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

The Effect of Optimism on Hiring. We first revisit our empirical strategy for measuring the effect of optimism on firms' hiring, using the CEO change event studies. 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

$$\Delta \log L_{it} = \delta^{CEO} \text{opt}_{it} + \psi \, \text{opt}_{i,t-1} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it} \tag{240}$$

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.[43] 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 $\chi_{j(i),t}$, 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.

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

We present our results in Table A21. 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 its interaction with the new management's optimism. Specifically, we estimate

$$\begin{aligned}\Delta \log L_{it} = \delta^{\text{NoChange}} \text{opt}_{it} + \delta^{\text{Change}} (\text{opt}_{it} \times \text{ChangeCEO}_{it}) + \alpha^{\text{Change}} \text{ChangeCEO}_{it} \\ + \psi \, \text{opt}_{i,t-1} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it}\end{aligned} \tag{241}$$

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.

Virality from CEO Change Spillovers. We next 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

$$\overline{\text{opt}}^{\text{ceo}}_{j(i),t-1} = \frac{1}{|M_{j(i),t}|} \sum_{k \in M^c_{j(i),t}} \text{opt}_{k,t-1} \tag{242}$$

where $M_{j(i),t}$ is the set of firms in industry $j(i)$ at time t, and $M^c_{j(i),t}$ is the subset that had plausibly exogenous CEO changes. We construct the peer-set instrument $\overline{\text{opt}}^{\text{ceo}}_{p(i),t-1}$ analogously. We use $(\overline{\text{opt}}^{\text{ceo}}_{j(i),t-1}, \overline{\text{opt}}^{\text{ceo}}_{p(i),t-1})$ as instruments for $(\overline{\text{opt}}_{j(i),t-1}, \overline{\text{opt}}_{p(i),t-1})$ in the estimation of Equation 19. We present the corresponding estimates in Table A22. We find similar point estimates under IV and OLS, although the IV estimates are significantly noisier.

E.2 Narrative Optimism, Beliefs, and Hiring

In this appendix, we study whether narrative optimism, measured using text-analysis methods, matters for firm decisions conditional on firm-manager beliefs, measured from recorded managerial guidance. We find that narrative optimism and measured expectations each have predictive power conditional on the other for explaining hiring and capital investment. These results suggest that textual optimism captures aspects of managers' latent beliefs not captured in traditional measurement of expectations (here, in guidance data).

Data. We collect data from IBES (the International Brokers' Estimate System) on quantitative forecasts by company managers for three statistics: sales, capital expenditures

(CAPX), and earnings per share (EPS). As described in Section 3.1, we restrict to the first recorded forecast per fiscal year of that year's variable. When managers' guidance is reported as a range, we code a point-estimate forecast as the range's midpoint. For each variable $Z \in \{\text{Sales}, \text{Capx}, \text{Eps}\}$, we calculated the manager's predicted growth for fiscal year t as

$$\text{ForecastGrowthZ}_{it} = \log \text{GuidanceForX}_{it} - \log Z_{i,t-1} \tag{243}$$

For example, $\text{GuidanceForSales}_{it}$ is the manager's earliest recorded guidance within fiscal year t for fiscal-year t sales, and $\text{Sales}_{i,t-1}$ are recorded sales from fiscal year $t-1$. Textual narrative optimism opt_{it} is measured as in our main analysis.

Empirical Strategy. We re-create our main regression model, predicting hiring by opt_{it} conditional on firm fixed effects and industry-by-time fixed effects. We now include, as control variables, each of the $\text{ForecastGrowthZ}_{it}$ variables:

$$\Delta \log L_{it} = \delta^{OP}\text{opt}_{it} + \delta^{Z}\,\text{ForecastGrowthZ}_{it} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{244}$$

The coefficient δ^{OP} measures the difference in hiring between textually optimistic and non-optimistic firms, holding fixed forecasted growth about variable Z (and the fixed effects). The coefficient δ^{Z} measures the marginal effect of forecasted growth, in variable Z, on hiring, holding fixed whether the firm is optimistic or pessimistic (and the fixed effects). We also estimate a variance with net investment, or $\Delta \log K_{it}$, on the left-hand side.

Results. Table A23 shows the results when hiring is the outcome. We find that forecasted sales, CAPX, and earnings growth have positive effects of hiring, the first two of which are statistically significant (columns 2-4). Nonetheless, conditional on these variables, optimism has a positive effect on hiring of comparable magnitude to the baseline (column 1). The effect is statistically significant conditional on forecasted sales and CAPX growth (respectively, $t = 1.80$ and $t = 4.54$). Both optimism and forecasted EPS growth are insignificant predictors of hiring on the small ($N = 1290$) sample for which we can obtain EPS growth forecasts.

To compare the magnitudes of effects, we can calculate standardized coefficients. These have units of the effect of a one-standard-deviation change in the regressor on standard deviations of the outcome. For column 2, the standardized coefficient on textual optimism is 0.057 (SE: 0.0030) and the coefficient on predicted sales growth is 0.213 (SE: 0.0329). In this sense, predicted sales growth, for the subset of firms for which it is available, explains *larger* variations in hiring than textual optimism; but nonetheless, textual optimism has a statistically and economically significant effect.

Table A24 shows analogous results when net capital investment is the outcome. As with hiring, we verify that predicted sales and CAPX growth have statistically significant, positive effects on capital investment, and that optimism has a positive effect conditional on these variables. Effects on the sub-sample with earnings guidance are noisy, for both the effects of optimism and the effects of forecasts.

Discussion. We interpret our results in a model in which textual optimism, opt_{it}, is one measurement of a non-fundamental shifter in firm managers' beliefs. We validate this interpretation in the paper by showing that opt_{it}: (i) predicts hiring, as reviewed above in this note; (ii) does not predict future positive firm performance; and (iii) does correlate with optimistic manager forecasts, when measured in a variety of ways. We interpret managerial forecasts, about a variety of firm-specific variables, as alternative possible measurements of beliefs and their non-fundamental component. We are agnostic, more or less, about which of these measures explains more variation in firm actions or does so more precisely.

More broadly, while forecasts are quantitative, and provide *hard information* about managers' beliefs, they also capture at best only one or two moments of a probability distribution. By contrast, our measures of text all us to capture information about managers' beliefs that they do not express numerically, *i.e.,* we capture *soft information* about managers' beliefs (Liberti and Petersen, 2019). We find that this soft aspect of managers' beliefs is important for explaining their decisions conditional on hard information. This is consistent with extensive economic and psychological evidence that humans do not naturally think probabilistically (see *e.g.,* Tversky and Kahneman, 1973). Language may reflect nuances not present in the forecasts. These nuances are actually what we want to map to economic models, where we (economists) introduce statistical beliefs to model sentiment. This is the sense in which language might measure aspects of beliefs that are not captured in "measured beliefs." In this way, our results relate to a literature focusing on the decision-relevance of measured beliefs (Gennaioli, Ma, and Shleifer, 2016). They are moreover consistent with the literature focusing on the decision-relevance of textually measured firm-level variables including reported risks (Hassan, Hollander, Van Lent, and Tahoun, 2019) and reported uncertainty (Handley and Li, 2020).

E.3 State-Dependent Effects of Sentiment

Our main empirical framework assumes that the effect of narrative sentiment on hiring does not depend on previous sentiment. As one concrete example, this rules out the possibility that switching from relative optimism to relative pessimism has a larger effect than remaining equally pessimistic for two consecutive periods. To test for such state-dependent effects, we estimate augmented regression equations of the form:

$$\begin{aligned}\Delta \log L_{it} = \delta_0 \text{sentiment}_{it} + \delta_1 \text{sentiment}_{i,t-1} + \\ \delta_2(\text{sentiment}_{it} \times \text{sentiment}_{i,t-1}) + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it}\end{aligned} \tag{245}$$

where sentiment_{it} is our continuous measure of firm sentiment in language, $(\gamma_i, \chi_{j(i),t})$ are fixed effects at the firm and industry-by-time levels, and X_{it} is a vector of controls. This model allows for the marginal effect of this fiscal year's sentiment to depend on the level of the previous fiscal year's sentiment. In particular, if $\delta_2 > 0$, and the marginal effect of sentiment is positive, then this marginal effect is higher for a previously positive firm; if $\delta_2 < 0$, and the marginal effect of sentiment is positive, then this marginal is lower for a previously positive firm.

Table A25 shows our results, for different choices of controls. We find significant evidence of positive marginal effects for sentiment_{it} and $\delta_2 < 0$, or larger marginal effects when lagged sentiment is low. This asymmetry is quantitatively small, however, in the following sense. The standard deviation of $\text{sentiment}_{i,t-1}$ is 1.14. Using the estimates of column 1, a one-standard-deviation increase in sentiment, starting from $\text{sentiment}_{i,t-1} = 0$, decreases the marginal effect of sentiment_{it} from 0.022 to 0.016.

E.4 Measuring Virality via Granular Instrumental Variables

As an alternative strategy to estimate virality, we apply the methods of Gabaix and Koijen (2020) to construct "granular variables" that aggregate idiosyncratic variation in large firms' narrative loadings. We find evidence that the idiosyncratic optimistic updating of large firms induces optimistic updating, a form of virality.

Constructing the Granular Measures. We construct our granular instruments via the following algorithm. We first estimate a firm-level updating regression that controls non-parametrically for aggregate trends and parametrically for firm-level conditions. Specifically, we estimate

$$\text{opt}_{it} = \tau' X_{it} + \chi_{j(i),t} + \gamma_i + u_{it} \tag{246}$$

where $\chi_{j(i),t}$ is an industry-by-time fixed effect (sweeping out industry-specific aggregate shocks), γ_i is a firm fixed effect (sweeping out compositional effects), and X_{it} is the largest vector of controls used in the analysis of Section 4.1, consisting of: lagged log employment, current and lagged log TFP, log stock returns, the log book to market ratio, and leverage. We construct the empirical residuals $\hat{u}_{it}$. To construct the aggregate granular variable, $\overline{\text{opt}}_t^{g,sw}$, we take a sales-weighted average of these residuals:

$$\overline{\text{opt}}_t^{g,sw} = \sum_i \frac{\text{sales}_{it}}{\sum_i \text{sales}_{it}} \hat{u}_{it} \tag{247}$$

To construct an industry-level granular variable, $\overline{\text{opt}}_{j(i),t}^{g,sw}$, we take the leave-one-out sales-weighted average of the $\hat{u}_{it}$:

$$\overline{\text{opt}}_t^{g,sw} = \sum_{i':j(i)=j(i'),i'\neq i} \frac{\text{sales}_{i't}}{\sum_i \text{sales}_{i't}} \hat{u}_{i't} \tag{248}$$

We also construct agggregate and industry (leave-one-out) averages of opt_{it} for comparison. We denote these variables as $\overline{\text{opt}}_t^{sw}$ and $\overline{\text{opt}}_{j(i),t}^{sw}$, respectively.

Empirical Strategy. At the aggregate level, we first consider a variant of our main model Equation 18, but with one of the sales-weighted variables $Z_t \in \{\overline{\text{opt}}_t^{sw}, \overline{\text{opt}}_t^{g,sw}\}$:

$$\text{opt}_{it} = u \ \text{opt}_{i,t-1} + s \ Z_{t-1} + r \ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it} \tag{249}$$

The coefficient s measures virality with respect to the sales-weighted measures of optimism. We estimate Equation 249 by OLS, and also estimate a version in which the granular variable $\overline{\text{opt}}_t^{g,sw}$ is an instrumental variable for the raw sales-weighted average $\overline{\text{opt}}_t^{sw}$.

Similarly, at the industry level, we estimate the model

$$\text{opt}_{it} = u_{\text{ind}} \ \text{opt}_{i,t-1} + s_{\text{ind}} \ Z_{j(i),t-1} + r_{\text{ind}} \ \Delta \log Y_{j(i),t-1} + \gamma_i + \chi_t + \varepsilon_{it} \tag{250}$$

for $Z_{j(i),t} \in \{\overline{\text{opt}}_{j(i),t}^{sw}, \overline{\text{opt}}_{j(i),t}^{g,sw}\}$. As above, we estimate this first via OLS for each outcome variable, and then via IV where the granular variable $\overline{\text{opt}}_{j(i),t}^{g,sw}$ is an instrument for the raw sales-weighted average $\overline{\text{opt}}_{j(i),t}^{sw}$.

Results. We present our results in Table A26. First, studying aggregate virality, we find strong evidence that $s > 0$ when measured with the raw sales-weighted average or its granular component (columns 1 and 2). We moreover find significant evidence of $s > 0$ in the IV estimation (column 3). Our IV point estimate of $\hat{s} = 0.308$ greatly exceeds the OLS estimate of $\hat{s} = 0.0847$.

At the industry level, we find strong evidence of virality via the sales-weighted measure (column 4). We find imprecise estimates, centered around 0, for virality measured with the granular variable (column 5) or via the granular IV (column 6). However, the granular IV estimate is noisily estimated and is not significantly different from the point estimate of column 4.

F Additional Details on Model Quantification

F.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 Appendix B.5, which allows for (among other features) persistent fundamentals.

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

$$\log \theta_t = (1-\rho)\mu + \rho \log \theta_{t-1} + \sigma \nu_t \tag{251}$$

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 13 establishes that equilibrium can be written as (f does not depend on θ_{t-1} here as all agents believe persistence is ρ)

$$\log Y_t = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t) \tag{252}$$

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

$$\log Y_t^f = a_1 \log \theta_t + a_2 \log \theta_{t-1} \tag{253}$$

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

$$\log Y_t^f - \rho \log Y_{t-1}^f = a_1 \sigma \nu_t + a_2 \sigma \nu_{t-1} \tag{254}$$

It remains to solve for the coefficients (a_1, a_2). In particular, Equations 172 and 173 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 $\delta_{t,k}$ for $k \in \{O, P\}$:

$$\begin{aligned}\delta_{t,k} = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \Bigg[& \log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left[\log \gamma_i + \kappa \log \theta_t + (1-\kappa)((1-\rho)\mu_k + \rho \log \theta_{t-1})\right] \\ & - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma\right)^2 \sigma^2_{\theta|s} \\ & + \left(\frac{1}{\epsilon} - \gamma\right)\left[a_0 + a_1\left(\kappa \theta_t + (1-\kappa)((1-\rho)\mu_k + \rho \log \theta_{t-1})\right) + a_2 \log \theta_{t-1} + f(Q_t)\right] \Bigg]\end{aligned} \tag{255}$$

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,

$$\alpha\delta^{OP} := \delta_{t,O} - \delta_{t,P} = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \right) (1-\kappa)(1-\rho)(\mu_O - \mu_P) \quad (256)$$

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 173 as

$$\log Y_t = \delta_{t,P} + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 + \frac{\epsilon}{\epsilon-1} \log \left(Q_t \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha\delta^{OP} \right\} \right) \quad (257)$$

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

$$\begin{aligned} \log Y_t = C &+ \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha}\kappa + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \kappa \right] \log \theta_t \\ &+ \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha}(1-\kappa)\rho + \left(\frac{1}{\epsilon} - \gamma \right) a_1 (1-\kappa)\rho \right] \log \theta_{t-1} \\ &+ \frac{\epsilon}{\epsilon-1} \log \left(Q_t \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha\delta^{OP} \right\} \right) \end{aligned} \quad (258)$$

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

$$f(Q_t) = \frac{\epsilon}{\epsilon-1} \log \left(Q_t \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha\delta^{OP} \right\} \right) - \frac{\epsilon}{\epsilon-1} \log \left(\frac{1}{2} \exp \left\{ \frac{\epsilon-1}{\epsilon} \alpha\delta^{OP} \right\} \right) \quad (259)$$

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 $\omega = (1/\epsilon - \gamma)/((1+\psi-\alpha)/\alpha + 1/\epsilon)$,

$$a_1 = \frac{1}{1-\kappa\omega} \frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \quad (260)$$

Finally, from matching coefficients for a_2,

$$a_2 = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \right] (1-\kappa)\rho \quad (261)$$

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

$$P_O^H(\log Y, Q) = [u + r\log Y + sQ + C_P]_0^1 \quad \text{and} \quad P_P^H(\log Y, Q) = [-u + r\log Y + sQ + C_P]_0^1 \tag{262}$$

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 $C_P = \frac{1-s}{2}$

F.2 Calibration Methodology

To calibrate the model, we proceed in four steps.

1. *Setting macro parameters.* We first set $(\epsilon, \gamma, \psi, \alpha)$. In Section 7.1 and Table 6, we describe our baseline method based on matching estimates of the deep parameters from the literature. We also consider two other strategies as robustness checks. First, to target estimated fiscal multipliers in the literature, we use the same external calibration of α (returns to scale) and ϵ (elasticity of substitution), and set (γ, ψ) to match the desired multiplier. Since the exact choice of these parameters is arbitrary subject to obtain the correct multiplier, we normalize $\gamma = 0$ and vary only ψ. Second, we match an estimate of the multiplier implied by our own data and an exact formula for the omitted variable bias incurred in estimating the effect of optimism on hiring without controlling for general-equilibrium effects via a time fixed effect. We outline that strategy for estimating the multiplier in Section F.3 below, and we map this to deep parameters exactly as described in our method for matching the literature's estimated multiplier.
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 259. 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 are estimated in column 1 of Table 4.
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 US real GDP (Baxter and King, 1999).[44] We apply

[44] Specifically, we filter to post-war quarterly US 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.

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

$$\log \hat{Y}_t^f = \log \hat{Y}_t - \hat{f}(\hat{Q}_t) \tag{263}$$

(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

$$\log \hat{Y}_t^f - 0.086 \log \hat{Y}_t^f = .0078(\nu_t + .32\,\nu_{t-1}) \tag{264}$$

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 6, which also is therefore the maximum likelihood estimate of (κ, σ).

F.3 Estimating a Demand Multiplier in Our Empirical Setting

Here, we describe a method for estimating a demand multiplier in our data on optimism and firm hiring. This circumvents the step of external calibration for the multiplier, but relies on correct specification of the time-series correlates of aggregate optimism. Reassuringly, this method yields a general-equilibrium demand multiplier that is comparable to our baseline calibration and our literature-derived calibration.

Mapping the Model to Data. Extending Corollary 1 with the calculations of Appendix B.5 and Appendix F.1, we first observe that firms' hiring can be written in equilibrium as

$$\Delta \log L_{it} = \tilde{c}_{0,i} + \tilde{c}_{10} \log \theta_t + \tilde{c}_{11} \log \theta_{t-1} + \tilde{c}_2 f(Q_t) + \tilde{c}_3 \log \theta_{it} + \tilde{c}_4 \log L_{i,t-1} + \delta^{OP} \lambda_{it} + \zeta_{it} \tag{265}$$

where ζ_{it} is an IID normal random variable with zero mean and λ_{it} is the indicator for having adopted the optimistic narrative.

In the data, our estimating equation without control variables had the following form

$$\Delta \log L_{it} = \gamma_i + \chi_{j(i),t} + \delta^{OP} \mathrm{opt}_{it} + z_{it} \tag{266}$$

This maps to the structural model with $\gamma_i = \tilde{c}_{0,i}$, $\chi_{j(i),t} = \tilde{c}_{10} \log \theta_t + \tilde{c}_{11} \log \theta_{t-1} + \tilde{c}_2 f(Q_t)$, and $z_{it} = \zeta_{it} + \tilde{c}_3 \log \theta_{it} + \tilde{c}_4 \log L_{i,t-1}$. Under the hypothesis that $\mathbb{E}[z_{it}\mathrm{opt}_{it}] = 0$, then the OLS regression of $\Delta \log L_{it}$ on opt_{it}, conditional on the indicated fixed effects, identifies δ^{OP}.

We consider now an alternative regression equation which is a variant of the above specification without the time fixed effect and with parametric controls for aggregate TFP:

$$\Delta \log L_{it} = \gamma_i + \delta^{OP} \text{opt}_{it} + \tilde{c}_{10} \log \theta_t + \tilde{c}_{11} \log \theta_{t-1} + \tilde{z}_{it} \tag{267}$$

Observe that the new residual, relative to the old residual, is contaminated by the equilibrium effect of optimism. That is, $\tilde{z}_{it} = z_{it} + \tilde{c}_2 f(Q_t)$. To refine this further, we apply the linear approximation $f(Q_t) \approx \frac{\alpha \delta^{OP}}{1-\omega} Q_t$ and the observation that $\tilde{c}_2 = \omega$, so we can write $\tilde{z}_{it} = z_{it} + \frac{\alpha\omega}{1-\omega} \delta^{OP} Q_t$.

We now derive a formula for omitted variables bias in the estimate of δ^{OP} from an OLS estimation of Equation 267. Let X denote a finite-dimensional matrix of data on opt_{it}, firm-level indicators (*i.e.,* the regressors corresponding to the firm fixed effects), and current and lagged aggregate TFP. Similarly, let Y be a finite-dimensional matrix of data on $\Delta \log L_{it}$. The OLS regression coefficient in this finite sample is $\hat{\delta} = ((X'X)^{-1} X'Y)_1$. Using the standard formula for omitted variables bias:

$$\begin{aligned} \mathbb{E}[\hat{\delta}|X] &= \delta^{OP} + \left((X'X)^{-1} \mathbb{E}[X'Q|X] \frac{\alpha\omega}{1-\omega} \delta^{OP} \right)_1 \\ &= \delta^{OP} \left(1 + \frac{\alpha\omega}{1-\omega} \left((X'X)^{-1} \mathbb{E}[X'Q|X] \right)_1 \right) \end{aligned} \tag{268}$$

where Q is the vector of observations of Q_t. We can then observe that:

$$(X'X)^{-1} \mathbb{E}[X'Q|X] = \mathbb{E}\left[(X'X)^{-1} X'Q | X \right] \tag{269}$$

Which is the (expected) OLS estimate of β in the following regression:

$$Q_t = \gamma_i + \beta_O^Q \text{opt}_{it} + \beta_\theta^Q \log \theta_t + \beta_{\theta_{-1}}^Q \log \theta_{t-1} + \varepsilon_t \tag{270}$$

But we observe that, averaging both sides, that $\gamma_i = \beta_\theta^Q = \beta_{\theta_{-1}}^Q = 0$ and $\beta_O^Q = 1$. Thus, $((X'X)^{-1} \mathbb{E}[X'Q|X])_1 = 1$. We therefore obtain that:

$$\mathbb{E}[\hat{\delta} \mid X] = \delta^{OP} \left(1 + \frac{\alpha\omega}{1-\omega} \right) \tag{271}$$

Hence, given a population estimate of the *biased* OLS estimate and an external calibration of α, we can pin down the complementarity ω and the multiplier $\frac{1}{1-\omega}$. Naturally this strategy relies on correctly measuring aggregate TFP as measurement error in that variable would contaminate this estimation. Moreover, it requires us to assume that all variation in aggregate output that is not due to TFP is due to optimism or forces entirely orthogonal to optimism; in view of our running assumption that the spread of optimism is associative, these other forces therefore also have to be completely transitory, lest they be incorporated into current optimism via associative updating in a previous period. These assumptions are very strong and are why we do not adopt this strategy for our main quantitative analysis. Nevertheless, we will find very similar results, as we now describe.

Empirical Application and Results. To operationalize this in practice, we compare estimates of Equation 266 and 267. For the latter, we proxy TFP using the cyclical component of both capacity adjusted and capacity un-adjusted TFP using the data of Fernald (2014).[45] We moreover maintain the assumption of $\alpha = 1$, or constant returns to scale, to map our estimates back to implied multipliers.

Our results are reported in Table A27, along with the associated values of complementarity ω and the multiplier $\frac{1}{1-\omega}$. Using capacity-adjusted and unadjusted TFP, we respectively obtain estimates of 1.46 and 1.37 for the multiplier. These are lower than our baseline estimate, but comparable to our estimates based on structural modeling in the literature. Both estimates are below our baseline calibration of 1.96 but above our multiplier-literature calibration of 1.33. In Table A18, we report our quantitative results under the assumed multiplier of 1.46. We find that, as expected, these estimates imply an role for optimism that is an intermediate between the baseline and multiplier-literature calibrations.

F.4 Simulation Methodology

For our baseline 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.

[45]Mirroring our filtering of US real GDP, we apply the Baxter and King (1999) band-pass filter to post-war quarterly data using 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.

G 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.

G.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 $\mathcal{I}$) and might regard such a story is 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').

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 ($Y^*(Q) : \Theta \to \mathcal{Y}$) 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 $N^*(\theta, Y) = N(\theta) \times \mathbb{I}[Y = Y^*(\theta)]$. Importantly, the distribution of narratives Q and endogenous outcomes Y then shape the distribution of narratives tomorrow Q'. 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.

G.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 $Q_t, \theta_t \mapsto Y_t$. 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 A13, 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

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{272}$$

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 $1996 \leq \tau \leq 2019$ in our sample:[46]

$$\Delta \log L_{it} = \sum_{\tau=1996}^{2017} \delta_{\tau}^{OP} (\text{opt}_{it} \times \mathbb{I}[t = \tau]) + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{273}$$

We show the coefficient series of δ_{τ}^{OP} graphically in Figure A11. 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,

[46] The 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.

G.3 The Perennial Economic Narratives: Our Empirical Findings

Shiller (2020) identifies nine perennial economic narratives based on 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 matters for firm decision-making. 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 A13). Moreover, we find evidence of virality and stubbornness in firms holding onto these narratives (see Appendix Figure A9).

H Additional Figures and Tables

List of Figures

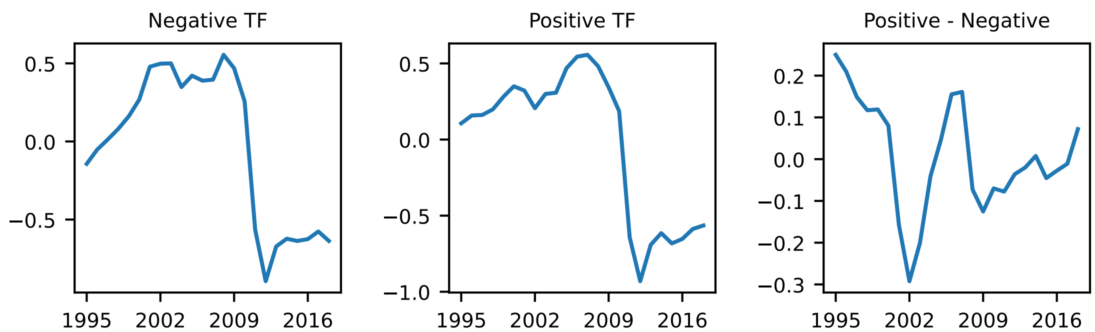
A1 Time-Series for Positive, Negative, and Their Difference . . . 129
A2 Time-Series for Shiller's Perennial Economic Narratives . . . 130
A3 Time-Series for the Selected LDA Topics . . . 131
A4 Heterogeneous Effects of Optimism on Hiring . . . 132
A5 Net Sentiment and Hiring . . . 132
A6 Dynamic Relationship between Optimism and Firm Fundamentals, Conference-Call Measurement . . . 133
A7 Dynamic Relationship between Optimism and Firm Fundamentals, Continuous Sentiment Measurement . . . 134
A8 Dynamic Relationship Between Optimism and Financial Variables . . . 135
A9 The Virality and Associativeness of Other Identified Narratives . . . 136
A10 Evaluating Potential for Hysteresis for All Decision-Relevant Narratives . . . 137
A11 Time-Varying Effects of Optimism on Hiring . . . 138
A12 Time-Varying Relationship Between Optimism and TFP . . . 138

List of Tables

A1 The Twenty Most Common Positive and Negative Words . . . 139
A2 The Twenty Most Common Words for Each Shiller Chapter . . . 140
A3 The Ten Most Common Words for Each Selected Topic . . . 141
A4 Persistence and Cyclicality of Narratives . . . 142
A5 Variance Decomposition of Narratives . . . 142
A6 Narrative Optimism Predicts Hiring, With More Adjustment-Cost Controls . 143
A7 Narrative Optimism Predicts Hiring, Alternative Standard Errors . . . 144
A8 Narrative Optimism Predicts Hiring, Instrumenting With Lag . . . 145
A9 Narrative Optimism Predicts Hiring, Conference-Call Measurement . . . 146
A10 The Effect of Narrative Optimism on All Inputs . . . 146
A11 The Effect of Narrative Optimism on Stock Prices, High-Frequency Analysis 147
A12 Textual Optimism and Optimistic Forecasts, Alternative Measurement . . . 147
A13 The Effects of All Selected Narratives on Hiring . . . 148
A14 Narrative Optimism is Viral and Associative, Alternative Standard Errors . . 149
A15 Narrative Sentiment is Viral and Associative . . . 150

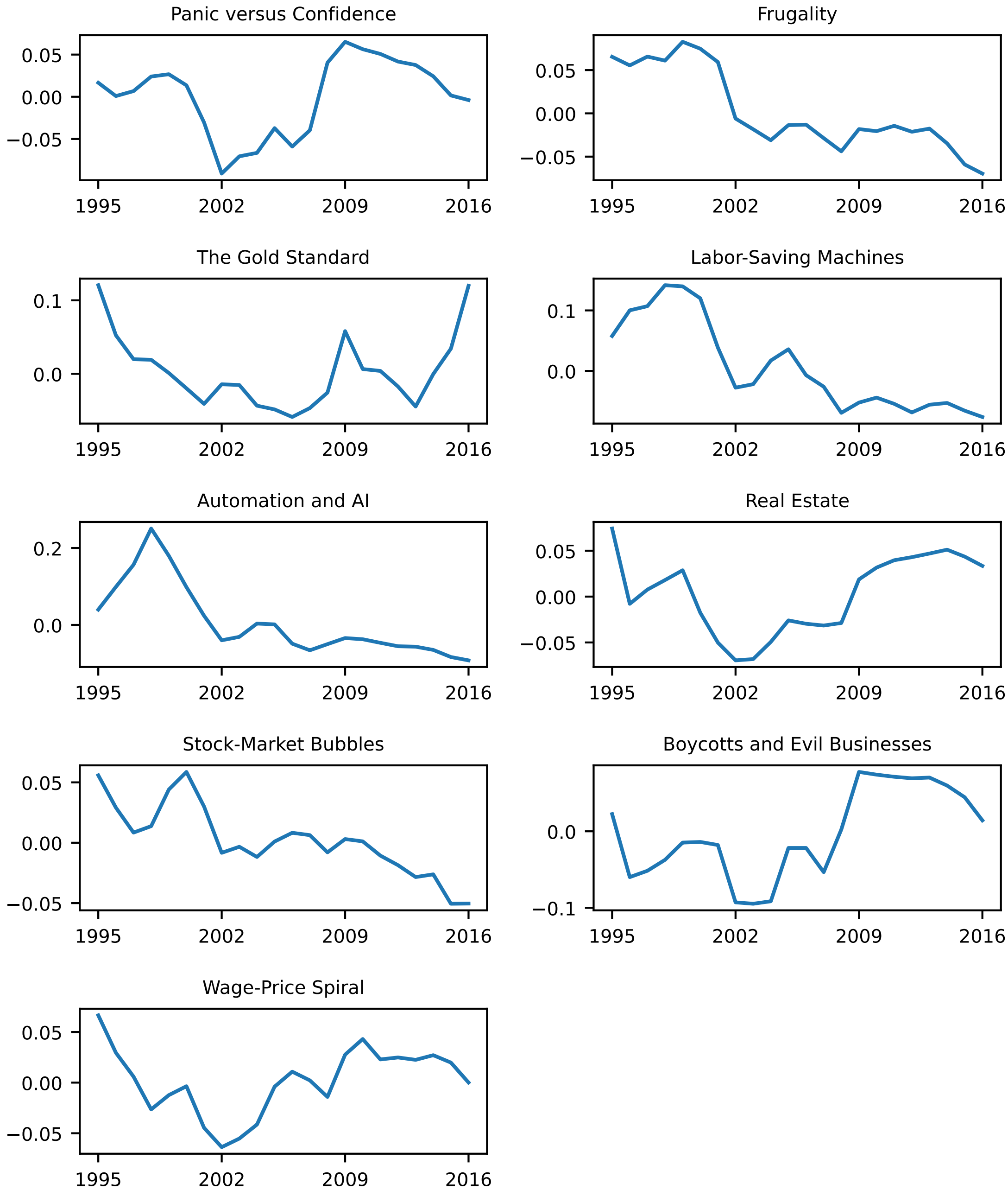
A16 Narrative Sentiment is Viral and Associative, Over-Controlling for Past and Future Outcomes . . . 151
A17 Narrative Optimism and Viral and Associative, Instrumented With Other Narratives . . . 152
A18 Sensitivity Analysis for the Quantitative Analysis . . . 153
A19 An Empirical Test for Cycles and Chaos . . . 154
A20 Data Definitions in Compustat . . . 154
A21 The Effect of Optimism on Hiring, CEO Change Strategy . . . 155
A22 The Virality of Optimism, CEO Change Strategy . . . 156
A23 Narrative Optimism Predicts Hiring, Conditional on Measured Beliefs . . . 156
A24 Narrative Optimism Predicts Investment, Conditional on Measured Beliefs . . . 157
A25 State-Dependent Effects of Sentiment on Hiring . . . 158
A26 Optimism is Viral and Associative, Granular IV Strategy . . . 159
A27 Multiplier Calibrations via Under-Controlled Regressions of Hiring on Optimism 160

Figure A1: Time-Series for Positive, Negative, and Their Difference



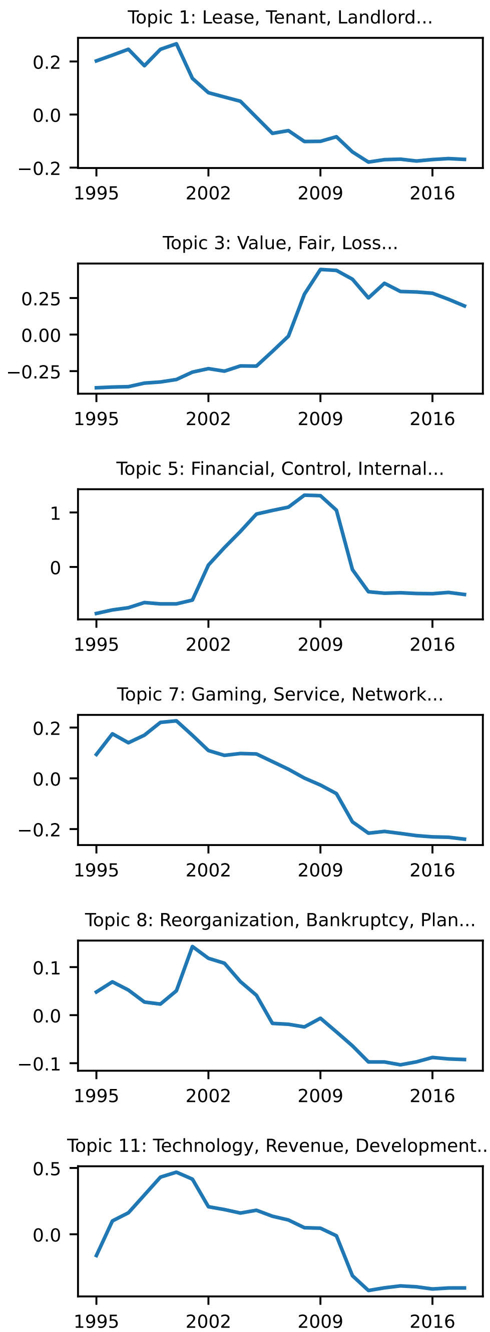
Notes: Negative and Positive term frequency (first two panels) are cross-sectional averages of z-score transformed variables. The third panel is the cross-sectional average of the difference between the two, or sentiment$_{it}$.

Figure A2: Time-Series for Shiller's Perennial Economic Narratives



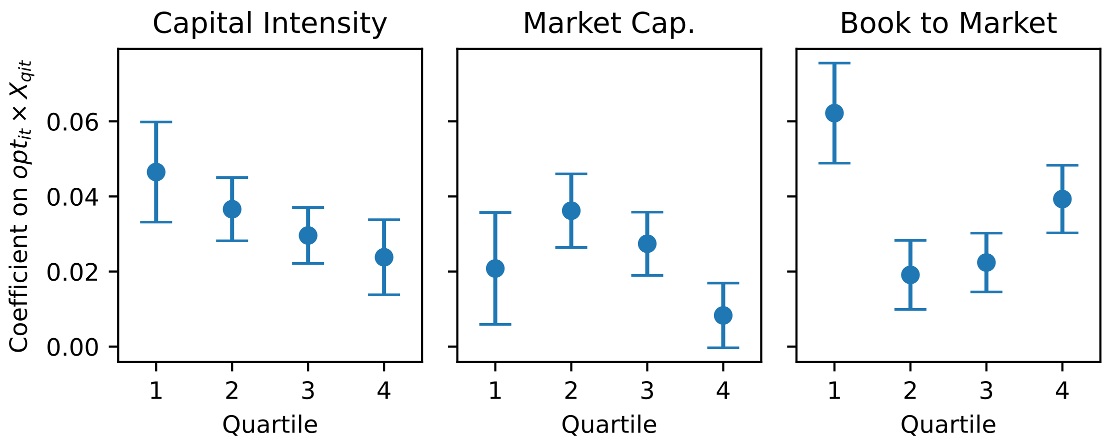
Notes: Each panel plots the time-series average of the narrative variable defined for the corresponding chapter of Shiller (2020)'s *Narrative Economics*. The units are cross-sectional averages of z-score transformed variables.

Figure A3: Time-Series for the Selected LDA Topics



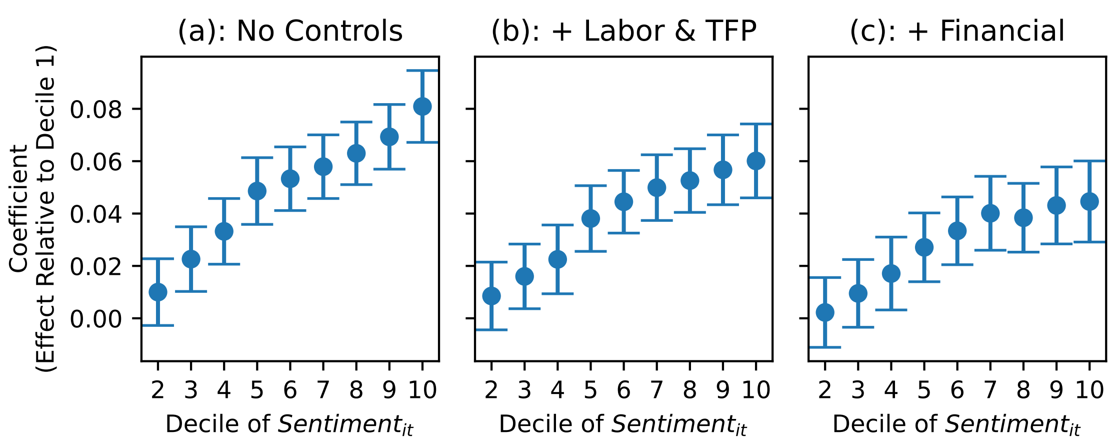
Notes: Each panel plots the time-series average of scores for the corresponding topics, identified by their three most common bigrams. The 11 topics are selected by the LASSO estimation described in Section 4.1, and estimates of which are reported in Table A13. The units in each panel are cross-sectional averages of z-score transformed variables.

Figure A4: Heterogeneous Effects of Optimism on Hiring



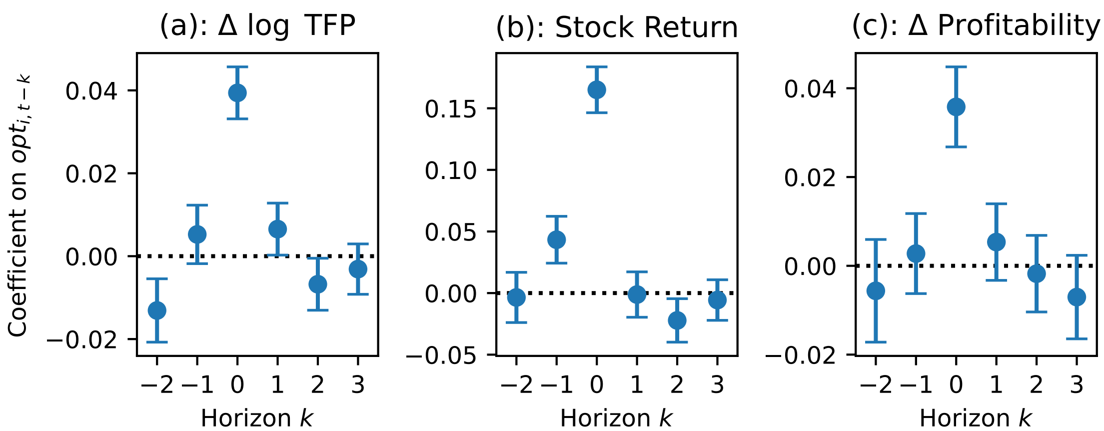
Notes: In each panel, we show estimates from the regression $\Delta \log L_{it} = \sum_{q=1}^{r} \beta_q \cdot (\text{opt}_{it} \times X_{qit}) + \gamma_i + \chi_{j(i),t} + \epsilon_{it}$, where X_{qit} indicates quartile q of the studied variable: one minus the variable cost share of sales, market capitalization, or book-to-market ratio. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are double-clustered by firm ID and industry-year.

Figure A5: Net Sentiment and Hiring



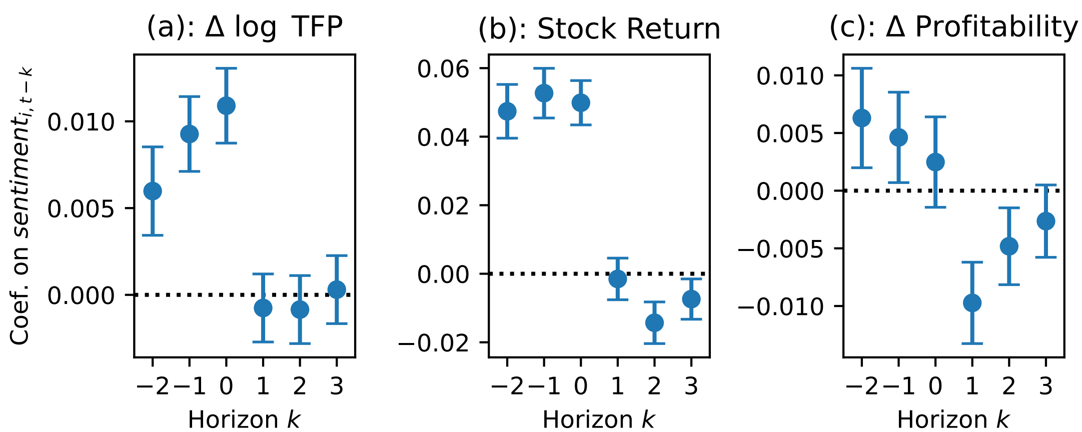
Notes: In each panel, we show estimates from the regression $\Delta \log L_{it} = \sum_{q=1}^{10} \beta_q \cdot (\text{sentiment}_{iqt}) + \tau' X_{it} + \gamma_i + \chi_{j(i),t} + \epsilon_{it}$, where sentiment_{iqt} indicates decile q of the continuous sentiment variable. Panel (a) estimates this equation without controls (like column 1 of Table 1); panel (b) adds controls for lagged labor and current and lagged log TFP (like column 2 of Table 1); and panel (c) adds controls for the log book to market ratio, log stock return, and leverage (like column 3 of Table 1). The excluded category in each regression is the first decile of sentiment_{it}. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are double-clustered by firm ID and industry-year.

Figure A6: Dynamic Relationship between Optimism and Firm Fundamentals, Conference-Call Measurement



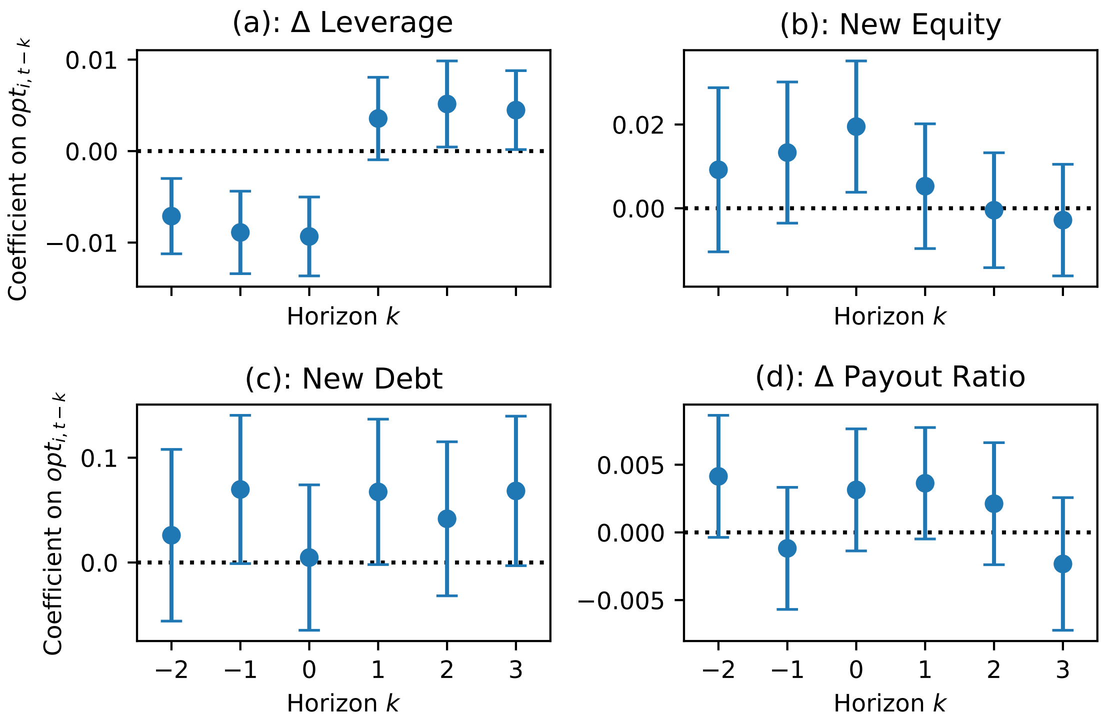
Notes: The regression model is Equation 14 (as in Figure 2), but measuring optimism from sales and earnings conference calls. Each coefficient is estimated from a separate projection regression. The outcomes are (a) the log change in TFP, calculated as described in Appendix D.2, (b) the log stock return inclusive of dividends, and (c) changes in profitability, defined as earnings before interest and taxes (EBIT) as a fraction of the previous fiscal year's variable costs. In all specifications, we trim the 1% and 99% tails of the outcome variable. Each coefficient is estimated from a separate projection regression. Error bars are 95% confidence intervals.

Figure A7: Dynamic Relationship between Optimism and Firm Fundamentals, Continuous Sentiment Measurement



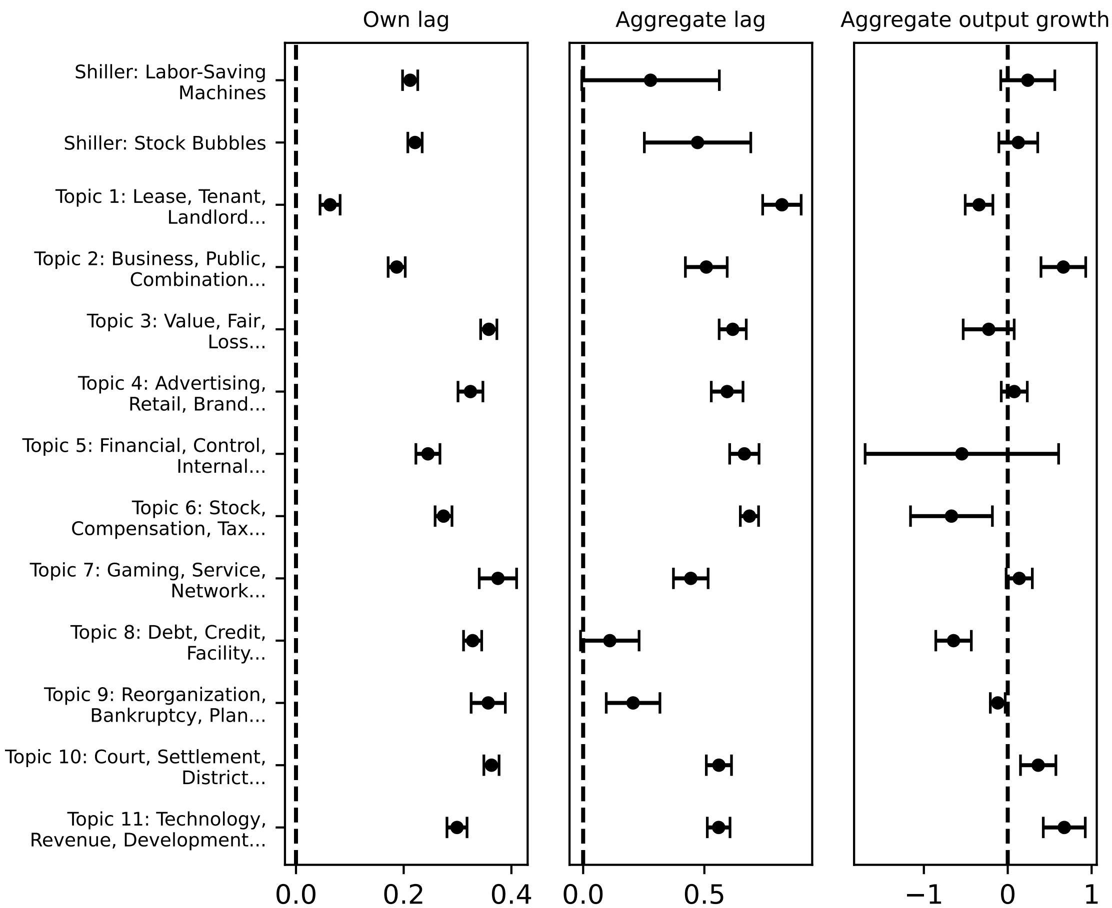
Notes: The regression model is Equation 14 (as in Figure 2), but using the continuous variable sentiment_{it}. Each coefficient is estimated from a separate projection regression. The outcomes are (a) the log change in TFP, calculated as described in Appendix D.2, (b) the log stock return inclusive of dividends, and (c) changes in profitability, defined as earnings before interest and taxes (EBIT) as a fraction of the previous fiscal year's variable costs. In all specifications, we trim the 1% and 99% tails of the outcome variable. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are two-way clustered by firm ID and industry-year.

Figure A8: Dynamic Relationship Between Optimism and Financial Variables



Notes: The regression model is Equation 14 (as in Figure 2), but with financial fundamentals as outcomes. Each coefficient is estimated from a separate projection regression. The outcome variables are: (a) the fiscal-year-to-fiscal-year difference in leverage, which is total debt (short-term debt plus long-term debt); (b) sale of common and preferred stock minus buybacks, normalized by the total equity outstanding in the previous fiscal year; (c) short-term debt plus long-term debt issuance, normalized by the total debt in the previous fiscal year; and (d) total dividends divided by earnings before interest and taxes (EBIT). In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are two-way clustered by firm ID and industry-year.

Figure A9: The Virality and Associativeness of Other Identified Narratives



Notes: For each narrative, we estimate the analogue of Equation 18. We first transform each continuous narrative loading variable $\hat{\lambda}_{k,it}$ (indexed by narrative identifier k, firm i, and time period t) into a binary indicator for being above the sample median,

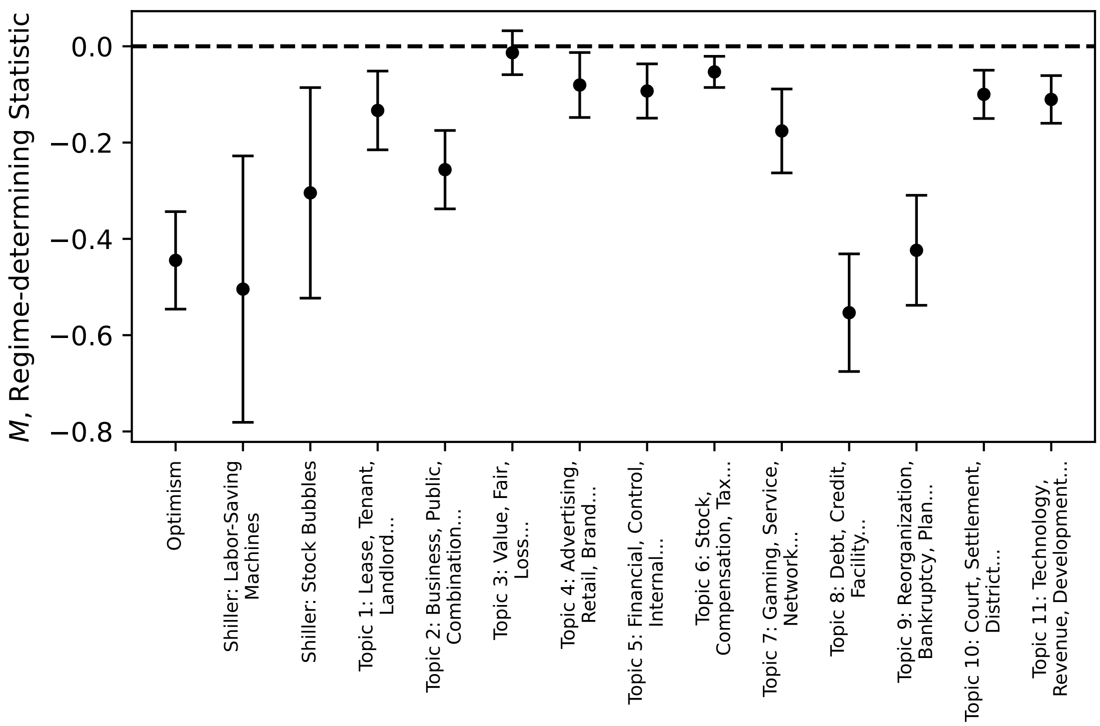
$$\hat{\lambda}^b_{k,it} = \mathbb{I}\left[\hat{\lambda}_{k,it} \geq \text{med}\left(\text{sentiment}_{it}\right)\right] \in \{0, 1\}$$

We then estimate

$$\hat{\lambda}^b_{k,it} = u_k \ \hat{\lambda}^b_{k,i,t-1} + s_k \ \overline{\hat{\lambda}^b}_{k,it} + r_k \ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it}$$

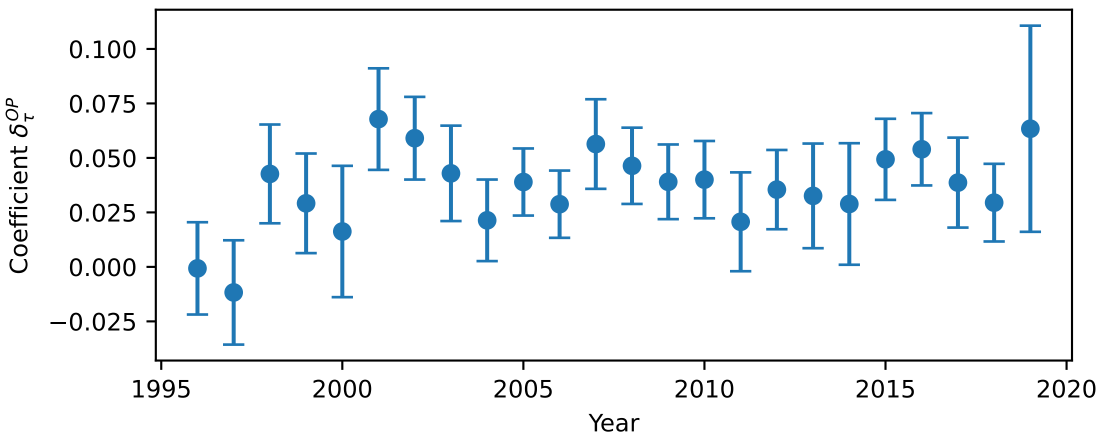
where $\overline{\hat{\lambda}^b}_{k,it}$ denotes an aggregate average of the binary variable. The three panels respectively show our estimates, for each narrative k, of (u_k, s_k, r_k). The error bars are 95% confidence intervals based on double-clustered (firm and industry-year) standard errors.

Figure A10: Evaluating Potential for Hysteresis for All Decision-Relevant Narratives



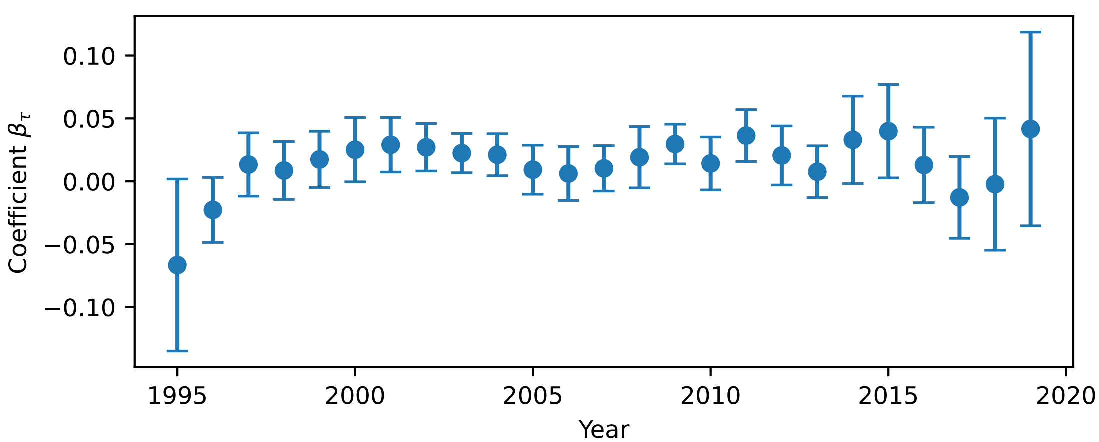
Notes: For the binary construction of each narrative, we estimate the parameters of the updating rule (see Figure A9) and the partial-equilibrium effect on hiring via a variant of Equation 12 (not reported for brevity). We then calculate the statistic $M = \hat{u} + \hat{r} + \hat{s} + \hat{r}\hat{f}(1)$ (see Section 7.3). We report our point estimates for each narrative along with 95% confidence intervals error bars, calculated using the delta method.

Figure A11: Time-Varying Effects of Optimism on Hiring



Notes: Each dot is a coefficient δ_τ^{OP} estimated from Equation 273, capturing a year-specific effect of binary optimism (opt_{it}) on hiring ($\Delta \log L_{it}$). Error bars are 95% confidence intervals, based on standard errors clustered by firm and industry-year.

Figure A12: Time-Varying Relationship Between Optimism and TFP



Notes: Each dot is a coefficient β_τ estimated from Equation 186, corresponding to a year-specific effect of binary optimism (opt_{it}) on log TFP ($\log \hat{\theta}_{it}$). The outcome variable is firm-level log TFP, $\log \theta_{it}$, and the regressors are indicators for binary optimism interacted with year dummies. In the regression, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals, based on standard errors clustered by firm and industry-time.

Table A1: The Twenty Most Common Positive and Negative Words

Positive	Negative
well	loss
good	decline
benefit	disclose
high	subject
gain	terminate
advance	omit
achieve	defer
improve	claim
improvement	concern
opportunity	default
satisfy	limitation
lead	delay
enhance	deficiency
enable	fail
able	losses
best	damage
gains	weakness
improvements	adversely
opportunities	against
resolve	impairment

Notes: The twenty most common lemmatized words among the 230 positive words and 1354 negative words. They are listed in the order of their document frequency. The words are taken from the Loughran and McDonald (2011) dictionary, as described in Section 3.2.

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
bank	help	standard	replac	replac	price	chapter	price	countri
consum	hous	book	produc	appear	appear	peopl	profit	labor
appear	buy	money	technolog	show	real	specul	good	union
show	home	run	appear	question	find	drop	consum	ask
forecast	famili	paper	book	suggest	hous	play	start	wage
economi	lost	peopl	power	labor	estat	depress	fall	inflat
suggest	display	metal	save	ask	buy	warn	buy	strong
run	job	depress	problem	run	home	peak	wage	world
concept	peopl	eastern	labor	worker	citi	great	inflat	mile
peopl	explain	almost	innov	vacat	land	today	world	peopl
grew	phrase	depositor	run	autom	movement	get	cut	happen
around	depress	young	wage	human	world	decad	shop	depress
weather	postpon	today	worker	univers	tend	reaction	peopl	war
figur	car	want	electr	world	peopl	newspap	explain	tri
confid	justifi	went	mechan	machin	never	news	campaign	wrote
wall	cultur	decad	human	job	search	storm	play	peak
happen	fashion	idea	world	peopl	specul	saw	depress	great
depress	unemploy	man	machin	answer	explain	memori	behavior	recess
tri	great	newspap	job	around	popul	interview	postpon	went
unemploy	fault	popular	invent	figur	phrase	watch	war	get

Notes: The twenty most common lemmatized words among the 100 words that typify each Shiller (2020) narrative. Our selection procedure is described in Section 3.2.

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	financial	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	affect	0.012	income	0.023
lessor	0.014	offer	0.014	debt	0.02	distribution	0.016	officer	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	financial	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. The LDA model and our estimation procedure are described in Section 3.2.

Table A4: Persistence and Cyclicality 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 2019 for optimism and the topics, and 1995 to 2017 for the Shiller Narratives. u_t is the US 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.

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 $1 - \text{Var}[N_{it}^{\perp}]/\text{Var}[N_{it}]$, where N_{it} is the narrative intensity and $N_{it}^{\perp}$ is the same after projecting out means at the indicated level. The last column ("All") partials out industry-by-time means and firm means.

Table A6: Narrative Optimism Predicts Hiring, With More Adjustment-Cost Controls

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0305	0.0257	0.0235	0.0184
	(0.0030)	(0.0034)	(0.0037)	(0.0039)
Firm FE	✓	✓	✓	✓
Industry-by-time FE	✓	✓	✓	✓
$\log L_{i,t-1}$	✓	✓	✓	✓
$(\log \hat{\theta}_{it}, \log \hat{\theta}_{i,t-1})$	✓	✓	✓	✓
$(\log L_{i,t-2}, \log \hat{\theta}_{i,t-2})$		✓	✓	✓
$(\log L_{i,t-3}, \log \hat{\theta}_{i,t-3})$			✓	✓
Log Book to Market				✓
Stock Return				✓
Leverage				✓
N	39,298	31,236	25,156	21,913
R^2	0.401	0.395	0.396	0.415

Notes: The regression model is Equation 12. Column 1 replicates column 2 of Table 1. Columns 2 and 3 add more lags of firm-level log employment and firm-level log TFP, and column 4 introduces the baseline financial controls (*i.e.,* those in column 3 of Table 1). In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A7: Narrative Optimism Predicts Hiring, Alternative Standard Errors

	(1)	(2)	(3)	(4)	(5)
	Outcome is				
	$\Delta \log L_{it}$				$\Delta \log L_{i,t+1}$
opt_{it}	0.0355	0.0305	0.0250	0.0322	0.0216
	(0.0030)	(0.0030)	(0.0032)	(0.0028)	(0.0037)
	[0.0031]	[0.0026]	[0.0031]	[0.0040]	[0.0034]
	{0.0035}	{0.0026}	{0.0025}	{0.0043}	{0.0036}
Firm FE	✓	✓	✓		✓
Industry-by-time FE	✓	✓	✓	✓	✓
Lag labor		✓	✓	✓	✓
Current and lag TFP		✓	✓	✓	✓
Log Book to Market			✓		
Stock Return			✓		
Leverage			✓		
N	71,161	39,298	33,589	40,580	38,402
R^2	0.259	0.401	0.419	0.142	0.398

Notes: This Table replicates the analysis of Table 1 with alternative standard error constructions. Standard errors in parentheses are two-way clustered by firm ID and industry-year; those in square brackets are two-way clustered by firm ID and year; and those in braces are two-way clustered by industry and year. For columns 1-4, the regression model is Equation 12 and the outcome is the log change in firms' employment from year $t-1$ to t. The main regressor is a binary indicator for the optimistic narrative, defined in Section 3.2. In all specifications, we trim the 1% and 99% tails of the outcome variable. In column 5, the regression model is Equation 13, the outcome is the log change in firms' employment from year t to $t+1$, and control variables are dated $t+1$.

Table A8: Narrative Optimism Predicts Hiring, Instrumenting With Lag

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0925	0.106	0.102	0.0470
	(0.0130)	(0.0160)	(0.0168)	(0.0053)
Firm FE	✓	✓	✓	
Industry-by-time FE	✓	✓	✓	✓
Lag labor		✓	✓	✓
Current and lag TFP		✓	✓	✓
Log Book to Market			✓	
Stock Return			✓	
Leverage			✓	
N	63,302	35,768	31,071	36,953
First-stage F	773	478	366	3,597

Notes: All columns come from a two-stage-least-squares (2SLS) estimate of Equation 12, using $\text{opt}_{i,t-1}$ as an instrument for opt_{it}. Specifically, the structural equation is

$$\Delta \log L_{it} = \delta^{OP} \cdot \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it}$$

and the first-stage equation is

$$\text{opt}_{it} = \alpha \cdot \text{opt}_{i,t-1} + \tilde{\gamma}_i + \tilde{\chi}_{j(i),t} + \tilde{\tau}' X_{it} + \tilde{\varepsilon}_{it}$$

In the last row, we report the first-stage F statistic associated with this equation. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A9: Narrative Optimism Predicts 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	0.0173	0.0121	0.0237	0.0123
	(0.0038)	(0.0040)	(0.0038)	(0.0038)	(0.0044)
Industry-by-time FE	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓		✓
Lag labor		✓	✓	✓	✓
Current and lag TFP		✓	✓	✓	✓
Log Book to Market			✓		
Stock Return			✓		
Leverage			✓		
N	19,625	11,565	10,851	11,919	11,416
R^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. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year. In column 5, control variables are dated $t + 1$.

Table A10: The Effect of Narrative 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	0.0305	0.0397	0.0193	0.0370	0.0273
	(0.0030)	(0.0030)	(0.0034)	(0.0033)	(0.0034)	(0.0036)
Industry-by-time FE	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Lag input		✓		✓		✓
Current and lag TFP		✓		✓		✓
N	71,161	39,298	66,574	39,366	68,864	36,005
R^2	0.259	0.401	0.298	0.418	0.276	0.383

Notes: $\Delta \log M_t$ 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 $t - 1$ and t. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A11: The Effect of Narrative 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	-0.000142	0.00106	0.000963	0.00173	0.00249
	(0.0007)	(0.0007)	(0.0011)	(0.0014)	(0.0012)	(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 $R_{i,w(t)} = \beta \text{opt}_{it} + \gamma_i + \chi_{j(i),y(i,t)} + \varepsilon_{it}$ 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.

Table A12: Textual Optimism and Optimistic Forecasts, Alternative Measurement

	(1)	(2)	(3)	(4)
	$\text{GuidanceOptExPost}_{i,t+1}$		$\text{GuidanceOptExPostC}_{i,t+1}$	
opt_{it}	0.0354		-0.000169	
	(0.0184)		(0.0049)	
sentiment_{it}		0.0152		-0.000219
		(0.0095)		(0.0025)
N	3,817	3,780	3,754	3,719
R^2	0.173	0.174	0.139	0.141

Notes: The regression model is Equation 15. The outcome in columns 1 and 2 is a binary indicator for *ex post* optimism in guidance, and the outcome in columns 3 and 4 is the difference between the log guidance value and the log realized sales. opt_{it} is the binary measure of narrative optimism, and sentiment_{it} is the underlying continuous measure from which opt_{it} is constructed. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A13: The Effects of All Selected Narratives on Hiring

	(1)	(2)	(3)	(4)
		Outcome is $\Delta \log L_{it}$		
	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	0.0597
			(0.0030)	(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: The first-stage equation for column 4 is described in the main text. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A14: Narrative Optimism is Viral and Associative, Alternative Standard Errors

	(1)	(2)	(3)
	Outcome is opt_{it}		
Own lag, $\text{opt}_{i,t-1}$	0.209	0.214	0.135
	(0.0071)	(0.0080)	(0.0166)
	[0.0214]	[0.0220]	[0.0281]
	{0.0218}	{0.0221}	{0.0273}
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290		
	(0.0578)		
	[0.180]		
	{0.179}		
Real GDP growth, $\Delta \log Y_{t-1}$	0.804		
	(0.2204)		
	[0.635]		
	{0.627}		
Industry lag, $\overline{\text{opt}}_{j(i),t-1}$		0.276	0.207
		(0.0396)	(0.0733)
		[0.0434]	[0.0563]
		{0.0496}	{0.0656}
Industry output growth, $\Delta \log Y_{j(i),t-1}$		0.0560	0.0549
		(0.0309)	(0.0632)
		[0.0328]	[0.0668]
		{0.0428}	{0.0772}
Peer lag, $\overline{\text{opt}}_{p(i),t-1}$			0.0356
			(0.0225)
			[0.0259]
			{0.0329}
Firm FE	✓	✓	✓
Time FE		✓	✓
N	64,948	52,258	8,514
R^2	0.481	0.501	0.501

Notes: This Table replicates the analysis of Table 4 with alternative standard error constructions. Standard errors in parentheses are two-way clustered by firm ID and industry-year; those in square brackets are two-way clustered by firm ID and year; and those in braces are two-way clustered by industry and year. Aggregate, industry, and peer average optimism are averages of the narrative optimism variable over the respective sets of firms. Industry output growth is the log difference in sectoral value-added calculated from BEA data, linked to two-digit NAICS industries.

Table A15: Narrative Sentiment is Viral and Associative

	(1)	(2)	(3)
	Outcome is sentiment$_{it}$		
Own lag, sentiment$_{i,t-1}$	0.259	0.279	0.226
	(0.0091)	(0.0106)	(0.0166)
Aggregate lag, $\overline{\text{sentiment}}_{t-1}$	0.253		
	(0.0519)		
Real GDP growth, $\Delta \log Y_{t-1}$	2.632		
	(0.5305)		
Industry lag, $\overline{\text{sentiment}}_{j(i),t-1}$		0.175	0.108
		(0.0360)	(0.0763)
Industry output growth, $\Delta \log Y_{j(i),t-1}$		0.108	0.142
		(0.0522)	(0.1312)
Peer lag, $\overline{\text{sentiment}}_{p(i),t-1}$			0.0234
			(0.0188)
Firm FE	✓	✓	✓
Time FE		✓	✓
N	63,881	51,555	8,338
R^2	0.568	0.599	0.602

Notes: The regression model is a variant of Equation 18 for column 1, and a variant of Equation 19 for columns 2 and 3, with the continuous variable sentiment$_{it}$ (and averages thereof) substituted for binary optimism. Aggregate, industry, and peer average sentiment are averages of the narrative sentiment variable over the respective sets of firms. Industry output growth is the log difference in sectoral value-added calculated from BEA data, linked to two-digit NAICS industries. In all specifications, we trim the 1% and 99% tails of sentiment$_{it}$. Standard errors are two-way clustered by firm ID and industry-year.

Table A16: Narrative Sentiment is Viral and Associative, Over-Controlling for Past and Future Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Outcome is sentiment_{it}						
Aggregate lag, $\overline{\text{sentiment}}_{t-1}$	0.253	0.385	0.410	0.340			
	(0.0519)	(0.0651)	(0.1103)	(0.1785)			
Ind. lag, $\overline{\text{sentiment}}_{j(i),t-1}$					0.175	0.151	0.213
					(0.0360)	(0.0409)	(0.0654)
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	63,881	48,889	37,643	13,112	51,555	37,643	13,112
R^2	0.568	0.578	0.599	0.640	0.599	0.601	0.642

Notes: The regression model is a variant of Equation 20 for column 1-4, and an analogous variant of industry-level specification for columns 5-7 (*i.e.*, Equation 19 with past and future controls), with the continuous variable sentiment_{it} (and averages thereof) substituted for binary optimism. Columns 1 and 5 are "baseline estimates" corresponding, respectively, with columns 1 and 3 of Table A15. The added control variables are two leads, two lags, and the contemporaneous value of: real GDP growth (columns 2-4), industry-level output growth (columns 3-4 and 6-7), and firm-level TFP growth (columns 4 and 7). In all specifications, we trim the 1% and 99% tails of sentiment_{it}. Standard errors are two-way clustered by firm ID and industry-year.

Table A17: Narrative Optimism and Viral and Associative, 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	0.207	0.214	0.200
	(0.0071)	(0.0072)	(0.0080)	(0.0084)
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290	0.393		
	(0.0578)	(0.0597)		
Real GDP growth, $\Delta \log Y_{t-1}$	0.804	0.672		
	(0.2204)	(0.2153)		
Ind. lag, $\overline{\text{opt}}_{j(i),t-1}$			0.276	0.437
			(0.0396)	(0.0748)
Ind. output growth, $\Delta \log Y_{j(i),t-1}$			0.0560	0.0390
			(0.0309)	(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: 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 3. 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. Standard errors are two-way clustered by firm ID and industry-year.

Table A18: Sensitivity Analysis for the Quantitative Analysis

	Parameters						Results			
	α	γ	ψ	ϵ	ω	$\frac{1}{1-\omega}$	$\hat{c}_Q(0)$	$\hat{c}_Q(1)$	2000-02	2007-09
Baseline	1.0	0.0	0.4	2.6	0.490	1.962	0.048	0.170	0.316	0.181
High ψ	1.0	0.0	2.5	2.6	0.133	1.154	0.020	0.132	0.186	0.106
High γ	1.0	1.0	0.4	2.6	-0.784	0.560	0.006	0.096	0.090	0.052
Empirical Mult.	1.0	0.0	1.154	2.6	0.250	1.333	0.025	0.142	0.215	0.123
Calibrated Mult.	1.0	0.0	0.845	2.6	0.313	1.455	0.029	0.146	0.235	0.134
High ϵ	1.0	0.0	0.208	5.0	0.490	1.962	0.047	0.172	0.317	0.181
Decreasing RtS	0.75	0.0	0.05	2.6	0.490	1.962	0.033	0.132	0.237	0.135

Notes: This table summarizes the quantitative results under alternative calibrations of the macroeconomic parameters, which we report along side their implied complementarity ω and demand multiplier $\frac{1}{1-\omega}$. We report four statistics as the "results" in the last four columns. The first two are the fraction of output variance explained statically and at an eight-year horizon by optimism. The second two are the fraction of output losses in the 2000-02 downturn and 2007-09 downturn explained by fluctuations in narrative optimism. Baseline corresponds to our main calibration. High ψ increases the inverse Frisch elasticity to 2.5, or decreases the Frisch elasticity to 0.4. High γ increases the curvature of consumption utility (indexing income effects in labor supply) from 0.0 to 1.0. Empirical Multiplier adjusts ψ to match an output multiplier in line with estimates from Flynn, Patterson, and Sturm (2022). Calibrated multiplier adjusts ψ to match our own calculation of the multiplier in our setting in Appendix F.3. High ε increases the elasticity of substitution from 2.6 to 5.0, with ψ adjusting to hold fixed the multiplier. Decreasing RtS reduces the returns-to-scale parameter α from 1.0 to 0.75, with ψ adjusting to hold fixed the multiplier.

Table A19: An Empirical Test for Cycles and Chaos

	(1) Outcome is opt_{it}
α: Constant	-0.051
	(0.244)
α_1: $\text{opt}_{i,t-1}$	0.655
	(0.062)
β_1: $\text{opt}_{i,t-1} \cdot \overline{\text{opt}}_{i,t-1}$	0.052
	(1.021)
β_2: $(1 - \text{opt}_{i,t-1}) \cdot \overline{\text{opt}}_{i,t-1}$	0.952
	(1.006)
τ: $(\overline{\text{opt}}_{i,t-1})^2$	-0.062
	(1.034)
η: Logistic parameter	1.443
	(0.698)
Firm FE?	✓
N	67,648
R^2	0.480

Notes: The regression model is Equation 203. η is a function of the regression coefficients defined in Equation 204, and interpretable in the model of cycles and chaos in Appendix B.7. Standard errors are two-way clustered by firm ID and industry-year. The standard error for η is calculated using the delta method.

Table A20: 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	—

Table A21: The Effect of Optimism on Hiring, CEO Change Strategy

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0253	0.0404	0.0362	0.0253
	(0.0131)	(0.0131)	(0.0132)	(0.0029)
$\text{opt}_{it} \times \text{ChangeCEO}_{it}$				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	905	36,953
R^2	0.243	0.375	0.375	0.134

Notes: The regression model is Equation 240 for columns 1-3, and Equation 241 for column 4. The outcome is the log change in firms' employment. opt_{it} is a binary indicator for the optimistic narrative, defined in Section 3.2. ChangeCEO_{it} is a binary indicator for whether firm i changed CEO in fiscal year t due to death, illness, personal issues or voluntary retirement. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A22: The Virality of Optimism, CEO Change Strategy

	(1)	(2)	(3)	(4)
		Outcome is opt_{it}		
	OLS	IV	OLS	IV
Industry lag, $\overline{\text{opt}}_{j(i),t-1}$	0.275	0.260	0.195	0.272
	(0.0407)	(0.2035)	(0.0760)	(0.5270)
Peer lag, $\overline{\text{opt}}_{p(i),t-1}$			0.0437	0.129
			(0.0236)	(0.1677)
Firm FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Industry output growth, $\Delta \log Y_{j(i),t-1}$	✓	✓	✓	✓
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: The IV strategies instrument the industry and/or peer lag with the CEO-change variation in those averages. Standard errors are two-way clustered by firm ID and industry-year.

Table A23: Narrative Optimism Predicts Hiring, Conditional on Measured Beliefs

	(1)	(2)	(3)	(4)
		Outcome is $\Delta \log L_{it}$		
opt_{it}	0.0355	0.0232	0.0311	0.0203
	(0.0030)	(0.0129)	(0.0068)	(0.0164)
$\text{ForecastGrowthSales}_{it}$		0.157		
		(0.0329)		
$\text{ForecastGrowthCapx}_{it}$			0.0564	
			(0.0062)	
$\text{ForecastGrowthEps}_{it}$				0.000961
				(0.0104)
Ind.-by-time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
N	71,161	2,908	7,312	1,290
R^2	0.259	0.506	0.425	0.638

Notes: opt_{it} is textual optimism from the 10-K for fiscal year t. $\text{ForecastGrowthZ}_{it}$ is defined in the text as the log difference between manager guidance about statistic Z, for fiscal year t, with last fiscal year's realized value. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A24: Narrative Optimism Predicts Investment, Conditional on Measured Beliefs

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log K_{it}$			
opt_{it}	0.0370	0.0238	0.0251	0.00503
	(0.0034)	(0.0177)	(0.0072)	(0.0193)
$\text{ForecastGrowthSales}_{it}$		0.172		
		(0.0423)		
$\text{ForecastGrowthCapx}_{it}$			0.0943	
			(0.0079)	
$\text{ForecastGrowthEps}_{it}$				-0.0147
				(0.0102)
Ind.-by-time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
N	68,864	2,748	7,048	1,245
R^2	0.276	0.496	0.472	0.661

Notes: This table is identical to Table A23, but has net capital investment ΔK_{it} as the outcome. opt_{it} is textual optimism from the 10-K for fiscal year t. $\text{ForecastGrowthZ}_{it}$ is defined in the text as the log difference between manager guidance about statistic Z, for fiscal year t, with last fiscal year's realized value. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A25: State-Dependent Effects of Sentiment on Hiring

	(1)	(2)	(3)
	Outcome is $\Delta \log L_{it}$		
sentiment_{it}	0.0218	0.0172	0.0130
	(0.0017)	(0.0018)	(0.0020)
$\text{sentiment}_{i,t-1}$	0.00605	0.00877	0.00830
	(0.0015)	(0.0016)	(0.0016)
$\text{sentiment}_{it} \times \text{sentiment}_{i,t-1}$	-0.00497	-0.00501	-0.00404
	(0.0008)	(0.0008)	(0.0008)
N	63,302	35,768	31,071
R^2	0.257	0.394	0.416
Ind.-by-time FE	✓	✓	✓
Firm FE	✓	✓	✓
Lag labor		✓	✓
Current and lag TFP		✓	✓
Log Book to Market			✓
Stock Return			✓
Leverage			✓

Notes: This table reports estimates from Equation 245 with our baseline sets of controls. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A26: Optimism is Viral and Associative, Granular IV Strategy

	(1)	(2)	(3)	(4)	(5)	(6)
			Outcome is opt_{it}			
	OLS	OLS	IV	OLS	OLS	IV
Own lag, $\text{opt}_{i,t-1}$	0.212	0.213	0.210	0.219	0.220	0.219
	(0.0071)	(0.0071)	(0.0073)	(0.0080)	(0.0081)	(0.0081)
Agg. sales-wt. lag, $\overline{\text{opt}}^{sw}_{t-1}$	0.0847		0.308			
	(0.0421)		(0.1044)			
Real GDP growth, $\Delta \log Y_{t-1}$	1.058	1.104	0.768			
	(0.2205)	(0.2110)	(0.2607)			
Agg. sales-wt. granular lag, $\overline{\text{opt}}^{g,sw}_{t-1}$		0.150				
		(0.0506)				
Ind. sales-wt. lag, $\overline{\text{opt}}^{sw}_{j(i),t-1}$				0.0728		0.0195
				(0.0209)		(0.0459)
Ind. output growth, $\Delta \log Y_{j(i),t-1}$				0.0851	0.0903	0.0886
				(0.0325)	(0.0336)	(0.0333)
Ind. sales-wt. granular lag, $\overline{\text{opt}}^{g,sw}_{j(i),t-1}$					0.00913	
					(0.0216)	
Firm FE	✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓
N	64,948	64,948	64,948	52,258	50,842	50,842
R^2	0.481	0.481	0.049	0.500	0.503	0.051
First-stage F	—	—	99.1	—	—	262.3

Notes: This table estimates Equations 18 and 19, respectively modeling the spread of optimism at the aggregate and industry level, using granular identification of spillovers (virality). $\overline{\text{opt}}^{sw}_{t-1}$ and $\overline{\text{opt}}^{sw}_{j(i),t-1}$ are sales-weighted averages of aggregate and industry optimism, respectively. $\overline{\text{opt}}^{g,sw}_{t-1}$ and $\overline{\text{opt}}^{g,sw}_{j(i),t-1}$ are (lagged) sales-weighted averages of the *non-fundamentally-predictable* components of firm-level optimism in the aggregate and in the industry, respectively, as explained in Appendix E.4. In columns 3 and 6, we use the granular variables as instruments for the raw sales-weighted averages. Standard errors are two-way clustered by firm ID and industry-year.

Table A27: Multiplier Calibrations via Under-Controlled Regressions of Hiring on Optimism

	(1)	(2)	(3)
	Outcome is ΔL_{it}		
opt_{it}	0.0355	0.0516	0.0486
	(0.0030)	(0.0034)	(0.0033)
Complementarity ω	—	0.313	0.270
Multiplier $\frac{1}{1-\omega}$	—	1.455	1.370
Industry-by-time FE	✓		
Firm FE	✓	✓	✓
Current and lagged adjusted TFP		✓	
Current and lagged unadjusted TFP			✓
N	71,161	65,508	65,508
R^2	0.259	0.207	0.216

Notes: The regression models are introduced in Appendix F.3. The first column replicates Column 1 of Table 1. The second two columns remove the industry-by-time FE and control for the contemporaneous and lagged value of seasonally adjusted log TFP, respectively with and without capacity utilization adjustment, as reported by the updated data series of Fernald (2014). The sample size is lower in columns 2 and 3 due to the band-pass filtering being impossible for the last part of the sample. The remaining rows give the implied complementarity ω and demand multiplier $\frac{1}{1-\omega}$, by comparing the coefficients with that of column 1 and applying the formula in Equation 271. Standard errors are double-clustered by industry-year and firm ID.