Contagious Business Cycles

Joel P. Flynn* Karthik A. Sastry†
Yale Princeton

July 24, 2025

Abstract

We study the implications of contagious beliefs for business cycles. We introduce a real business cycle model in which heterogeneous firms dynamically switch between models of the world. Models spread based on both their consistency with data, as in models of learning, and their prevalence, as in models of social contagion. The spread of models generates endogenously persistent belief-driven fluctuations. If contagion is sufficiently strong, models can "go viral" and induce hysteresis in the model's unique equilibrium. To take this framework to the data, we adopt a "micro-to-macro" approach in which we measure firms' models using the sentiment of their language in regulatory filings. We find that firms accelerate hiring when they are optimistic, even though this sentiment does not positively predict future firm fundamentals or performance. Moreover, sentiment spreads contagiously at the aggregate and industry levels. Mapping our microeconomic estimates into the model, we find that contagious beliefs account for one fifth of business cycle fluctuations in aggregate output.

*Yale University Department of Economics, 30 Hillhouse Avenue, New Haven, CT, 06511. Email: joel.flynn@yale.edu.

†Princeton University Department of Economics, 192A Julis Romo Rabinowitz Building, Princeton, NJ, 08544, Email: ksastry@princeton.edu.

This paper was formerly circulated under the title: "The Macroeconomics of Narratives." We are grateful to Daron Acemoglu, Mark Aguiar, Peter Andre, George-Marios Angeletos, Susanto Basu, Ricardo Caballero, Andrew Caplin, John Coglianese, Roberto Corrao, Xavier Gabaix, Alexandre Kohlhas, Chen Lian, Karel Mertens, Stephen Morris, Christina Patterson, Charlie Rafkin, Robert Shiller, Andrei Shleifer, David Thesmar, Aleh Tsyvinski, Adrien Verdelhan, Mirko Wiederholt, Christian Wolf, and seminar participants at Princeton, HBS, Northwestern, LSE, Oxford, Chicago Booth, Yale, Wharton, Boston U., Boston College, NYU, Maryland, Yale SOM, the NBER Summer Institute Behavioral Macroeconomics Meeting, ETH Zurich, Boston College, the VEAMS, the 2024 ASSA Meeting, the Leibniz Institute SAFE, the Columbia University Narratives Workshop, and MIT. First posted version: June 19, 2022.

1 Introduction

One of the oldest ideas in macroeconomics is that fluctuations in beliefs drive business cycles. Pigou (1923) wrote that the "varying expectations of business men ..., and not anything else, constitute the immediate and direct causes or antecedents of industrial fluctuations." Keynes (1936), despite his vehement disagreement with Pigou on most other matters, broadly concurred that the mercurial moods of managers—whose "spontaneous optimism" might just as suddenly give way to fear and panic—drove booms and busts.

Of course, such assertions beg the question: what drives belief fluctuations? The classics, while not fully precise about the details, suggest an important role for human impulses that are divorced from current economic fundamentals: the "psychological causes" of Pigou (1923) and "animal spirits" of Keynes (1936). A long tradition in macroeconomics has formalized these ideas in various ways: for example, with noise shocks that directly sway aggregate beliefs (*e.g.,* Lorenzoni, 2009; Angeletos and La'O, 2010, 2013; Wiederholt, 2015; Benhabib, Wang, and Wen, 2015; Schaal and Taschereau-Dumouchel, 2023; Guerreiro, 2023), news about future productivity (*e.g.,* Jaimovich and Rebelo, 2009; Barsky and Sims, 2012; Blanchard, L'Huillier, and Lorenzoni, 2013), or "sunspots" that allow agents to coordinate on one of many equilibria (e.g., Boldrin and Woodford, 1990; Benhabib and Farmer, 1994).

A different intellectual tradition has focused on the related question of how beliefs spread across people. One important hypothesis is that economic attitudes are *contagious* like a virus (Carroll, 2001; Burnside, Eichenbaum, and Rebelo, 2016; Carroll and Wang, 2022; Jamilov, Kohlhas, Talavera, and Zhang, 2024). Taking this argument further, Shiller (2020) puts contagion at the core of his theory of "Narrative Economics." Moreover, he argues that contagious economic attitudes are best detected in *language* rather than quantitative statistics, since the former better reveal emotional states—echoing Keynes' (1936) assertion that the animal spirits that drive decisions cannot be understood by merely considering "mathematical expectations."

This paper introduces a new framework to understand how contagious beliefs shape business cycles. Our contribution is to "study Keynes through the lens of Shiller": that is, to revisit belief-driven business cycles by integrating a macroeconomic model in which beliefs can spread contagiously with microeconomic data that connect the language of economic agents with their decisions.

To do this, we first develop a real business cycle model in which heterogeneous beliefs can gain or lose prevalence based on their past prevalence, consistent with contagion, or based on past economic outcomes, consistent with standard notions of learning. We characterize how belief-driven fluctuations are jointly shaped by contagiousness and other sources of strategic complementarity. Empirically, we introduce a new method to discipline models of belief-driven cycles using panel data on the sentiment of firm managers' language. At the firm-level, we find that textual sentiment leads to short-run hiring booms, but does not predict future productivity growth, consistent with the classical notion of an "animal spirit" and inconsistent with the notion that sentiment measures news. Moreover, sentiment spreads contagiously at the aggregate and industry levels. Quantitatively, using these microeconomic measurements to discipline our macroeconomic model, we find that fluctuations in sentiment account for about 20% of the US business cycle. Endogenous persistence via contagion explains much of these effects. Taken together, our results suggest that belief contagion underlies first-order features of business cycles.

Model. In order to make decisions, agents forecast economic fundamentals using one of a finite set of probability models. At an individual level, model adoption is a Markov process that depends on one's own past model, the relative prevalence of models in the population, and endogenous economic outcomes. The second feature incorporates the possibility of contagion emphasized by Carroll (2001) and Shiller (2017). The third feature allows agents to associate certain macroeconomic outcomes with certain models, consistent with theories of memory (Kahana, 2012) and learning (*e.g.,* Eusepi and Preston, 2011; Molavi, 2019).

We embed these dynamics of heterogeneous beliefs in a real business cycle model. The agents are heterogeneous firms, the fundamental is productivity, and the two models are *optimism* and *pessimism* about productivity. The remaining microfoundations of the consumption, labor supply, and production blocks of the model are intentionally standard, as in the textbook models of Woodford (2003b) or Galí (2008). Our goal is to study how contagious business cycles can manifest in this simple framework, intentionally abstracting from richer aspects of firm and household decision-making to isolate the role of contagious beliefs when they are the only state variable.

Our solution concept is a variant of rational expectations equilibrium in which agents "agree to disagree" while making correct inferences regarding equilibrium outcomes *conditional* on their misspecified models about fundamentals. That is, in equilibrium, models of productivity are also models of GDP, employment, the prevalence of optimism, and even how these quantities respond to shocks.

Theoretical Results. We theoretically study how contagious beliefs co-evolve with the business cycle. We first establish that there is a unique equilibrium in which aggregate output depends both on aggregate productivity and the fraction of optimists in the population. We refer to the latter effect as the *non-fundamental component* of aggregate output, because it is driven by the agents' choice of models rather than any fundamental change in productivity. We decompose this component into a partial-equilibrium effect of optimism on firms' expansion as well as a general-equilibrium multiplier driven by strategic complementarities: even if a firm is pessimistic, the presence of other, more optimistic firms causes them to produce more. The equilibrium effect of optimism on output is smaller if firms have access to more precise information, as this leads them to rely less on priors to form their beliefs. However, because of the multiplier, the power of the truth to stop false models of the world is weakest exactly when those models are strongest.

We next describe the dynamics of the fraction of optimists in the population ("optimism"), the key new state variable in our economy. For a fixed level of aggregate productivity, there always exists a steady-state level of optimism, but there may be multiple. We provide a necessary and sufficient condition for a particularly extreme type of steady-state multiplicity: if models are sufficiently contagious and/or the multiplier is sufficiently large, then unanimous optimism and unanimous pessimism are both stable steady states. That is, in the economy's unique equilibrium, contagious optimism or contagious pessimism can "go viral" or die out entirely, depending on their initial prevalence.

Contagious business cycles have three key properties. First, if contagiousness and the multiplier are sufficiently high, the economy can feature *hysteresis*: history determines whether the economy is optimistic and high-output or pessimistic and low-output. This prediction may remind of the sunspot-driven fluctuations in multiple-equilibrium models (Azariadis, 1981; Cooper and John, 1988; Benhabib and Farmer, 1994), but it differs in several crucial respects: it obtains in a *unique*-equilibrium model, it allows the analyst to make history-based predictions for the economy's regime switches, and it is underpinned by dynamic complementarity through belief dynamics rather than static complementarity through, for example, external returns to scale. Second, productivity shocks have *endogenously persistent* effects on output because of belief evolution. This can generate hump-shaped impulse responses to fully transitory shocks. In cases consistent with hysteresis, transitory shocks can even have permanent effects because a new model takes hold. This is an important difference relative to models of dispersed information (see *e.g.,* Woodford, 2003a; Lorenzoni, 2009; Angeletos and La'O, 2010), in which one-time shocks can have persistent effects only if fundamentals are themselves persistent. Third, when hit by repeated shocks, the economy can experience *boom-bust cycles* because of fluctuations between high and low optimism. In the hysteresis case, these can be slow oscillations between periods of stable extreme pessimism ("lost decades") and stable extreme optimism ("roaring decades").

We finally show how the model can be identified using data on the choices of optimistic versus pessimistic firms, the updating process of firms' models, and standard parameters that control equilibrium effects. We subsequently pursue this "micro-to-macro" approach.

Empirical Results. Motivated by both Shiller's (2017) argument that language reveals the most relevant information on individuals' economic attitudes and a burgeoning literature pioneering textual analysis as a means for understanding corporate decisions (Hassan, Hollander, Kalyani, van Lent, Schwedeler, and Tahoun, 2024), we measure firms' optimism using the text of US public firms' 10-K regulatory filings and their earnings conference calls. Specifically, we apply the standard method of Loughran and McDonald (2011) to measure positive versus negative sentiment in financial language. We use this to construct an empirical proxy for whether firms hold the optimistic or pessimistic model in the theory, noting that the theoretical predictions apply equally to contagious optimism about aggregate or firm-level conditions. We combine all these measures with Compustat data on firms' decisions and financial performance as well as data from IBES on managers' and equity analysts' beliefs about future firm performance.

We first estimate how optimism affects firm decisions. Our model implies that the partial-equilibrium effect of optimism on hiring is identified in a firm-level panel regression that controls for firm and time fixed effects, which sweep out the correlation of firms' optimism with aggregate fundamentals and aggregate sentiment. Implementing this strategy, we find that optimistic firms hire 3.6 percentage points more per year than their pessimistic counterparts. This effect is quantitatively robust to controlling for canonical measures of firm-level fundamentals: productivity, leverage, stock returns, and Tobin's q. To further isolate a plausibly exogenous shifter of mental models that is independent of news, we leverage plausibly exogenous changes in firm CEOs as natural experiments. We find that changes in optimism triggered by the arrival of a new CEO have, if anything, even greater effects on hiring.

We finally show that textual optimism predicts hiring, even conditional on standard proxies for the beliefs of firms and equity analysts used in the literature (e.g., Gennaioli, Ma, and Shleifer, 2016; Bordalo, Gennaioli, Shleifer, and Terry, 2021; Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer, 2024). This finding supports the idea that managers liberally use non-quantitative "soft information" when making decisions (Liberti and Petersen, 2019). And it is moreover consistent with the hypothesis of Shiller (2020)—and Keynes (1936) himself—that measuring animal spirits requires going beyond "mathematical expectations."

We next show two additional results that validate our interpretation of optimism as capturing a non-fundamental "animal spirit" and reject an alternative interpretation that the measure merely picks up news about future fundamentals. First, in firm-level local projection regressions, we find that optimism predicts no further TFP growth and *negatively* predicts future stock returns and profitability. That is, firms that are currently optimistic and accelerating hiring become no more productive, continue to expand operations, and ultimately do *worse*, not better, in the near future. This is inconsistent with a model in which measured optimism reflects news about growing productivity or demand: even if signals are noisy, the news would be true on average. But it is consistent with our model in which individual managers have tendencies toward optimism (and pessimism) that can be disconnected from fundamentals. Second, using managerial guidance data, we show that managers predictably overestimate firm performance after writing optimistic reports or giving optimistic earnings calls. This is inconsistent with a view that optimistic (pessimistic) managers are interpreting positive (negative) news using Bayes' rule, but consistent with our framework of model dynamics.

We finally estimate how optimism spreads across firms. 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—that is, optimism is contagious and associative. We also find evidence of contagiousness and associativeness at the industry level when we non-parametrically control for aggregate conditions with time fixed-effects. Moreover, both these aggregate and industry-level results are robust to controlling for future economic conditions. This finding is inconsistent with the key threat to our interpretation: that aggregate optimism drives future optimism through its correlation with omitted positive news about economic conditions. To further test the validity of our interpretation, we construct a granular instrumental variable (Gabaix and Koijen, 2020) for aggregate optimism based on idiosyncratic shocks to large firms. We find similar results using this approach.

Quantification. Finally, we employ these empirical results—disciplining both the partial equilibrium effects of optimism and its dynamic spread—in the model-derived identification scheme, along with a standard calibration for preference and technology parameters. Quantitatively, we find that contagious optimism contributes significantly to the US business cycle. Decomposing aggregate output into the components attributable to optimism versus fundamentals, we find that measured aggregate movements in optimism account for 32% of output loss during the early 2000s recession and 18% during the Great Recession. More systematically, fluctuations in optimism account for 19% of output variance as well as 33% of the short-run (one-year) and 79% of the medium-run (two-year) autocovariance in output. Thus, belief dynamics lead to strong endogenous persistence: the model generates persistent business cycles even with close to i.i.d. shocks.[1] This represents an important difference between our model and those of noise shocks or dispersed information (see *e.g.,* Woodford, 2003a; Lorenzoni, 2009; Angeletos and La'O, 2010), which require persistent exogenous shocks to explain business cycles. However, while internal propagation is strong enough to underlie this important feature of fluctuations, it is insufficient to generate hysteresis.

[1]Normatively, we show that contagious 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.

In an extension, we study an enriched model that allows multiple latent topics to form a basis for overall optimism. We return to the data to measure more granular topics of firms' language using two methods: (i) identifying the nine *Perennial Economic Narratives* introduced by Shiller (2020) and (ii) estimating an unsupervised Latent Dirichlet Allocation (LDA) model (Blei, Ng, and Jordan, 2003) that allows the text to speak flexibly about what firms are discussing. We find that the interaction of many jointly evolving and highly contagious topics—that can individually feature hysteresis—nevertheless underlies stable fluctuations in emergent aggregate optimism and output. These findings demonstrate the applicability of our methods to further studying how diffuse and diverse topics of firms' discussion propagate through the economy and shape business cycles.

Related Literature. The most closely related work is from a literature studying epidemiological contagion in the macroeconomy (Carroll, 2001; Burnside et al., 2016; Carroll and Wang, 2022; Jamilov et al., 2024) and the survival dynamics of competing models among economic agents (*e.g.*, Brock and Hommes, 1997; Molavi, Tahbaz-Salehi, and Vedolin, 2021; Bohren and Hauser, 2021). Theoretically, our work presents a new framework for understanding how competing models shape equilibrium business cycles; empirically and quantitatively, our work introduces a new and distinct method to measure economic agents' models, their decision relevance, and their spread. An important insight of our analysis is that, in equilibrium, business cycles and model choice are deeply intertwined: economic outcomes determine what models catch on, and the popularity of different models shapes economic outcomes.

Our work relates to a large literature on belief-driven business cycles. Some studies postulate that shocks directly to aggregate beliefs cause fluctuations (*e.g.*, Lorenzoni, 2009; Angeletos and La'O, 2010, 2013; Angeletos, Collard, and Dellas, 2018; Christiano, Ilut, Motto, and Rostagno, 2008; Benhabib et al., 2015; Nimark, 2014; Chahrour, Nimark, and Pitschner, 2021). Others focus on the interaction of other shocks with fixed belief differences (*e.g.*, Caballero and Simsek, 2020; Guerreiro, 2023), the tendency of agents to over-extrapolate (*e.g.*, Bordalo et al., 2021; Bianchi, Ilut, and Saijo, 2024), or the effects of slow and/or misspecified learning (*e.g.*, Marcet and Sargent, 1989a,b; Eusepi and Preston, 2011; Kozlowski, Veldkamp, and Venkateswaran, 2020). Our work can be understood as micro-founding both belief dynamics and belief disagreement through the transmission of models, as well as providing novel evidence for this mechanism.

Our analysis finally relates to a literature studying firms' language and outcomes (*e.g.*, Loughran and McDonald, 2011; Hassan et al., 2024). A methodological contribution of this paper is showing how to interpret text-derived objects in an equilibrium macroeconomic model that can be used to study counterfactuals.

2 Model

We first describe our framework, a real business cycle model with contagious beliefs.

2.1 Beliefs and Fundamentals

Time is discrete, indexed by $t \in \mathbb{N}$, and there is an exogenous macroeconomic state $\theta_t \in \Theta \subseteq \mathbb{R}$. We will interpret θ_t as aggregate productivity and assume that it follows:

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

where $\nu_t \sim N(0,1)$ is an i.i.d. shock.

Each agent uses a probability model to forecast θ_t. There is a finite set of possible such models. For our main analysis, we will suppose that there are two competing models that differ only in the perceived mean of productivity: an optimistic one under which $\mu = \mu_O$ and a pessimistic one under which $\mu = \mu_P$, where $\mu_O > \mu_P$. The *true* distribution of the fundamental need not coincide with either model. Firms either believe the optimistic model or the pessimistic model. Hence, each agent $i \in [0,1]$ has a prior belief regarding the fundamental that can be described as:

$$N\left(\left[\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right](1-\rho) + \rho \log \theta_{t-1}, \sigma^2\right) \tag{2}$$

where $\lambda_{it} = 1$ corresponds to optimism and $\lambda_{it} = 0$ corresponds to pessimism.

In equilibrium, as we will see shortly, agents' prior beliefs about θ_t will also govern their prior beliefs about endogenous objects (*e.g.*, GDP, wages, and what beliefs will catch on in the future). Thus, models by our definition will convey multidimensional information about the economy and how it will evolve—but, crucially for our purposes, in a way that can be made consistent with rational expectations equilibrium.

We focus on a setting with two models about the mean of fundamentals for parsimony. However, our analysis extends to richer spaces of models. In Appendix B.4, we consider a continuum of models. In Appendix B.5, we consider models of idiosyncratic (as opposed to aggregate) fundamentals. In Appendix B.6, we consider multiple models that differ in the mean, persistence, and volatility of productivity. Finally, in Section 6.4, we allow optimism to be driven by a large set of underlying and more specific topics.

2.2 Belief Dynamics

If individuals were each endowed with one unchanged model forever, ours would be a model of heterogeneous priors in the tradition of Miller (1977) and Harrison and Kreps (1978). Our goal instead is to describe how models gain or lose prevalence over time. We suppose that this can happen in two distinct and potentially complementary ways.

The first channel is *contagion*: models spread as a function of their prevalence. The contagion hypothesis is rooted in a large social science literature studying how social interaction affects beliefs (see Carroll and Wang, 2022, for a review). Contagion is also central to the "Narrative Economics" of Shiller (2020), who argues for its relevance in macroeconomics:

> We need to incorporate the contagion of narratives into economic theory. [...] If we do not understand the epidemics of popular narratives, we do not fully understand changes in the economy and in economic behavior.

The second channel is *association*: models spread when they better describe the world. This is the standard view in economic theories in which belief dynamics are governed by learning (*e.g.,* Eusepi and Preston, 2011; Molavi, 2019) or by associative memory (Kahana, 2012). Failing to account for association may lead an observer to spuriously attribute the spread of a model to contagion: the observation that people become more optimistic when the economy is doing well does not imply the existence of contagion.

Toward formalizing these notions, we summarize the prevalence of models by their cross-sectional distribution in the population. In our economy with optimists and pessimists, we need only to keep track of the fraction of optimists: $Q_t = \int_{[0,1]} \lambda_{it} \, \mathrm{d}i \in [0, 1]$. We also define Y_t as aggregate output, the endogenous outcome that will govern association.

At the individual level, adherence to a model follows a Markov process. We describe this process via a probability that optimists remain optimistic, P_O, and the probability that pessimists become optimistic, P_P. Both probabilities depend on aggregate output Y_t, the fraction of optimists in the population Q_t, and an aggregate *model* shock to how agents update $\varepsilon_t \sim G$. The first two features respectively capture associativeness and contagiousness. The shock captures shifts in models that are unrelated to economic conditions. Hence, the fraction of optimists evolves according to:

$$Q_{t+1} = Q_t P_O(\log Y_t, Q_t, \varepsilon_t) + (1 - Q_t) P_P(\log Y_t, Q_t, \varepsilon_t) \tag{3}$$

We assume that P_O and P_P are increasing in their first two arguments, continuous and almost everywhere differentiable. We do not take a stand on the details of what makes a model contagious or associative, noting only that microfoundations based on communication between people (Burnside et al., 2016) and retrieval of memories (Kahana, 2012; Bordalo, Gennaioli, and Shleifer, 2020) can underlie contagion and association, respectively. Our approach will instead be to take as given P_O and P_P and later empirically estimate them. In Appendix B.3, we discuss how our approach differs from Bayesian learning.

2.3 Technology and Preferences

The consumption, production, and labor supply blocks of the model are intentionally standard, following the textbook structure of Woodford (2003b) and Galí (2008). This approach intentionally abstracts from many realistic frictions and forces in the economy, consistent with our goal of understanding how contagious beliefs affect business cycles under the most simple and standard microfoundations. It is simple to see how one could extend our analysis to richer models and we leave that to future research.

There is a continuum of monopolistically competitive intermediate goods firms of unit measure, indexed by i, and uniformly distributed on the interval $[0,1]$. 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:

$$x_{it} = \theta_{it} L_{it}^{\alpha} \tag{4}$$

where $\alpha \in (0,1]$ is the return-to-scale in production and θ_{it} is the firm's productivity.

Models correspond to beliefs about the common component of firms' productivity. Concretely, 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 $\tilde{\theta}_{it}$:

$$\theta_{it} = \tilde{\theta}_{it} \gamma_i \theta_t \tag{5}$$

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 (i.i.d.) across firms and time. We assume that firms can observe all previous macroeconomic outcomes. Firms receive idiosyncratic Gaussian signals about $\log \theta_t$ with noise $e_{it} \sim N(0, \sigma_e^2)$ that is i.i.d. across firms and time: $s_{it} = \log \theta_t + e_{it}$. We define the signal-to-noise ratio as $\kappa = 1/(1 + \frac{\sigma_e^2}{\sigma_{\tilde{\theta}}^2})$, which indexes how much firms update their beliefs about aggregate productivity upon receiving the signal s_{it}. Importantly, by allowing for these signals, our model nests the case in which beliefs are fully driven by models ($\kappa = 0$) as well as the case in which models have no bearing on the posterior beliefs that are relevant for decisions ($\kappa = 1$).

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

$$Y_t = \left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}} \, \mathrm{d}i \right)^{\frac{\epsilon}{\epsilon-1}} \tag{6}$$

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:

$$\mathcal{U}\left(\{C_t, \{L_{it}\}_{i\in[0,1]}\}_{t\in\mathbb{N}}\right) = \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\left(\frac{C_t^{1-\gamma}}{1-\gamma} - \int_{[0,1]}\frac{L_{it}^{1+\psi}}{1+\psi}\,\mathrm{d}i\right)\right] \tag{7}$$

where household expectations are arbitrary (and potentially correct), $\gamma \in \mathbb{R}_+$ indexes the size of income effects in the household's supply of labor, and $\psi \in \mathbb{R}_+$ is the inverse Frisch labor supply elasticity to each firm, and $\beta \in (0,1)$ is the household's discount factor.

Finally, we define the composite parameter:

$$\omega = \frac{\frac{1}{\epsilon} - \gamma}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \tag{8}$$

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. This condition holds, as we will later see, under standard calibrations of the relevant parameters.

2.4 Equilibrium

We study a rational expectations equilibrium. All agents optimize, firms form their expectations by combining all available information with their models, models spread dynamically as described earlier, and all markets clear.

Definition 1 (Model-Based Rational Expectations Equilibrium)**.** *An equilibrium is a path:*

$$\mathcal{E} = \left\{Y_t, C_t, Q_t, \theta_t, \varepsilon_t, \{L_{it}, x_{it}, p_{it}, w_{it}, \lambda_{it}, s_{it}, \tilde{\theta}_{it}\}_{i\in[0,1]}\right\}_{t\in\mathbb{N}} \tag{9}$$

1. *Models weights λ_{it} and the fraction of optimists Q_t follow a Markov process consistent with Equation 3.*
2. *Firms' production x_{it} maximizes expected profits under the household's stochastic discount factor given their model weights λ_{it}, signals s_{it}, and knowledge of $\mathcal{E}$.*
3. *Consumption C_t and labor supply $\{L_{it}\}_{i\in[0,1]}$ are consistent with household expected utility maximization.*
4. *All markets clear.*

As we will shortly see, in a model-based REE, agents' models endogenously take on the role of forecasting equilibrium outcomes and thus embed this multi-dimensional information.

3 Macroeconomic Dynamics with Contagious Models

We now study the equilibrium dynamics of models and output in our framework. We find that models induce non-fundamental fluctuations in the economy and have the potential to generate endogenous persistence and hysteresis. Moreover, we show how to use firm-level panel data to identify the model's parameters and test its predictions.

3.1 Characterizing Equilibrium Dynamics

We solve for equilibrium by first characterizing the choices of intermediate goods firms, the key agents whose models affect hiring and production. These firms maximize expected profits priced by the representative household, or $\mathbb{E}_{it}[C_t^{-\gamma}(p_{it}x_{it} - w_{it}L_{it})]$. Each firm takes as given the demand generated by the competitive final goods firm, $p_{it} = Y_t^{\frac{1}{\epsilon}} x_{it}^{-\frac{1}{\epsilon}}$. Since firms are monopsonists, they internalize movements in the wage that are pinned down by the household's labor supply curve $L_{it}^{\psi} = w_{it}C_t^{-\gamma}$. Finally, the firm's labor requirement to produce x_{it} units of output is $L_{it} = \theta_{it}^{-\frac{1}{\alpha}} x_{it}^{\frac{1}{\alpha}}$. Putting these pieces together (along with market clearing $C_t = Y_t$), we obtain that each intermediate goods firm solves:

$$\max_{x_{it}} \mathbb{E}_{it}\left[Y_t^{-\gamma}\left(Y_t^{\frac{1}{\epsilon}} x_{it}^{1-\frac{1}{\epsilon}} - Y_t^{\gamma}\theta_{it}^{-\frac{1+\psi}{\alpha}} x_{it}^{\frac{1+\psi}{\alpha}}\right)\right] \tag{10}$$

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

$$\left(1 - \frac{1}{\epsilon}\right)\mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right] x_{it}^{-\frac{1}{\epsilon}} = \frac{1+\psi}{\alpha}\mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right] x_{it}^{\frac{1+\psi-\alpha}{\alpha}} \tag{11}$$

where the left-hand side is the marginal expected revenue from expanding production and the right-hand side is the marginal expected cost of this expansion. Models affect the expected marginal costs of production, via the expectation of productivity, and the expected marginal benefits of production, via the expectation of aggregate output (which itself matters due to aggregate demand externalities, asset pricing forces, and wage pressure). In equilibrium, these beliefs about output will also depend on the models held by *other* firms.

Substituting this into the final goods production function, any equilibrium conditional on any process of model evolution solves the following functional fixed-point equation:

$$\begin{aligned}\log Y_t = \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\Bigg[\exp\Bigg\{\frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg(\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\ - \log\mathbb{E}_{it}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right] + \log\mathbb{E}_{it}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)\log Y_t\right\}\right]\Bigg)\Bigg\}\Bigg]\end{aligned} \tag{12}$$

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

By employing a functional guess-and-verify argument, we characterize equilibrium output and how it depends on both productivity and the prevalence of optimism:

Theorem 1 (Equilibrium Characterization)**.** *There exists a unique quasi-loglinear equilibrium:*

$$\log Y(\log\theta_t, \log\theta_{t-1}, Q_t) = a_0 + a_1 \log\theta_t + a_2 \log\theta_{t-1} + f(Q_t) \tag{13}$$

Moreover, in the unique quasi-loglinear equilibrium, we have that:

$$f(Q_t) = \frac{1}{1-\omega}\frac{\epsilon}{\epsilon-1}\log\left(1 + Q_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\right\} - 1\right]\right) \tag{14}$$

where δ^{OP} is given by:

$$\delta^{OP} = \frac{1}{\alpha}\frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}\left(1 + \frac{\kappa\omega}{1-\kappa\omega}\right)(1-\kappa)(1-\rho)(\mu_O - \mu_P) \tag{15}$$

Proof. See Appendix A.1, which also provides the formulas for $a_0, a_1 > 0$, and $a_2 > 0$. □

Remark 1. Theorem 1 establishes uniqueness within the quasi-loglinear class. As best replies and aggregation are non-linear and the space of fundamentals is not compact, one cannot use classical arguments based on Blackwell's sufficient conditions to ensure that the fixed point operator in Equation 12 is a contraction and thereby establish uniqueness in a larger class. Of course, studying an economy with an unbounded support for productivity is merely a mathematically convenient approximation: productivity is certainly bounded, perhaps for a very large bound. In Appendix A.1, we show that there is a unique equilibrium when fundamentals are restricted to lie in a compact set (Lemma 1). Moreover, the claimed quasi-loglinear equilibrium is an ε-equilibrium for any $\varepsilon > 0$ for some sufficiently large support for fundamentals (Lemma 2). Hence, the quasi-loglinear equilibrium is the limit of the unique equilibrium with bounded fundamentals as the bound becomes large. This refinement may be of independent interest and could be used in dispersed information economies in which the uniqueness of equilibrium is an open question outside of the log-linear class, such as those studied by Angeletos and La'O (2013) and Benhabib et al. (2015).

3.2 The Static Effects of Optimism on Output

We now unpack the economics of Theorem 1 to study how models drive non-fundamental fluctuations. We also study the extent to which information can limit the effects of models.

The Effect of Optimism on Output. Optimism affects output in a way that is separable from fundamentals via the function f. This function is non-linear because firms' heterogeneous priors induce heterogeneity in production conditional on productivity and hence also misallocation. Notwithstanding this non-linearity, it turns out that it is useful to summarize the role of optimism by computing the difference in aggregate output when everyone in the economy is optimistic versus when everyone in the economy is pessimistic:

$$\Delta_t \equiv \log Y(\log\theta_t, \log\theta_{t-1}, 1) - \log Y(\log\theta_t, \log\theta_{t-1}, 0) \tag{16}$$

By Theorem 1, we observe that $\Delta_t = \Delta = f(1) - f(0)$, which has an intuitive structure:

Corollary 1 (The Effect of Optimism on Output)**.** *The effect on aggregate output of moving from a fully pessimistic economy to a fully optimistic economy, Δ_t, is invariant to time and the state of the economy and is given by:*

$$\Delta = \frac{1}{1-\omega} \times \alpha \times \delta^{OP} \tag{17}$$

In this expression, as we will later justify formally, δ^{OP} is the partial equilibrium effect of a firm's optimism on the amount of labor the firm hires when we hold fixed the behavior of all other firms and fundamentals. To find the general equilibrium effect of this on aggregate output, Theorem 1 implies that we can first convert the effect of hiring into the output effect via the returns-to-scale parameter α and then apply a *multiplier* $\frac{1}{1-\omega}$. This multiplier is large when strategic complementarities in production arising from aggregate demand externalities are much larger than strategic substitutability that arises from income effects in household labor supply. This multiplier captures the intuitive idea that even a pessimistic firm will produce more if a large fraction of *other* firms is optimistic, as this optimism increases aggregate demand.

The Power of the Truth. This result formalizes and provides nuance for Shiller's (2020) argument that "the truth is not enough to stop false narratives." Specifically, let us define the *power of the truth* as $|\frac{\partial\Delta}{\partial\kappa}|$. This measures how the effect of optimism on firms' hiring and aggregate output changes as they receive more precise information about productivity.

Corollary 2. *The power of the truth $|\frac{\partial\Delta}{\partial\kappa}|$ is positive and increasing in the precision of private information κ. The power of truth is strictly increasing in the precision of private information if and only if strategic complementarity is strictly positive ($\omega > 0$).*

Information is least effective at stopping models exactly when models are at their most powerful. Conversely, when private information is precise and models are weak, the marginal effects of better private information are strong.

This result depends critically on the presence of general equilibrium interactions and strategic complementarity. When there is no strategic complementarity ($\omega = 0$), the power of the truth is constant: $\Delta \propto (1-\kappa)(\mu_O - \mu_P)$, where the constant of proportionality does not depend on κ. In this case, the power of the truth is $|\frac{\partial \Delta}{\partial \kappa}| \propto (\mu_O - \mu_P)$, which depends on the differences in beliefs across models but not the precision of agents' information. Intuitively, providing more information simply scales down how much agents rely on their models. When there is strategic complementarity ($\omega > 0$), increasing the precision of private information now has a second effect: agents know that other agents will be responding more to their signals and relying less on their priors. Because agents want to produce more when others produce more, agents' models about fundamentals become paradoxically *more* important as they now use models more aggressively in forecasting the actions that others will take. This effect dampens the ability of information to stop false models and it does so by more exactly when information is weakest and models are strongest.

3.3 The Dynamics of Models and Output

We now use the characterization of Theorem 1 to describe the economy's dynamics:

Corollary 3 (Model Dynamics)**.** *In the unique quasi-loglinear equilibrium, the fraction of optimists Q_t evolves according to $Q_{t+1} = T(Q_t, \theta_t, \theta_{t-1}, \varepsilon_t)$, where*

$$\begin{aligned} T(Q_t, \theta_t, \theta_{t-1}, \varepsilon_t) = Q_t P_O(a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t), Q_t, \varepsilon_t) + \\ (1 - Q_t) P_P(a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t), Q_t, \varepsilon_t) \end{aligned} \tag{18}$$

This result has two important implications. First, models can be self-propagating. Formally, holding fixed the fundamental and optimism shocks $(\theta_t, \theta_{t-1}, \varepsilon_t)$, the spread of models is shaped by individuals' proclivity to hold onto their current model, social contagiousness, and associativeness, as embodied in P_O and P_P. Second, models can generate persistent effects of one-time fundamental shocks: a one-time productivity shock today can increase aggregate output and thereby increase future optimism.

Steady-States and Hysteresis. To isolate and study these ideas, we first isolate the propagation of models without shocks. Formally, let $T_\theta(Q) = T(Q, \theta, \theta, 0)$ denote the transition map for aggregate optimism with fixed productivity and no optimism shock. We say that a level of optimism Q_θ^* is a deterministic steady state for the 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.

Theorem 2 (Steady State 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:*

$$\theta_P = \exp\left\{\frac{P_P^{-1}(0;0) - a_0}{a_1 + a_2}\right\} \quad \text{and} \quad \theta_O = \exp\left\{\frac{P_O^{-1}(1;1) - a_0 - \Delta}{a_1 + a_2}\right\} \tag{19}$$

   where $P_P^{-1}(x;Q) = \sup\{Y : P_P(Y,Q,0) = x\}$*,* $P_O^{-1}(x;Q) = \inf\{Y : P_O(Y,Q,0) = x\}$*.*
3. *Extreme pessimism is stable if* $\theta < \theta_P$ *and* $P_O(P_P^{-1}(0;0),0,0) < 1$ *and extreme optimism is stable if* $\theta > \theta_O$ *and* $P_P(P_O^{-1}(1;1),1,0) > 0$*.*

Proof. See Appendix A.3. □

If extreme optimism or extreme pessimism is a stable steady state, then the optimistic (or pessimistic) model has a tendency to "go viral" and fully infect the entire population. The conditions under which this occurs can be checked with only a few parameters, which we will later be able to discipline empirically: the responsiveness of output to productivity (a_1, a_2), the impact of all agents being optimistic on output Δ, 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 $P_O^{-1}(1;1)$.

Of particular interest is the case in which, for fixed values of other parameters and fundamentals, either extreme optimism or extreme pessimism could go viral depending on initial conditions. This can induce fully history-dependent, long-run changes in output, a property which we refer to as *hysteresis.* The following corollary characterizes exactly when this can happen:

Corollary 4 (Hysteresis)**.** *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*

$$P_O^{-1}(1;1) - P_P^{-1}(0;0) \leq \Delta \tag{20}$$

Intuitively, this condition is more likely to hold if the optimistic model has a large effect on output (high Δ), if a relatively low output can be consistent with self-fulfilling optimism (low $P_O^{-1}(1;1)$), or if a relatively high output can be consistent with self-fulfilling pessimism (high $P_P^{-1}(0;0)$).

3.4 Hysteresis, Endogenous Persistence, and Boom-Bust Cycles

To more concretely illustrate the features of contagious business cycles and, later, enable us to take the model to the data, we now introduce a parametric family of model updating rules. The *linear-associative-contagious* (LAC) model for updating rules is:

$$\begin{aligned} P_O(\log Y, Q, \varepsilon) &= \left[\frac{u}{2} + r\log Y + sQ + \varepsilon\right]_0^1 \\ P_P(\log Y, Q, \varepsilon) &= \left[-\frac{u}{2} + r\log Y + sQ + \varepsilon\right]_0^1 \end{aligned} \tag{21}$$

where $[z]_0^1 = \max\{\min\{z, 1\}, 0\}$ and ε is i.i.d. $N(0, \sigma_\varepsilon^2)$. The parameter $u \geq 0$ captures *stubbornness*, or all agents' proclivity not to change models. The parameter $r \geq 0$ captures *associativeness*, or the extent to which agents associate high output with the optimistic model. The parameter $s \geq 0$ captures *contagiousness*, or the direct effect of peers' models on one's own. Finally, the aggregate shock allows models to fluctuate autonomously.

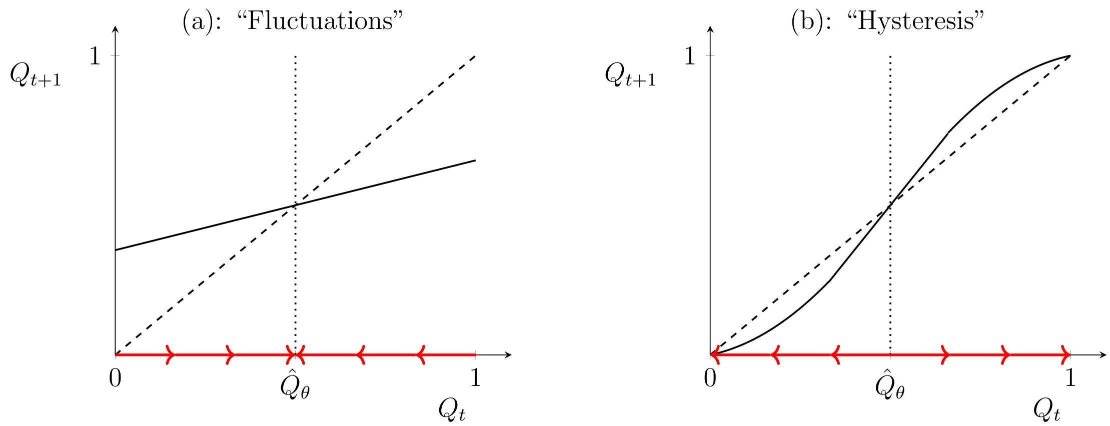
We now illustrate three qualitative properties of contagious business cycles:

1. *Hysteresis and the criticality threshold*: despite equilibrium uniqueness, there can be multiple steady states and a critical level of model adoption away from which the economy diverges.
2. *Endogenous persistence of output*: stubbornness, associativeness, and contagiousness generate state-dependent and size-dependent persistence of one-time shocks.
3. *Boom-bust cycles*: even when hit by i.i.d. stochastic shocks, the economy features a tendency toward boom-bust cycles.

In Appendix B.1, we formalize these properties of shock responses in a larger class of non-parametric updating rules. Below, for purposes of simplest exposition, we describe them using examples from the LAC class.

Hysteresis and the Criticality Threshold. Figure 1 visualizes the transition map for two example calibrations of the updating rule, fixing the state θ and the calibration of other parameters. In panel (a), stubbornness, contagiousness, and associativeness are relatively low. The transition map T_θ crosses the 45-degree line once, from above. Therefore, the interior steady state denoted by $\hat{Q}_\theta$ is stable, and optimism tends to converge to this level if perturbed away from it. For this reason, we refer to this as a case that admits "fluctuations" if hit by shocks. In panel (b), stubbornness, contagiousness, and associativeness are relatively high. The transition map T_θ intersects the 45-degree line three times: twice at the extremes of $Q = 0$ and $Q = 1$ and once from below at an interior level $\hat{Q}_\theta$. Paths for Q that start slightly to the left or right of $\hat{Q}_\theta$ converge, respectively, to the stable points of $Q = 0$ or $Q = 1$. In

Figure 1: Fluctuations *vs.* Hysteresis in the Linear-Associative-Contagious Model



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. Both correspond to the linear-associative-contagious model with different calibrations for the underlying parameters. In panel (a) ("fluctuations"), the condition for extremal multiplicity (Equation 22) does not hold. In panel (b) ("hysteresis"), the condition does hold, and $Q = 0$ and $Q = 1$ are stable steady states.

this sense, dynamics of optimism display hysteresis: holding fixed fundamentals, the long-run behavior of the economy depends on initial conditions. Lemma 3 in the Appendix formalizes these ideas by showing exactly when $\hat{Q}_\theta$ is on the boundary of two basins of attraction for, respectively, extreme pessimism and extreme optimism.

We can analytically compute the condition in Corollary 4, which determines when extreme optimism and pessimism are both stable steady states.[2] Thus, extreme optimism and pessimism are steady states if and only if:

$$M = u + s + r\Delta - 1 \geq 0 \tag{22}$$

which is to say that stubbornness, associativeness, contagiousness, and the equilibrium impact of optimism on output are sufficiently large. This expression clarifies that strong *static complementarities*, which would manifest in high Δ, are sufficient but not necessary for extremal multiplicity. In particular, stubbornness and contagiousness contribute *dynamic complementarity* that can also induce extremal multiplicity. Thus, the parameter M, which incorporates both static and dynamic complementarity, is the correct gauge for the "strength" of models to generate hysteresis.

Moreover, as suggested by panel (b) of Figure 1, the model with stable extremal steady

[2]In particular, $P_O^{-1}(1;1)$ solves $1 = \frac{u}{2} + rP_O^{-1}(1;1) + s$, so $P_O^{-1}(1;1) = \frac{1}{r} - \frac{u}{2r} - \frac{s}{r}$; $P_P^{-1}(0;0)$ solves $0 = -\frac{u}{2} + rP_P^{-1}(0;0)$, so $P_P^{-1}(0;0) = \frac{u}{2r}$.

states has an unstable, intermediate steady state $\hat{Q}_\theta \in (0,1)$ that solves $\hat{Q}_\theta = T_\theta(\hat{Q}_\theta)$, or

$$\hat{Q}_\theta = \frac{u}{2}(2\hat{Q}_\theta - 1) + s\hat{Q}_\theta + r(a_0 + (a_1 + a_2)\log\theta + f(\hat{Q}_\theta)) \tag{23}$$

We refer to this value of optimism as the *criticality threshold* because it separates regions of the state space that are attracted to extreme optimism versus extreme pessimism. Under the approximation that $f(Q) \approx \Delta Q$, which we later find to be quantitatively accurate,

$$\hat{Q}_\theta \approx \frac{\frac{u}{2} - r(a_0 + (a_1 + a_2)\log\theta)}{M} \tag{24}$$

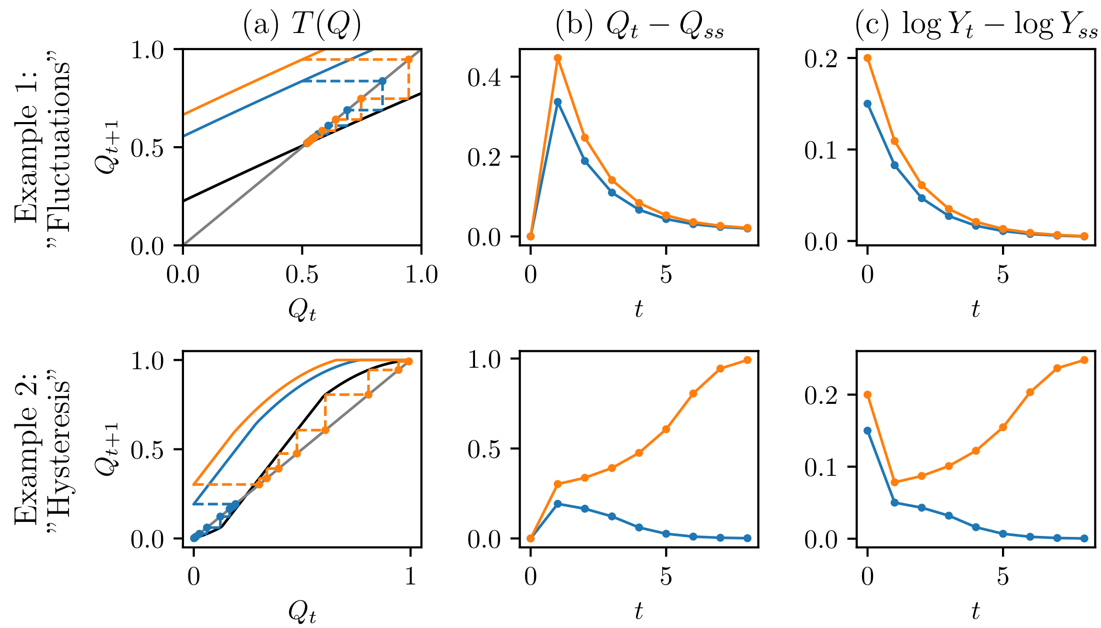
Hence, greater contagiousness (s), associativeness (r), and static economic impact of optimism (Δ) reduce $\hat{Q}_\theta$, or equivalently decrease the lower bound of initial optimism that is consistent with the optimistic model eventually going viral.

Endogenous Persistence. We now study how the economy responds to shocks. Figure 2 illustrates how the economy responds to shocks under a "fluctuations" versus "hysteresis" calibration of the model. In both calibrations, productivity shocks are perfectly transitory. The blue and orange lines of each plot respectively illustrate smaller and larger one-time productivity shocks at $t = 0$. In both cases, because of positive associativeness, these correspond to one-time upward shifts in the transition maps $T(Q)$ (panel (a)). The dots and dashed lines in panel (a) trace out the dynamic response of optimism to each shock using the transition map. Panels (b) and (c) illustrate the impulse responses of optimism and output.

In the fluctuations case (row 1), persistence of models creates endogenous persistence in the economic boom. Because of positive associativeness ($r > 0$), stubbornness ($u > 0$), and contagiousness ($s > 0$), optimism remains elevated for several periods before smoothly converging back to the steady state. At $t = 0$, output is elevated above its steady-state value *only* because of the productivity shock; for $t \geq 1$, output is elevated because of the persistent increase in optimism, even though productivity has returned to its steady-state value. The large shock (orange) leads to a larger and more persistent boom than the small shock (blue).

In the hysteresis case (row 2), the small shock leads to a highly persistent boom, whereas the large shock leads to a regime shift. This discontinuity of shock responses as a function of shock size emerges because large shocks can push the economy above the unstable interior steady state (panel (a)). Intuitively, the large shock seeds enough optimism for the optimistic model to "go viral." This also induces a non-monotone response of output (panel (c)): while the direct effect of productivity disappears after one period, the effect of viral optimism grows over time.

Figure 2: Endogenous Persistence in Shock Responses

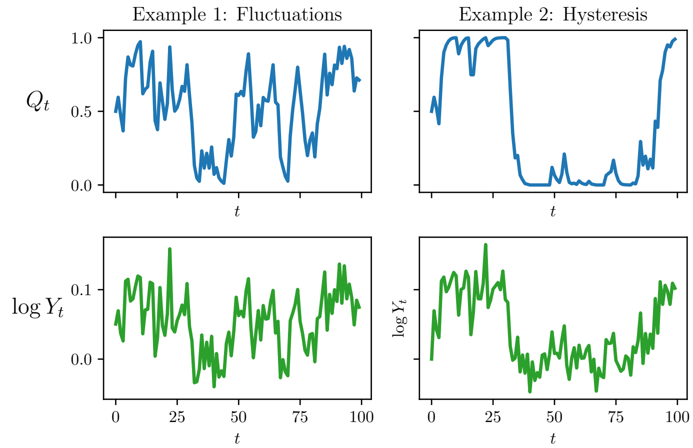


Notes: This figure illustrates the response of the economy to transitory productivity shocks under two different model calibrations. The top row corresponds to a "fluctuations" calibration and the bottom row corresponds to a "hysteresis" calibration, as defined in the main text. The orange lines correspond to a larger productivity shock and the blue lines correspond to a smaller productivity shock. Column (a) shows the transition map for aggregate optimism (black), its perturbations under each shock (colors), and the paths of optimism (dots and dashed lines). Column (b) shows the impulse response of optimism relative to the interior steady-state value (top row) and relative to extreme pessimism (bottom row). Column (c) shows the impulse response of log output relative to the respective steady-state values.

In sum, both regimes feature *endogenous persistence* of output even in response to one-time shocks. In the hysteresis regime, there is the possibility also of permanent economic effects of temporary shocks. Propositions 1 and 2 in the Appendix formalize these properties, and moreover characterize when endogenous persistence is large enough to generate "hump-shaped" impulse response functions of output in response to perfectly transitory shocks.

Boom-Bust Cycles. We finally study how these different cases map to time-series properties of the macroeconomy. To visualize this, we simulate from "fluctuations" and "hysteresis" calibrations of the model for 100 periods in Figure 3. Productivity shocks are *common* across the two simulations and there are no direct shocks to model evolution ($\sigma_\varepsilon = 0$). In both simulations, output is persistent despite the lack of persistent driving shocks. This arises because of persistent variation in optimism. In the fluctuations case, optimism (and therefore output) is mean-reverting. In the hysteresis case, optimism swings between extremes at low

Figure 3: Simulated Paths of the Economy under "Fluctuations" *vs.* "Hysteresis"



Notes: Each panel corresponds to a simulated time series of the model, with identical time paths of i.i.d. productivity shocks but different calibrations. The condition for extremal multiplicity (Equation 22) does not hold in example 1 ("Fluctuations") but does hold in example 2 ("Hysteresis").

frequencies. This induces sudden but persistent "booms and busts." Proposition 3 in the Appendix formalizes this and provides analytical bounds on the period of boom-bust cycles.

3.5 Additional Results and Extensions

Persistent Idiosyncratic Shocks and Model Updating. In Appendix B.5, we show that a model in which models concern the probability distribution of *idiosyncratic* revenue total factor productivity (TFPR) yields the same predictions as our main model. Thus, as long as there are aggregate model dynamics, models need describe neither the macroeconomy nor physical productivity *per se*. This is an important motivation for our measurement strategy, which will not distinguish between optimism about micro versus macro conditions.

Multi-dimensional Models. In Appendix B.6, we generalize the model to allow for arbitrarily many models regarding the mean, persistence, and volatility of fundamentals, which is essentially exhaustive within the Gaussian class. In Appendix B.4, we generalize our analysis to feature a continuum of models. In each case, we generalize our results to characterize equilibrium output and model evolution.

Bayesian Updating. An important model that is ruled out by our conditions on the updating rule 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.3, this "Bayesian benchmark" contradicts the dependence of firms' updating on Q_t and ε_t conditional on $\log Y_t$ (respectively, contagiousness and shocks). Moreover, this "Bayesian benchmark" predicts that agents converge to holding the better-fitting empirical model exponentially quickly. Later, we will show that such a prediction is at odds with our finding of cyclical dynamics for aggregate optimism (Figure A1). However, in principle, richer Bayesian models that are consistent with our empirical results might be nested by the model updating process.

Contrarianism, Endogenous Cycles, and Chaos. While this model generates non-fundamental fluctuations, it cannot generate fully endogenous cycles and chaotic dynamics. In Appendix B.8, 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; we reject at the 95% confidence level that either cycles or chaos obtain.

Welfare Implications. In Appendix B.2, we study the normative implications of optimism and provide conditions under which its presence is welfare improving, despite its being misspecified. Intuitively, optimism acts as if it were an *ad valorem* price subsidy for firms, which induces firms to hire more and can undo distortions caused by market power.

3.6 From Theory to Measurement

We have shown that the theoretical properties of contagious business cycles hinge on: (1) the effect of optimism on hiring, (2) agents' updating rules, (3) the persistence and volatility of exogenous shocks, and (4) the extent of private information. We now show how to identify these objects conditional on calibrating four standard macroeconomic production and preference parameters. This will form the basis for our empirical strategy and quantification.

Step I: Identification of The Effect of Optimism. Theorem 1 implies that f, the effect of optimism on output in the unique quasi-loglinear equilibrium, is identified given knowledge of both δ^{OP} and the standard macroeconomic parameters $(\alpha, \epsilon, \gamma, \psi)$. Moreover, δ^{OP} can be recovered via a simple regression of firms' hiring on their optimism:

Corollary 5 (Firm Hiring Regression)**.** *In the unique quasi-loglinear equilibrium*

$$\Delta \log L_{it} = \gamma_i + \chi_t + \tau_1 \log \theta_{it} + \tau_2 \log L_{i,t-1} + \delta^{OP} \lambda_{it} + \zeta_{it} \tag{25}$$

where $\chi_t = c_1 \log \theta_t + c_2 \log \theta_{t-1} + c_3 f(Q_t)$ *for some constants* c_1, c_2, *and* c_3, *and* ζ_{it} *is an i.i.d. normal random variable with zero mean. Moreover, conditional on* $(\alpha, \epsilon, \gamma, \psi)$, δ^{OP} *uniquely identifies* f, *the equilibrium effect of optimism on output.*

Proof. See Appendix A.4. □

The time fixed-effect χ_t absorbs two aggregate equilibrium forces: the general-equilibrium effect of optimism on hiring, $c_3 f(Q_t)$, and the effects of aggregate productivity on aggregate output, $c_1 \log \theta_t + c_2 \log \theta_{t-1}$. Without the time fixed effect, the regression would produce a biased estimate for δ^{OP} because of the correlation of optimism with aggregate economic fundamentals. This underscores the necessity of combining cross-sectional variation and a structural model for general-equilibrium interaction to identify the effect of models on economic outcomes.[3]

Step II: Identification of Updating Rules. In the linear-associative-contagious (LAC) model for updating rules, we can identify u, r, and s by estimating a linear probability model for the evolution of optimism at the firm level. The residual term in this regression corresponds to idiosyncratic shocks to updating (since the model is probabilistic) plus the aggregate shock ε.

Step III: Identification of Private Information and the Shock Processes. To obtain the law of motion of aggregate output, we require the four parameters that govern the persistence of productivity ρ, the volatility of productivity innovations σ, the signal-to-noise ratio for productivity κ, and the volatility of optimism shocks σ_ε. The key to our identification strategy for the first three parameters is the following observation: after removing the non-fundamental component of output identified by Step I, the fundamental component follows an ARMA(1,1) process.

Corollary 6 (Fundamental Output is an ARMA(1,1))**.** *In the unique quasi-loglinear equilibrium, the fundamental component of output,* $\log Y_t^f = \log Y_t - f(Q_t) - a_0$*, follows an ARMA(1,1) process:*

$$\log Y_t^f = \rho \log Y_t^f + a_1 \sigma \nu_t + a_2 \sigma \nu_{t-1} \tag{26}$$

where (a_0, a_1, a_2) *are the coefficients characterized in Theorem 1 and* ν_t *is i.i.d.* $N(0,1)$.

The coefficients of this ARMA(1,1) process impose three restrictions on the three parameters (ρ, σ, κ). We finally observe that, conditional on all other parameters, the scaling of the optimism shock σ_ε is pinned down by the time-series variance of aggregate optimism $\mathrm{Var}[Q_t]$.

[3] We note that our argument for identifying the partial-equilibrium effect of optimism on hiring and aggregating these effects via the model also applies if models concern firms' idiosyncratic productivity (Appendix B.5). Thus, our empirical and quantitative analysis is not sensitive to this modeling choice. As we explain formally in the Appendix, only the exact identification of the underlying parameters κ and $\mu_O - \mu_P$ would change, while the outcome-relevant objects δ^{OP} and f remain the same.

4 Data and Measurement

To implement these steps and take the model to the data, we construct a panel dataset on firms' sentiment and decisions.

4.1 Data

Text. We measure proxies for firms' models using textual data. Our main sources of data are Forms 10-K, annual reports submitted to the Securities and Exchange Commission (SEC) by each publicly traded firm in the US. 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" (SEC, 2011). Moreover, relevant to our purposes, "company management also discusses its perspective on the business results and what is driving them" (SEC, 2011). We download all SEC Forms 10-K from the SEC Edgar database from 1995 to 2018. 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 (see Appendix C.1 for details). 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' quarterly sales and earnings conference calls. Our initial sample consists of 158,810 documents from 2002 to 2014. We collapse the data to 25,589 firm-by-fiscal-year observations (see Appendix C.2 for details).

Fundamentals and Choices. We measure 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. As is standard, we also drop 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.

Manager and Analyst Beliefs. We collect data from the International Brokers' Estimate System (IBES) on firm-level quantitative forecasts. Specifically, we use the IBES Guidance dataset to measure management's forecasts of sales and capital expenditures, which IBES records from press releases or transcripts of corporate events. We use the IBES Estimates dataset to measure the median forecast of equity analysts, focusing on the long-term growth (LTG) expectations of earnings over three to five years as in Bordalo et al. (2024).

4.2 Recovering Optimism from Language

Inspired by the linguistic analysis in the "Narrative Economics" of Shiller (2020) and the burgeoning literature using textual data to understand the attitudes of firms (see the survey by Hassan et al., 2024), we measure firms' models using their language. In particular, to map to our model notion of optimism and pessimism, we apply the standard and widely used sentiment-scoring approach of Loughran and McDonald (2011). These authors construct dictionaries of positive and negative words suitable for financial documents, in which certain words (*e.g.*, the leading example "liability") have definitions that differ from their use in common English and their interpretation in standard (non-financial) sentiment scoring.[4] We calculate positive and negative sentiment as:

$$\text{pos}_{it} = \sum_{w \in \mathcal{W}_P} \text{tf}(w)_{it} \qquad\qquad \text{neg}_{it} = \sum_{w \in \mathcal{W}_N} \text{tf}(w)_{it} \tag{27}$$

where $\mathcal{W}_P$ is the set of positive words, $\mathcal{W}_N$ is the set of negative words, and $\text{tf}(w)_{it}$ is the term frequency of all bigrams including word w in the time-t 10-K of firm i.[5] 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:

$$\text{opt}_{it} = \mathbb{I}\left[\text{sentiment}_{it} \geq \text{med}\left(\text{sentiment}_{it}\right)\right] \tag{28}$$

Aggregating this measure across firms, we find that aggregate optimism is persistent, with an autocorrelation of 0.75, and cyclical, with a correlation of -0.37 with the contemporaneous level of unemployment (see Figure A1). Both features are to be expected in our model. The former is a result of stubbornness, contagiousness, associativeness, and autocorrelation of fundamentals. The latter could reflect either direction of causality: current conditions could shape the adoption of models, or models might affect economic outcomes. Because the time series evidence cannot distinguish among these different explanations, we will use cross-sectional variation in optimism to isolate its effects.

Measuring Topics of Discussion. Measuring optimism is sufficient for describing firms' prior beliefs in our model, as there is only one dimension of fundamental uncertainty. But this strategy abstracts from the richer details of what firms discuss. We revisit this point and describe strategies to measure (and interpret) more specific topics in Section 6.4.

[4] 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 methods in Appendix C.3.

[5] For reference, we print the 20 most common words in each set in Table A1.

5 Empirical Results

We now use our micro data to estimate how optimism affects decisions and spreads. In the process, we apply tests to distinguish whether optimism arises from non-fundamental beliefs (as in our model) or news about future fundamentals.

5.1 Optimism Drives Hiring

We first estimate the relationship between optimism in language and hiring. The estimating equation is derived in Corollary 5. Specifically, we estimate the following firm-by-fiscal-year model:

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \tag{29}$$

The outcome variable is the log difference of the firm's employment across fiscal years ("hiring") and the main regressor, opt_{it}, is the binary indicator for optimism whose construction is described in Section 4. We control for firm and industry-by-time fixed effects to sweep out fixed differences across firms and non-parametric trends and cycles within industries. We include a suite of firm-level time-varying controls X_{it} including current and past TFP, lagged labor, and financial variables.[6] Viewed through the lens of the model, the estimated effect δ^{OP} combines two margins: the effect of optimism on beliefs and the effect of beliefs on input choices. We could obtain a null result of $\delta^{OP} = 0$ if optimism in language has no influence on firms' choices over and above other measured fundamentals.

We find that optimistic firms hire more than pessimistic firms holding fixed other observed fundamentals (Table 1). We first estimate the model with no additional controls other than fixed effects and estimate $\hat{\delta}^{OP} = 0.0355$ 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, adjustment costs.[7] Our point estimate $\hat{\delta}^{OP} = 0.0305$ (SE: 0.0030) is quantitatively comparable to our uncontrolled estimate. To formalize this, in Appendix E.1 we report the robustness of our estimate to selection on unobservables using the method of Oster (2019). We find that our finding of a positive effect of optimism on hiring is robust by the benchmark suggested by Oster (2019) (see Table A2).

In column 3, we add measures of firms' financial characteristics, the (log) book-to-market

[6] 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. More details are provided in Appendix D.2.

[7] To evaluate robustness to richer adjustment dynamics, in Table A3, 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: 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 29 and the outcome is the change in firms' log employment from year $t-1$ to t. The main regressor is a binary indicator for optimism, defined in Section 4.2. In column 5, the regression model is Equation 30, the outcome is the log change in firms' employment from year t to $t+1$, and control variables are dated $t+1$. 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.

ratio, last fiscal year's log stock return (inclusive of dividends), and leverage (total debt over total assets). These controls proxy for Tobin's q and firm-level financial frictions, features absent from our model but potentially relevant in practice. These controls are conservative in that they may absorb variation in both omitted firm fundamentals and optimism itself. The point estimate remains positive and quantitatively similar. In column 4, we estimate a specification with the controls from column 2 but no firm fixed effects to guard against small-sample bias from strict exogeneity violations (Nickell, 1981). We find similar results.[8]

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:

$$\Delta \log L_{i,t+1} = \delta_{-1}^{OP} \text{opt}_{it} + \tau' X_{i,t+1} + \gamma_i + \chi_{j(i),t+1} + \varepsilon_{i,t+1} \tag{30}$$

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

[8] In Table A4, we report standard errors for the estimates of Table 1 under alternative clustering approaches.

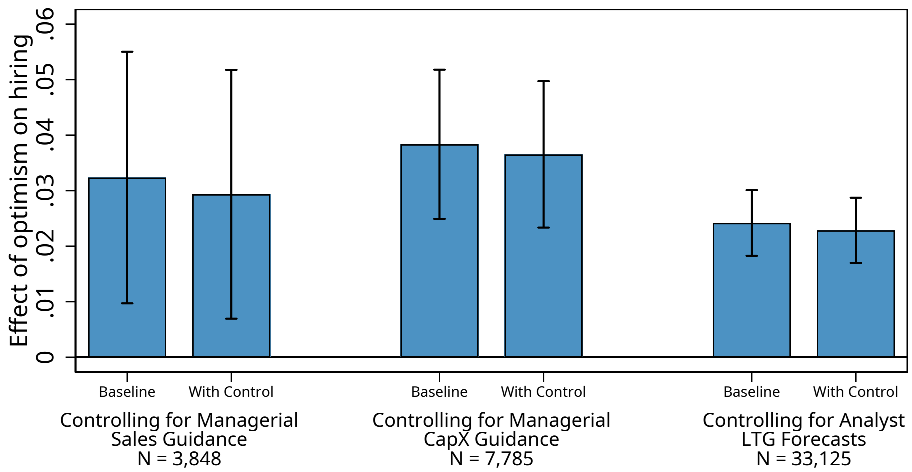
similar in magnitude to our comparable baseline estimate (column 2). In Table A5, we report results from our baseline regression Equation 29, using $\text{opt}_{i,t-1}$ as an instrument for opt_{it}. This is robust to any identification concern arising from the simultaneous determination of opt_{it} and $\Delta \log L_{it}$, but estimates the original parameter δ^{OP} rather than δ^{OP}_{-1}. Our estimates are positive, statistically significant, and larger than our baseline estimates.

Leveraging Exogenous CEO Exits as Natural Experiments. To further isolate plausibly exogenous variation in firms' optimism, we study the effects of changes in optimism induced by plausibly exogenous CEO turnover. We provide the details in Appendix E.2. Specifically, we estimate a variant of Equation 29 over firm-year observations corresponding to the death, illness, or voluntary retirement of a CEO, as measured by Gentry, Harrison, Quigley, and Boivie (2021). Employing this strategy, we find even larger effects of optimism on hiring than those in our baseline specification.

Language Matters Over and Above Measured Beliefs. In our mapping from theory to data, we treat sentiment in language as a proxy for the optimistic and pessimistic models that shift beliefs. Thus, if we are trying to obtain a model-consistent estimate for the effect of optimism on beliefs (the structural parameter δ^{OP} in Corollary 5), other measures of managers' beliefs could be a "bad control" in our regression model. Nonetheless, it may be of independent interest to check whether text-based measures of optimism are useful to predict firms' decisions conditional on other measures of beliefs that have been pioneered by prior work on belief-driven fluctuations (Gennaioli et al., 2016; Bordalo et al., 2021, 2024). To gauge this, we construct three firm-level variables directly measuring beliefs. The first two are based on managerial guidance and respectively measure the forecasted growth rate of sales and capital expenditures. The third is based on equity analysts' "long-term growth" forecast of corporate earnings growth over three to five years. We note that managers and analysts do not always provide forecasts, whereas firms always write a 10-K; therefore, these beliefs data are available for only a subset of our full firm-year sample.

Even where forecasts are available, we find that optimism in text has a large and quantitatively stable effect on hiring conditional on beliefs (Figure 4). We also find a similar result for capital investment (Figure A2). The importance of words over and above quantitative forecasts is consistent with the hypothesis of Shiller (2017) that language contains important information about economic actions—and even with Keynes's (1936) claim in his *General Theory* that the emotional states that spur managers to action are not reflected in "mathematical expectation." Our finding is also consistent with the hypothesis from corporate finance that managers rely heavily on non-quantitative *soft information* (Liberti and Petersen, 2019) to make decisions. Finally, from a methodological perspective, our analysis

Figure 4: Language Matters Conditional on Beliefs



Notes: The regression model is Equation 29, the outcome is the change in firms' log employment from year $t-1$ to t, the main regressor is a binary indicator for optimism, and all specifications include firm and industry-by-time fixed effects. In each panel, we add a different control variable measuring beliefs: managerial guidance for sales growth (log of guidance value minus log of last year's sales), managerial guidance for capital expenditures growth (log of guidance value minus log of last year's capital expenditures), and analysts' long-term growth forecasts (both contemporaneous and first lag). The two bars show the coefficient on optimism on a common sample without and with the controls, respectively. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals based on standard errors clustered by firm ID.

suggests that textual data—which are available more comprehensively in standard datasets than forecast data—may be a viable alternative to quantitative forecasts for many economic analyses.

Robustness and Alternative Strategies. In the Appendix, we also report several additional results that probe the robustness of our main specification. We summarize them briefly here. First, Table A6 repeats the analysis of Table 1 with our conference-call-based optimism measure, and finds similar results. Second, Table A7 repeats our main analysis for different measured inputs—employment (the baseline), total variable input expenditure, and investment—and demonstrates a positive and comparably sized effect of optimism on all three. Third, in Figure A3 we re-create the regression models of the first three columns of Table 1 with indicators for each decile of our continuous sentiment measure. We find monotonically increasing associations of hiring with sentiment, implying that our binary construction is not masking non-monotone effects of the continuous measure.

5.2 Non-Fundamental Beliefs *vs.* News

We have interpreted our text-based measurement of optimism as a proxy for firms' believing in an optimistic model: that is, a shift in priors toward being more optimistic about future economic conditions. An alternative interpretation is that text-based optimism measures news about future firm-level fundamentals. This would be inconsistent with our model, in which optimism conveys no news about those fundamentals. Moreover, while our CEO-exits design should handle such concerns, this news-based interpretation could create an identification threat that is not handled by controlling for measured *past* fundamentals of the firm.

We perform two tests that distinguish between the news and non-fundamental beliefs interpretations. First, we test whether a firm's optimism predicts positive future fundamentals and performance. This must be the case under the news interpretation, but need not be the case under the non-fundamental beliefs interpretation. Second, we test how a firm's optimism affects its forecasts. Under both models, optimism should predict optimistic forecasts. The news model predicts, if firms are Bayesian, that optimism efficiently enters forecasts and therefore does not predict forecast errors; our model, by contrast, implies that optimism predicts systematically *over*-optimistic forecasts.

Test I: Optimism Predicts Poor(er) Future Performance. To conduct the first test, we estimate projection regressions of firm fundamentals and performance Z_{it}, 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:[9]

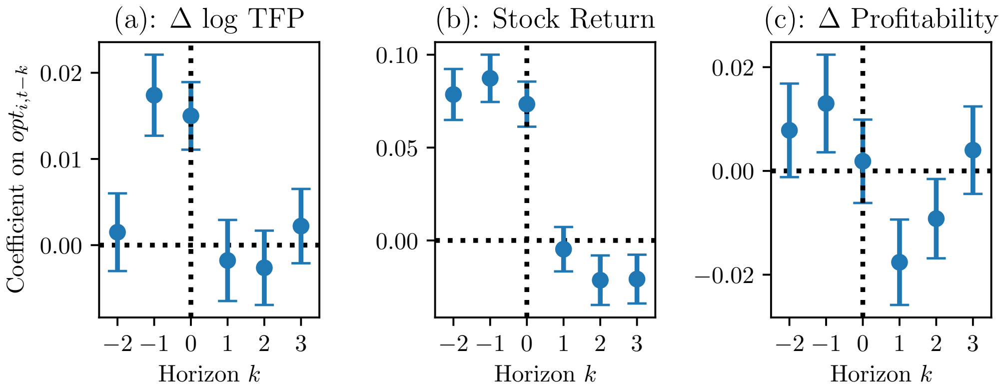
$$Z_{it} = \beta_k \, \mathrm{opt}_{i,t-k} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{31}$$

Under the "news" hypothesis, we would expect $\beta_k > 0$ for $k > 0$: that is, optimistic firms are both more productive and more successful than their pessimistic counterparts in the future. Under the "non-fundamental beliefs" hypothesis, we should expect to see that $\beta_k = 0$ for $k > 0$ for firm productivity (as a measure of fundamentals) and that $\beta_k < 0$ for $k > 0$ for firm performance.

Our findings, reported in Figure 5, strongly contradict this "news" hypothesis and are, instead, consistent with the "non-fundamental beliefs" hypothesis. 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.

[9] 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).

Figure 5: The Dynamic Relationship between Optimism and Firm Performance



Notes: The model is Equation 31 and each dot shows the coefficient on binary optimism from a different projection regression. The outcomes are (a) the log change in TFP (b) the log stock return inclusive of dividends over the fiscal year, 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.

That is, a firm that is optimistic this year does not on average do better next year.[10] We find, in sharp contrast, that optimistic firms have *worse* stock returns and 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 5). Figure A4 replicates this analysis with conference-call-based optimism and finds similar results.[11] Finally, in Figure A5 we replicate this analysis with other financial fundamentals (leverage, capital structure, payout policy, and stock return volatility): consistent with the findings above, optimistic firms face relatively *tighter* future financial conditions.

Test II: Optimism Predicts Over-Optimistic Beliefs. We next directly test whether optimistic firms hold over-optimistic beliefs. We do this by linking a subset of our data on optimism with data on managerial guidance forecasts. We construct $\text{GuidanceOverOpt}_{it}$ as an indicator of managers' guidance minus the realization exceeding the sample median.[12]

[10] To further investigate the effects on stock prices, we also estimate the correlation of optimism with stock returns near the 10-K filing date (Table A8). We find essentially no evidence of stock response on or before the filing day, and weak evidence of positive returns (about 15-25 basis points) in the four days after. The latter finding is consistent with those in Loughran and McDonald (2011).

[11] Jiang, Lee, Martin, and Zhou (2019) relatedly find that positive textual sentiment in firm disclosures, by their own measure, predicts negative excess returns over the subsequent year.

[12] When managers' guidance is reported as a range, we code a point-estimate forecast as the range's midpoint. The method of comparing to the median corrects for the fact that, in more than half of our observations, guidance is lower than the realized value, presumably due to asymmetric incentives.

Table 2: Optimism Predicts Over-Optimistic Forecasts

	(1)	(2)
	Outcome is $\text{GuidanceOverOpt}_{i,t+1}$	
opt_{it}	0.0310	0.0390
	(0.0168)	(0.0228)
Ind.-by-time FE	✓	✓
Lag labor		✓
Current and lag TFP		✓
N	4,773	2,735
R^2	0.161	0.180

Notes: The regression model is Equation 32. The main regressor is a binary indicator for optimism, defined in Section 4.2. The outcome is a binary indicator for whether sales guidance was high relative to realized sales. Standard errors are two-way clustered by firm ID and industry-year.

We estimate the following regression model:

$$\text{GuidanceOverOpt}_{i,t+1} = \beta\, \text{opt}_{it} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it} \tag{32}$$

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 and those firms do not always provide guidance, we do not include firm fixed effects. Our findings are reported in Table 2. We find a positive relationship that gets stronger when we add the aforementioned control variables (columns 1 and 2). That is, textual optimism corresponds to forecasts that are predictably more likely to exceed subsequent performance. This is exactly what we would predict under the "non-fundamental beliefs" hypothesis and not what we would predict under the "news" hypothesis.[13]

Summary. Based on these tests, we argue that there is strong evidence in favor of the non-fundamental beliefs interpretation suggested by our model. To be concrete, to argue against this interpretation of the data, one would have to argue that firms that use positive language, subsequently expand hiring and investment, have predictably over-optimistic forecasts, and perform worse in the future were somehow correct in their optimism.

[13]In an analogous regression in which the outcome measures managerial optimism relative to contemporaneous *analyst* forecasts, we find an imprecise positive effect in an uncontrolled model and a zero effect in the controlled model (Table A9). These findings are consistent with a story in which models are shared between management and investors, potentially due to persuasion in communications.

5.3 Optimism is Contagious and Associative

We next estimate how optimism spreads across firms. Specifically, we estimate a version of the linear-associative-contagious updating rule (Equation 21) in our panel data:

$$\mathrm{opt}_{it} = u\ \mathrm{opt}_{i,t-1} + s\ \overline{\mathrm{opt}}_{t-1} + r\ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it} \tag{33}$$

where $\overline{\mathrm{opt}}_{t-1}$ 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. In our model interpretation, u measures stubbornness, s measures contagiousness, and r measures associativeness.

We find strong evidence of all three forces (Table 3). 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 finding of $s > 0$, in particular, is consistent with Shiller's (2020) hypothesis that optimism is contagious.

Our estimation of Equation 33 leverages only time-series variation. While this is the level of variation that is relevant for calibrating the model, one may worry that the small sample size leaves open the door to spurious correlation. We therefore also study a model that allows for contagiousness and associativeness at finer levels. Specifically, we estimate the following variation of the original regression at the industry level:

$$\mathrm{opt}_{it} = u_{\mathrm{ind}}\ \mathrm{opt}_{i,t-1} + s_{\mathrm{ind}}\ \overline{\mathrm{opt}}_{j(i),t-1} + r_{\mathrm{ind}}\ \Delta \log Y_{j(i),t-1} + \gamma_i + \chi_t + \varepsilon_{it} \tag{34}$$

where $\overline{\mathrm{opt}}_{j(i),t-1}$ is the leave-one-out mean of optimism within industry $j(i)$ and $\Delta \log Y_{j(i),t-1}$ is the growth of sectoral value-added, measured by linking BEA sector-level data to our NAICS-based classification.[14] The time fixed effect χ_t absorbs aggregate contagiousness and associativeness. We find strong evidence for contagiousness and weaker evidence for associativeness within industries (column 2 of Table 3).[15]

Robustness. As a robustness check, we also measure contagiousness at a finer level by defining narrow sets of peers that share equity analysts for firms listed on the New York Stock Exchange, following Kaustia and Rantala (2021). We find both a quantitatively similar industry-level effect and an independent peer-set effect (Table A11). Moreover, we find consistent evidence of stubbornness, contagiousness, and associativeness for the continuous measure of sentiment (Table A12).

Inspecting the Mechanism: Spillovers are Not Driven by Common Shocks. The coefficients of interest (u, r, and s) identify stubbornness, associativeness, and contagious-

[14] These data are available only from 1997.

[15] In Table A10, we report standard errors for Table 3 under alternative clustering.

Table 3: Optimism is Contagious and Associative

	(1)	(2)
	Outcome is opt_{it}	
Own lag, $\text{opt}_{i,t-1}$	0.209	0.214
	(0.0071)	(0.0080)
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290	
	(0.0578)	
Real GDP growth, $\Delta \log Y_{t-1}$	0.804	
	(0.2204)	
Industry lag, $\overline{\text{opt}}_{j(i),t-1}$		0.276
		(0.0396)
Industry output growth, $\Delta \log Y_{j(i),t-1}$		0.0560
		(0.0309)
Firm FE	✓	✓
Time FE		✓
N	64,948	52,258
R^2	0.481	0.501

Notes: The regression model is Equation 33 for column 1 and Equation 34 for column 2. Aggregate and industry optimism are averages of the 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.

ness, when idiosyncratic optimism, aggregate optimism, and GDP are unrelated to other factors that affect changes in optimistic sentiment at the firm level. Since the key regressor is *lagged* aggregate optimism, 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.

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 GDP or value-added growth at the aggregate and industry levels. Specifically, we estimate

$$\text{opt}_{it} = u\ \text{opt}_{i,t-1} + s\ \overline{\text{opt}}_{t-1} + \gamma_i + \sum_{k=-2}^{2} \left(\eta_k^{\text{agg}} \Delta \log Y_{t+k} + \eta_k^{\text{ind}} \Delta \log Y_{j(i),t+k}\right) + \varepsilon_{it} \tag{35}$$

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 contagiousness 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.

Table 4: Optimism is Contagious, Controlling for Past and Future Outcomes

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

Notes: The regression model is Equation 35 for columns 1-3, and an analogous industry-level specification for columns 4 and 5 (*i.e.,* Equation 34 with past and future controls). Columns 1 and 4 correspond, respectively, with columns 1 and 3 of Table 3. The added control variables are two leads, two lags, and the contemporaneous value of: real GDP growth (columns 2-3) and industry-level output growth (columns 3 and 6). Standard errors are two-way clustered by firm ID and industry-year.

We report our estimates of the contagiousness coefficients in Table 4, adding the "bad controls" one at a time (columns 2 and 3) and find similar results to our baseline (column 1). Similarly, for our industry-level estimates, we find no statistically significant evidence of coefficient attenuation as additional controls are added (columns 4 and 5). In Table A13, we report analogous estimates with the continuous sentiment variable and find similar results. These estimates indicate that our results are not driven by omitted aggregate conditions.

Alternative Identification Strategy: Granular IV. While the previous strategy suggests that aggregate shocks do not bias our findings, it remains possible that unmeasured aggregate shocks to fundamentals could bias our results. To further test whether our measure of contagiousness captures spillovers, we pursue a granular instrumental variables strategy. The idea underlying this strategy is that larger firms are more likely to influence the views of other firms than smaller firms. Under this view, we can measure idiosyncratic changes in the optimism of firms (which we have shown are non-fundamental in Test I above) and weight this by firm size to construct a granular instrumental variable for past optimism (Gabaix and Koijen, 2020). Although not comparable to our main estimates because the spillover measure is different, we recover a large and statistically significant contagiousness effect. Contemporaneously, Jamilov et al. (2024) pursue a similar strategy and find similar results.

6 Quantification

We now combine our model and empirical results to measure the quantitative effects of contagious beliefs on business cycles and decompose the mechanisms underlying this.

6.1 Estimating the Model

In Section 3.6, we showed that we could estimate the model in three steps. We now combine our empirical estimates from Section 5 with this three-step approach to estimate the model. We provide the point estimates of model parameters in Table 5 and provide additional details in Appendix F.

Step I: Estimation of the Effect of Optimism. To estimate 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). 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.

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) and Foster, Haltiwanger, and Syverson (2008) 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 $\gamma = 0$. Third, we calibrate the inverse Frisch elasticity of labor supply at $\psi = 0.4$ based on 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 $\omega = 0.49$. In Section 6.2, we study the sensitivity of our results to this external calibration, and we introduce two other estimation strategies for complementarity: using estimates of demand multipliers from the literature and inferring a demand multiplier for optimism using our own firm-level regressions.

Step II: Estimation of Updating Rules. To estimate the parameters of the LAC updating rules, we use the linear probability model estimated in Table 3.[16] This yields values of $u = 0.208$ for stubbornness, $r = 0.804$ for associativeness, and $s = 0.290$ for contagiousness.

[16]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% (holding fixed ε_t), *i.e.,* there is a five-standard-deviation optimism shock.

Table 5: 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	Contagiousness	0.290
	σ_ε	Std. dev. of the optimism shock	0.044

Notes: "Fixed" parameters are externally set. "Calibrated" parameters are chosen to hit empirical targets. For more details, see Section 6.1.

Step III: Estimation of Private Information and the Shock Processes. To estimate the extent of private information and the persistence and volatility of productivity shocks, we showed that we need to estimate the model-implied ARMA(1,1) process for fundamental output. Now that we have estimated f, we can compute fundamental output as:

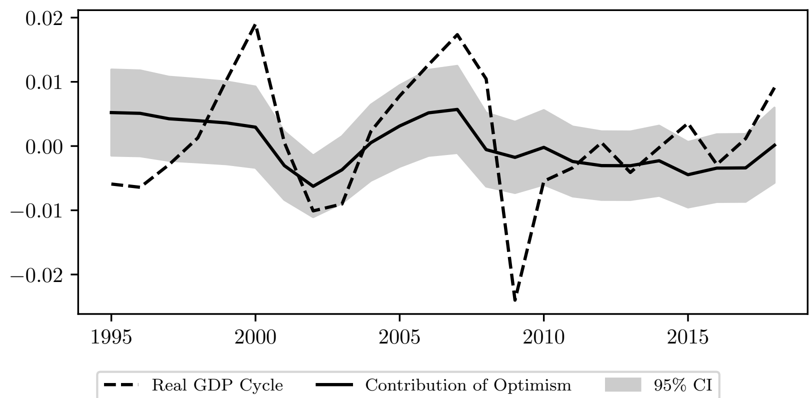
$$\log Y_t^f = \log Y_t - f(Q_t) \tag{36}$$

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 A1), and f as our calibrated function.[17] 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 $\mu_O - \mu_P$. Finally, we estimate the variance of optimism shocks, σ_ε^2, to match the time-series variance of optimism.

The Estimated Model Features Almost i.i.d. Shocks. Before proceeding to the quantitative results, we observe an important property of the estimated model: our point estimate for the persistence of exogenous productivity shocks is $\rho = 0.086$. As we have only allowed for i.i.d. optimism shocks, this means that our model only requires *almost* i.i.d. exogenous shocks to match the time-series properties of output. Thus, our estimates imply

[17] 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 6: 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 estimation of δ^{OP} using the delta method.

that contagious beliefs generate strong internal propagation. This represents an important difference between our theory of model dynamics and theories based on learning and dispersed information (see *e.g.,* Woodford, 2003a; Lorenzoni, 2009; Angeletos and La'O, 2010), all of which require exogenously persistent fundamentals about which agents slowly learn.

6.2 How Does Optimism Shape the Business Cycle?

Using the calibrated model, we now study the effects of optimism on the business cycle via two complementary approaches: (i) gauging the historical effect of swings in business optimism on US GDP and (ii) exploring the full dynamic implications of contagious and associative optimism.

The Effects of Optimism on US GDP. In our empirical exercise, which leveraged cross-sectional data on US firms' optimism, the general-equilibrium effect of optimism on total production was the unidentified "missing intercept." Now, equipped with the model calibration of general-equilibrium forces, we can return to the question of how changes in optimism have historically affected the US business cycle. We calculate the time series of $f(Q_t)$, where f is the calibrated function mapping aggregate optimism to aggregate output, which depends on the partial-equilibrium effect of optimism on hiring, returns-to-scale, and the demand multiplier, and Q_t is the observed annual time series for business optimism. We take the observed time path of aggregate optimism as given, and therefore use the estimated dynamics of optimism only to determine the shocks that rationalize this observed path.

Figure 6 illustrates our findings by plotting 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). 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. The decline in the optimism component of GDP explains 31.65% (SE: 2.68%) of the output loss between 2000 and 2002 and 18.06% (SE: 1.53%) of the output loss between 2007 and 2009.

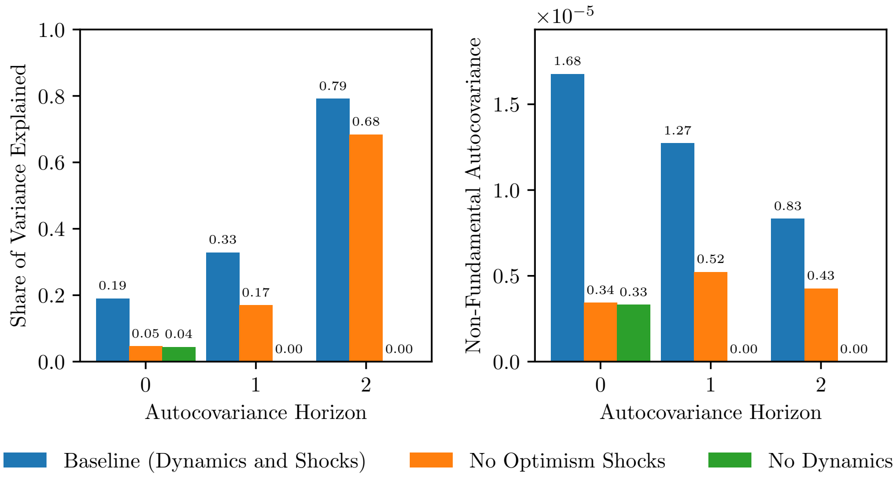
To unpack the model-implied causes of the historical business cycle, we plot the sequence of fundamental output and optimism shocks that our model requires to match the realized optimism and output time series in Figure A6. Our model accounts for the early 2000s recession with a large negative optimism shock ($\varepsilon_{2001} = -0.08$, or -1.8 standard deviations in our calibration) and a moderate-sized shock to fundamental output. For the Great Recession, our model implies a larger shock to fundamentals along with a smaller optimism shock ($\varepsilon_{2008} = -0.06$ or -1.4 standard deviations). The larger contribution of, and shock to, optimism at the outset of the early 2000s recession is consistent with a story that a break in confidence, associated with the "dot com" crash in the stock market, spurred a recession despite sound economic fundamentals. This is further consistent with independent textual evidence that "crash narratives" in financial news were especially rampant in this period (Goetzmann, Kim, and Shiller, 2022).

Contagious Models and Economic Fluctuations. We now fully describe the role of model dynamics in shaping the business cycle via the estimated process for how optimism spreads. To produce a summary statistic for the contribution of optimism toward the covariance structure of output, we observe that the covariance of output at lag $\ell \geq 0$ can be decomposed into four terms:

$$\begin{aligned} \text{Cov}[\log Y_t, \log Y_{t-\ell}] = \text{Cov}[\log Y_t^f, \log Y_{t-\ell}^f] + \text{Cov}[f(Q_t), f(Q_{t-\ell})] \\ + \text{Cov}[f(Q_t), Y_{t-\ell}] + \text{Cov}[f(Q_{t-\ell}), Y_t] \end{aligned} \tag{37}$$

The first term captures the volatility and persistence of exogenous fundamentals (*i.e.,* the driving productivity shocks). The second term captures the volatility and persistence of the non-fundamental component of output. The last two terms capture the relationship of optimism with past and future fundamentals, which arises from the co-evolution of the prevalence of models with economic outcomes. We therefore define non-fundamental variance as the total autocovariance arising from endogenous optimism as the sum of the last three

Figure 7: The Contribution of Optimism to Output Variance



Notes: The left panel plots the fraction of variance, one-year autocovariance, and two-year autocovariance explained by endogenous optimism in model simulations. The right panel plots the total non-fundamental autocovariance. Both quantities are defined in Equation 38. In each figure, we plot results under three model scenarios: the baseline model with optimism shocks and optimism dynamics (blue), a variant model with no optimism shocks, or $\sigma_\varepsilon^2 = 0$ (orange), and a variant model with shocks but no dynamics for model spread, or $u = r = s = 0$ (green).

terms, as well as its fraction of total variance, at each lag ℓ:

$$\begin{aligned} \text{Non-Fundamental Autocovariance}_\ell &= \text{Cov}[\log Y_t, \log Y_{t-\ell}] - \text{Cov}[\log Y_t^f, \log Y_{t-\ell}^f] \\ \text{Share of Variance Explained}_\ell &= \frac{\text{Non-Fundamental Autocovariance}_\ell}{\text{Cov}[\log Y_t, \log Y_{t-\ell}]} \end{aligned} \tag{38}$$

We calculate these statistics at horizons $\ell \in \{0, 1, 2\}$ and under three model variants: the baseline model with optimism shocks, a variant model which turns off the shocks (or sets $\sigma_\varepsilon^2 = 0$), and a variant model that keeps optimism shocks but shuts down the endogenous evolution of models (by setting $u = r = s = 0$).[18]

Optimism explains 19% of contemporary variance ($\ell = 0$), and this fraction increases with the lag (Figure 7). At one-year and two-year lags, optimism explains 33% and 79% of output autocovariance, respectively. Thus, most medium-frequency (two-year) dynamics are produced by contagious optimism instead of fundamentals. The model without endogenous dynamics of optimism explains only 4% of output variance and, as optimism shocks are i.i.d., 0% of output auto-covariance. Moreover, while the model without optimism shocks matches only 5% of output variance, it accounts for 17% and 69% of one-year and two-year output

[18] As discussed in Appendix F, we always add a constant to LAC updating so 0.5 is the interior steady-state when output is at its steady state. Thus, the "no dynamics" variant sets $Q_{t+1} = 0.5 + \epsilon_t$.

autocovariance. Interestingly, the separate contributions to output variance of shocks and endogenous dynamics sum to less than one-half of their joint explanatory power. This result establishes that the contagiousness and associativeness of models are amplifying propagation mechanisms for exogenous sentiment shocks.

Sensitivity Analysis. In Table A16, we report a sensitivity analysis of the conclusions above to different calibrations for the macroeconomic parameters. 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 $\frac{1}{1-\omega} = 1.96$. In rows 1, 2, 3, and 4 we vary the multiplier by: (i) adjusting the inverse-Frisch elasticity to 2.5 to match microeconomic estimates (Peterman, 2016), (ii) allowing for greater income effects in labor supply $\gamma = 1$, (iii) matching the empirical estimates of the demand multiplier of 1.33 from Becko, Flynn, and Patterson (2024), and (iv) estimating the general equilibrium multiplier semi-structurally by using the extent of omitted variables bias from omitting a time fixed effect in the regression of hiring on optimism (see Appendix F.4 for the details). Our numerical results from adjusting the multiplier, holding fixed $(\delta^{OP}, \alpha, \epsilon)$, convey that the contribution of optimism is increasing in this number. We finally consider sensitivity to the calibrations of the elasticity of substitution ϵ (row 5 of Table A16) and the returns-to-scale α (row 6 of Table A16) 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 hiring.

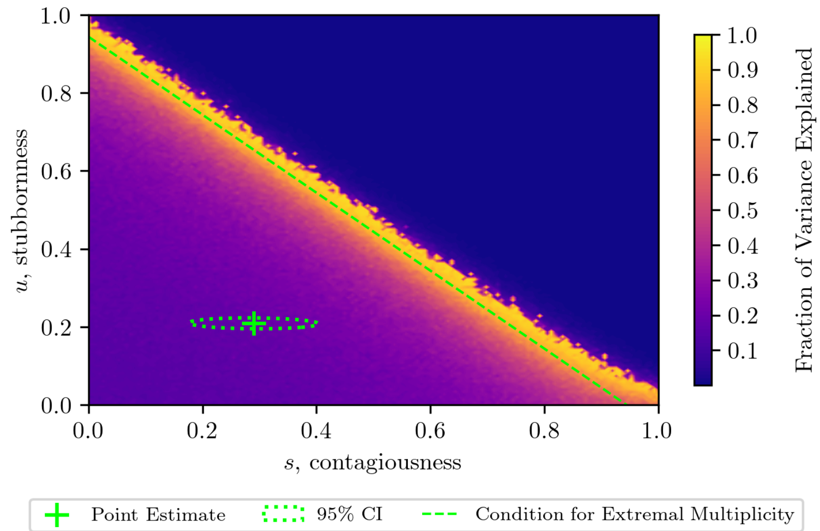
6.3 Can Contagious Optimism Generate Hysteresis?

We have shown that the dynamics of optimism generate quantitatively significant business cycles. However, we have not yet explored the implications of contagious models for hysteresis and long-run movements in output. Our theoretical analysis delimited two qualitatively different regimes for macroeconomic dynamics with contagious optimism: one with stochastic fluctuations around a stable steady state, and one with hysteresis and (almost) global convergence to extreme steady states. Are models contagious enough to generate hysteresis?

For the LAC case which we have taken to the data, the necessary and sufficient condition for extremal multiplicity is given by Equation 22. We compute the empirical analog of this condition:

$$\hat{M} = \hat{u} + \hat{s} + \hat{r}\frac{\alpha}{1-\omega}\hat{\delta}^{OP} - 1 \tag{39}$$

Figure 8: Variance Decomposition for Different Values of Stubbornness and Contagiousness



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 Share of Variance Explained$_0$ defined in Equation 38. The plus is our calibrated value from Table 5, and the dotted line is the boundary of a 95% confidence set. The dashed line is the condition of extremal multiplicity from Corollary 4 and Equation 22.

If $\hat{M} > 0$, the calibrated model features hysteresis in the dynamics of optimism and output; if $\hat{M} < 0$, the model features oscillations around a stable steady state. We find $\hat{M} = -0.44 < 0$ with a standard error of 0.052, implying stable oscillations and ruling out hysteresis dynamics. This reflects the fact that decision-relevance, stubbornness, contagiousness, and associativeness are sufficiently small for optimism.

We explore the sensitivity of this conclusion to our calibration of the two parameters to which it is most sensitive: stubbornness and contagiousness. In Figure 8, we plot our point estimate of contagiousness and stubbornness as a plus and its 95% confidence interval as a dotted ellipse. We also plot, as a dashed line, the condition for $M = 0$; to the left of this line, $M < 0$, and to the right of this line, $M > 0$. In the Figure, we shade the fraction of variance explained by non-fundamental optimism. Given the statistical precision in the estimates of stubbornness and contagiousness, we are confident that contagious models contribute stable fluctuations to the economy and explain about 20% of the variance in output. To reverse this and enter the regime of extremal multiplicity ($M > 0$) would require, for example, about 2.5 times the contagiousness that we observe.

How does extremal multiplicity interact with our model's predictions for non-fundamental volatility? To isolate the role of endogenous propagation, our theoretical discussion of extremal multiplicity considered paths of the economy without shocks. In the quantitative model, the economy is constantly buffeted with shocks that move optimism away from its steady state(s). Near the condition for extremal multiplicity, non-fundamental variance reaches essentially 100% of total variance. This is because even small shocks have the potential to "go viral," and the force pulling the economy toward an interior steady state (*i.e.*, balanced optimism and pessimism) is weak.[19]

Finally, far to the right of the extremal multiplicity condition, contagious optimism explains little output variance. This is because the economy quickly settles into an extreme steady state, fully optimistic or fully pessimistic, and moves quickly back to this steady state in response to shocks. Thus, the "M test" provides an accurate diagnostic for whether economic models can "go viral" even in the presence of shocks.

6.4 Emergent Optimism in a Multidimensional World

In our main analysis, we restricted attention to a case in which agents' models of the world are one-dimensional: firm managers are either optimistic or pessimistic about their overall economic prospects. Consequently, our perspective on how models spread was focused on the dynamics of overall optimism. Of course, in practice, managers hold many different views on a wide variety of topics, and their adoption of each individual view may be subject to contagious and associative dynamics. This perspective reminds of Shiller's (2020) hypothesis that *constellations* of many small and semantically related narratives reinforce one another to create stronger economic and social effects, and that the *confluence* of seemingly unrelated narratives may explain business-cycle fluctuations. We now explore this idea in an extension of our analysis.

Model. We first describe an enriched model in which managers hold a different *view* regarding each of many underlying *topics*, and these views together determine their overall optimism. There is a latent space of K topics. Agents either do or do not hold a (binary) view about each topic, and we denote individuals' views by $\lambda_{it} = (\lambda_{1,it}, \ldots, \lambda_{K,it}) \in \{0,1\}^K$. We let $Q_t^k = \int_0^1 \lambda_{k,it} \, \mathrm{d}i \in [0,1]$ denote the share of population that adopts each view.

Optimism emerges from the confluence of many views. To model this tractably, we assume that the aggregate fraction of optimists, Q_t, depends linearly on the fraction of agents adopting each view:

$$Q_t = \left[\sum_{k=1}^{K} \zeta^k Q_t^k \right]_0^1 \tag{40}$$

[19] Due to the presence of shocks to optimism, this prediction is symmetric around the extremal multiplicity threshold; in the variant model which turns off optimism shocks, the extremal multiplicity condition sharply delineates the regime in which optimism fluctuations contribute to output variance from the regime in which there is complete hysteresis (Figure A7).

where $(\zeta^k)_{k=1}^K$ are weights controlling the marginal effect of each view on emergent optimism.

Each manager's view about each topic k evolves via a linear-associative-contagious process. That is, we let (P_1^k, P_0^k) respectively denote functions returning the probability that an agent who currently does or does not hold view k at time t holds the view at time $t+1$:

$$\begin{aligned} P_1^k(\log Y, Q, \varepsilon) &= \left[\frac{u^k}{2} + r^k \log Y + s^k Q^k + \varepsilon^k\right]_0^1 \\ P_0^k(\log Y, Q, \varepsilon) &= \left[-\frac{u^k}{2} + r^k \log Y + s^k Q^k + \varepsilon^k\right]_0^1 \end{aligned} \tag{41}$$

We allow for k-specific stubbornness, associativeness, and contagiousness, as well as independent shocks $\varepsilon^k \sim N(0, \sigma^2_{\varepsilon,k})$. The dynamics of different views interact through associativeness. For example, views may contribute to optimism, boosting the economy, and thereby indirectly promoting other views associated with a good economy.

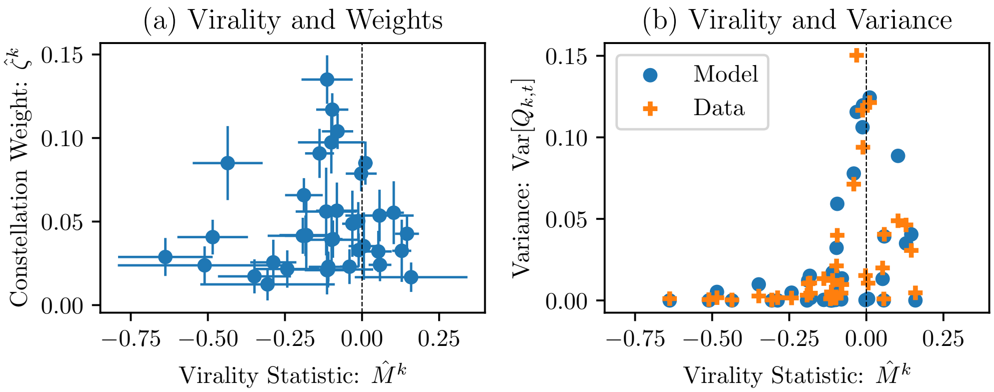
The rest of the model is the same as the baseline. Thus, while dynamics are the same *conditional* on the process for optimism, the process for emergent optimism through the latent evolution of views may differ.

Estimation. We employ two strategies to measure topics and managers' views. The first is a partially supervised method that detects firms' discussion of the nine *Perennial Economic Narratives* described by Shiller (2020). The second is an unsupervised Latent Dirichlet Allocation model (Blei et al., 2003), which flexibly identifies clusters of topics discussed by firms. We describe the details behind these approaches in Appendix F.3. We estimate the weights ζ by running the following firm-level regression:

$$\text{opt}_{it} = \sum_{k=1}^{9} \zeta^k_{\text{Shiller}} \cdot \text{Shiller}^k_{it} + \sum_{k=1}^{100} \zeta^k_{\text{topic}} \cdot \text{topic}^k_{it} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{42}$$

This model estimates the marginal effects of each granular topic on the propensity toward optimism. As throughout our analysis, we control for firm fixed effects and non-parametric sector-by-time trends. We apply the Rigorous Square-Root post-LASSO method of Belloni, Chen, Chernozhukov, and Hansen (2012) to account for the likely fact that not all extracted topics are relevant for optimism. Applying this method yields a relevant subset of 30 LDA topics and 8 Shiller topics. Table A22 in the Appendix prints each of the selected topics and their respective ζ^k. Next, we estimate stubbornness, associativeness, and contagiousness for each topic just as in the main analysis, by estimating variants of Equation 33. This step fixes the parameters (u^k, r^k, s^k) for each selected topic. These estimates are also reported in Table A22. We calibrate the variance of view shocks, $\sigma^2_{\varepsilon,k}$, to match the time-series variance of each

Figure 9: The Viral Components of Emergent Optimism



Notes: Panel (a) plots our estimates of the virality statistic $\hat{M}^k$, defined in Equation 43, against our estimates of the constellation weights $\hat{\zeta}^k$, from Equation 42. The solid lines are 95% confidence intervals. Panel (b) plots our estimated virality statistics against their simulated variances (blue circles) and their empirical time-series variances (orange crosses).

granular topic. Specifically, we minimize the sum of square deviations of model-generated variances from measured time-series variances. Finally, to calibrate the rest of the model, we proceed exactly as described in Section 6.1.

Results. Comparing the model with granular topics to the baseline, we find that emergent optimism explains a comparable amount of the variance and autocorrelation of output. For example, optimism explains 16% of output variance and 31% of the first-lag autocovariance, compared to 19% and 33% in our baseline calibration (Figure A10).

However, this similarity belies significant heterogeneity in how the granular topics spread, which is in turn related to each topic's tendency to "go viral." To assess this, we observe that the topic-specific M statistics,

$$\hat{M}^k = \hat{u}^k + \hat{s}^k + \hat{r}^k \frac{\alpha \hat{\zeta}^k \hat{\delta}^{OP}}{1-\omega} - 1 \tag{43}$$

correspond to the correct hysteresis test statistic if topic k were the only component of emergent optimism. Intuitively, $\hat{M}^k$ captures each granular topic's "tendency toward virality."

We find that many topics have M statistics that exceed (or nearly exceed) the criticality threshold of zero (see Panel (a) of Figure 9). These topics, unlike aggregate optimism, can therefore go viral. Moreover, emergent optimism places large weights on many of these viral topics (see Panel (a) of Figure 9). Thus, aggregate optimism is significantly driven by viral topics. Finally, the topics that our model predicts as being close to the threshold ($M^k = 0$) are precisely the highest-variance topics in the data (see Panel (b) of Figure 9). This provides empirical validation of the M statistic as a diagnostic for virality.

Taken together, we find that emergent optimism is largely driven by viral and volatile topics. But, despite the virality of its underlying components, emergent optimism is stable and its effect on the business cycle is almost unchanged relative to our baseline model.

7 Conclusion

This paper studies the macroeconomic implications of contagious beliefs. We first introduce a real business cycle model in which competing models of the world gain and lose prevalence based on their match with reality (associativeness) and their existing prevalence (contagiousness). Contagious optimism can generate non-fundamentally driven boom-bust cycles and hysteresis. To take this model to the data, we extract firms' sentiment from their language in regulatory reports and earnings calls. We find that contagious and associative optimism affects firms' decisions and beliefs without representing news about fundamentals. When we calibrate the model to match the data, we find 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. Finally, we show that the interaction of many simultaneously evolving and highly contagious topics, some of which are individually prone to hysteresis, can nevertheless underlie stable fluctuations in emergent optimism and output. Taken together, our analysis shows that belief contagion may underlie many important features of the business cycle.

Our analysis leaves open at least two important areas for future study. First, we have analyzed how firms' contagious beliefs matter and abstracted away from studying similar dynamics on the household side, which may influence spending and saving. Moreover, coevolving models on both the "supply side" and the "demand side" of the economy might have mutually reinforcing effects. Second, there remains much more to study about what makes a model contagious. Probing these deeper origins of models and their relationship to economic narratives could help further account for the full economic, semantic, and psychological interactions among economic agents who are trying to make sense of a complex world.

References

George-Marios Angeletos and Jennifer La'O. Noisy business cycles. *NBER Macroeconomics Annual*, 24(1):319–378, 2010.

George-Marios Angeletos and Jennifer La'O. Sentiments. *Econometrica*, 81(2):739–779, 2013.

George-Marios Angeletos, Fabrice Collard, and Harris Dellas. Quantifying confidence. *Econometrica*, 86(5):1689–1726, 2018.

Costas Azariadis. Self-fulfilling prophecies. *Journal of Economic Theory*, 25(3):380–396, 1981.

Robert B Barsky and Eric R Sims. Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review*, 102(4):1343–1377, 2012.

Susanto Basu and John G Fernald. Returns to scale in US production: Estimates and implications. *Journal of Political Economy*, 105(2):249–283, 1997.

Marianne Baxter and Robert G King. Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4):575–593, 1999.

John S Becko, Joel P Flynn, and Christina Patterson. Fiscal policy in a networked economy. Working Paper 29619, National Bureau of Economic Research, 2024.

Alexandre Belloni, Daniel Chen, Victor Chernozhukov, and Christian Hansen. Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429, 2012.

Jess Benhabib and Roger EA Farmer. Indeterminacy and increasing returns. *Journal of Economic Theory*, 63(1):19–41, 1994.

Jess Benhabib, Pengfei Wang, and Yi Wen. Sentiments and aggregate demand fluctuations. *Econometrica*, 83(2):549–585, 2015.

Francesco Bianchi, Cosmin Ilut, and Hikaru Saijo. Diagnostic business cycles. *Review of Economic Studies*, 91(1):129–162, 2024.

Olivier J Blanchard, Jean-Paul L'Huillier, and Guido Lorenzoni. News, noise, and fluctuations: An empirical exploration. *American Economic Review*, 103(7):3045–3070, 2013.

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, pages 993–1022, 2003.

J Aislinn Bohren and Daniel N Hauser. Learning with heterogeneous misspecified models: Characterization and robustness. *Econometrica*, 89(6):3025–3077, 2021.

Michele Boldrin and Michael Woodford. Equilibrium models displaying endogenous fluctuations and chaos: a survey. *Journal of Monetary Economics*, 25(2):189–222, 1990.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Memory, attention, and choice. *The Quarterly journal of economics*, 135(3):1399–1442, 2020.

Pedro Bordalo, Nicola Gennaioli, Andrei Shleifer, and Stephen J Terry. Real credit cycles. Working Paper 28416, National Bureau of Economic Research, 2021.

Pedro Bordalo, Nicola Gennaioli, Rafael La Porta, Matthew O'Brien, and Andrei Shleifer. Long-term expectations and aggregate fluctuations. *NBER Macroeconomics Annual*, 38 (1):311–347, 2024.

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

Craig Burnside, Martin Eichenbaum, and Sergio Rebelo. Understanding booms and busts in housing markets. *Journal of Political Economy*, 124(4):1088–1147, 2016.

Ricardo J Caballero and Alp Simsek. A risk-centric model of demand recessions and speculation. *The Quarterly Journal of Economics*, 135(3):1493–1566, 2020.

Christopher D Carroll. The epidemiology of macroeconomic expectations. Working Paper 8695, National Bureau of Economic Research, 2001.

Christopher D. Carroll and Tao Wang. Epidemiological expectations. In *Handbook of Economic Expectations*, volume 1, page Forthcoming. Elsevier, 2022.

David Cesarini, Erik Lindqvist, Matthew J Notowidigdo, and Robert Östling. The effect of wealth on individual and household labor supply: evidence from Swedish lotteries. *American Economic Review*, 107(12):3917–46, 2017.

Ryan Chahrour, Kristoffer Nimark, and Stefan Pitschner. Sectoral media focus and aggregate fluctuations. *American Economic Review*, 111(12):3872–3922, December 2021.

Lawrence J Christiano, Cosmin L Ilut, Roberto Motto, and Massimo Rostagno. Monetary policy and stock market boom-bust cycles. *ECB working paper*, 2008.

Russell Cooper and Andrew John. Coordinating coordination failures in keynesian models. *The Quarterly Journal of Economics*, 103(3):441–463, 1988.

Jan De Loecker, Jan Eeckhout, and Gabriel Unger. The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2):561–644, 2020.

Stefano Eusepi and Bruce Preston. Expectations, learning, and business cycle fluctuations. *American Economic Review*, 101(6):2844–72, October 2011.

Lucia Foster, John Haltiwanger, and Chad Syverson. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1): 394–425, 2008.

Xavier Gabaix and Ralph SJ Koijen. Granular instrumental variables. Working Paper 28204, National Bureau of Economic Research, 2020.

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

Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. Expectations and investment. *NBER Macroeconomics Annual*, 30(1):379–431, 2016.

Richard J Gentry, Joseph S Harrison, Timothy J Quigley, and Steven Boivie. A database of ceo turnover and dismissal in s&p 1500 firms, 2000–2018. *Strategic Management Journal*, 42(5):968–991, 2021.

William N Goetzmann, Dasol Kim, and Robert J Shiller. Crash narratives. Working Paper 30195, National Bureau of Economic Research, July 2022.

Joao Guerreiro. Belief disagreement and business cycles. *Manuscript, February*, 2023.

J Michael Harrison and David M Kreps. Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics*, 92(2):323–336, 1978.

Tarek Alexander Hassan, Stephan Hollander, Aakash Kalyani, Laurence van Lent, Markus Schwedeler, and Ahmed Tahoun. Economic surveillance using corporate text. Technical report, National Bureau of Economic Research, 2024.

Nir Jaimovich and Sergio Rebelo. Can news about the future drive the business cycle? *American Economic Review*, 99(4):1097–1118, 2009.

Rustam Jamilov, Alexandre Kohlhas, Oleksandr Talavera, and Mao Zhang. Granular sentiments. Working paper, 2024.

Fuwei Jiang, Joshua Lee, Xiumin Martin, and Guofu Zhou. Manager sentiment and stock returns. *Journal of Financial Economics*, 132(1):126–149, 2019.

Michael Jacob Kahana. *Foundations of human memory*. Oxford University Press, United States, 2012.

Markku Kaustia and Ville Rantala. Common analysts: method for defining peer firms. *Journal of Financial and Quantitative Analysis*, 56(5):1505–1536, 2021.

John Maynard Keynes. *The General Theory of Employment, Interest and Money*. Palgrave Macmillan, 1936.

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

José María Liberti and Mitchell A Petersen. Information: Hard and soft. *Rev. of Corporate Finance Studies*, 8(1):1–41, 2019.

Guido Lorenzoni. A theory of demand shocks. *American Economic Review*, 99(5):2050–84, 2009.

Tim Loughran and Bill McDonald. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65, 2011.

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

Albert Marcet and Thomas J Sargent. Convergence of least-squares learning in environments with hidden state variables and private information. *Journal of Political Economy*, 97(6): 1306–1322, 1989a.

Albert Marcet and Thomas J Sargent. Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, 48(2):337–368, 1989b.

Edward M Miller. Risk, uncertainty, and divergence of opinion. *The Journal of finance*, 32 (4):1151–1168, 1977.

Pooya Molavi. Macroeconomics with learning and misspecification: A general theory and applications. Working paper, Northwestern University, 2019.

Pooya Molavi, Alireza Tahbaz-Salehi, and Andrea Vedolin. Model complexity, expectations, and asset prices. Technical report, National Bureau of Economic Research, 2021.

Stephen Nickell. Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426, 1981.

Kristoffer P. Nimark. Man-bites-dog business cycles. *American Economic Review*, 104(8): 2320–67, August 2014.

Emily Oster. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204, 2019.

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

Arthur Cecil Pigou. *Industrial Fluctuations*. Macmillan and Company, London, 1923.

Edouard Schaal and Mathieu Taschereau-Dumouchel. Herding through booms and busts. *Journal of Economic Theory*, 210:105669, 2023.

SEC. Investor bulletin: How to read a 10-k. Article, Securities and Exchange Commission Office of Investor Education and Advocacy, 2011. Available at: https://www.sec.gov/files/reada10k.pdf.

Robert J Shiller. Narrative economics. *American Economic Review*, 107(4):967–1004, 2017.

Robert J Shiller. *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press, 2020.

Mirko Wiederholt. Empirical properties of inflation expectations and the zero lower bound. *manuscript, Goethe University*, 2015.

Michael Woodford. Imperfect common knowledge and the effects of monetary policy. In Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford, editors, *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, pages 25–58. Princeton University Press, 2003a.

Michael Woodford. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, 2003b.

Appendices

Contents

A Omitted Derivations and Proofs

A.1 Proof of Theorem 1

Proof. We guess and verify that there exists a unique quasi-loglinear equilibrium. That is, there exists a unique equilibrium of the following form:

$$\log Y(\theta_t, \theta_{t-1}, Q_t) = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t) \tag{44}$$

for some parameters $a_0, a_1, a_2 \in \mathbb{R}$ and function $f : [0,1] \to \mathbb{R}$. 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 tuple (a_0, a_1, a_2, f).

From the arguments in the main text, we have Equation 12 holds. Thus, 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 weight λ_{it}, we have that the distribution of the aggregate component of productivity is:

$$\log \theta_t | s_{it}, \lambda_{it} \sim N\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s}\right) \tag{45}$$

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

$$\mu(\lambda_{it}, \theta_{t-1}) = (1-\rho)(\mu_O \lambda_{it} + \mu_P(1-\lambda_{it})) + \rho \log \theta_{t-1}\,,\ \kappa = \frac{1}{1+\frac{\sigma_e^2}{\sigma_\theta^2}}\,,\ \sigma^2_{\theta|s} = \frac{1}{\frac{1}{\sigma_\theta^2}+\frac{1}{\sigma_e^2}} \tag{46}$$

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

$$\log \theta_{it} | s_{it}, \lambda_{it} \sim N\left(\log \gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) \tag{47}$$

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:

$$\begin{aligned} \log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right] &= \log \mathbb{E}_{it}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right] \\ &= -\frac{1+\psi}{\alpha}\mu_{it} + \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \eta^2 \end{aligned} \tag{48}$$

Under our conjecture (Equation 44), we can moreover compute:

$$\begin{aligned}\log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right] &= \log \mathbb{E}_{it}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)(a_0 + a_1\log\theta_t + a_2\log\theta_{t-1} + f(Q_t)\right\}\right] \\ &= \left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1(\mu_{it} - \log\gamma_i) + a_2\log\theta_{t-1} + f(Q_t)\right] \\ &\quad + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\left[\eta^2 - \sigma_{\tilde{\theta}}^2\right]\end{aligned} \tag{49}$$

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

$$\begin{aligned}\log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\mu_{it} - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\eta^2 \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1(\mu_{it}-\log\gamma_i) + a_2\log\theta_{t-1} + f(Q_t)\right] + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\left[\eta^2-\sigma_{\tilde{\theta}}^2\right]\Bigg]\end{aligned} \tag{50}$$

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:

$$\begin{aligned}\log Y_t &= \log\left[\left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}\right] \\ &= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}\right] \\ &= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right]\end{aligned} \tag{51}$$

Moreover, expanding the terms in Equation 50, we have that:

$$\begin{aligned}\log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\ &+ \frac{1+\psi}{\alpha}\left[\log\gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it},\theta_{t-1})\right] \\ &- \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma_{\theta|s}^2 + \sigma_{\tilde{\theta}}^2\right) \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it},\theta_{t-1})\right) + a_2\log\theta_{t-1} + f(Q_t)\right] \\ &+ \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma_{\theta|s}^2\Bigg]\end{aligned} \tag{52}$$

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:

$$\begin{aligned}
\delta_t(\lambda_{it}) = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \Bigg[& \log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\
& + \frac{1+\psi}{\alpha}\left[\mu_\gamma + \kappa \log\theta_t + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right] - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) \\
& + \left(\frac{1}{\epsilon} - \gamma\right)\left[a_0 + a_1\left(\kappa\log\theta_t + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + a_2 \log\theta_{t-1} + f(Q_t)\right] \\
& + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon} - \gamma\right)^2 \sigma^2_{\theta|s}\Bigg]
\end{aligned} \tag{53}$$

and:

$$\hat{\sigma}^2 = \left(\frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}\right)^2 \left[\left(\frac{1+\psi}{\alpha}\right)^2 \sigma^2_\gamma + \kappa^2\left[\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon} - \gamma\right)\right]^2 \sigma^2_e\right] \tag{54}$$

Thus, we have that:

$$\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right] = \exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(\lambda_{it}) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} \tag{55}$$

and so:

$$\begin{aligned}
\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right] &= Q_t \exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(1) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} \\
+ (1-Q_t)\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(0) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} & \\
= \left[Q_t\exp\left\{\frac{\epsilon-1}{\epsilon}(\delta_t(1) - \delta_t(0))\right\} + (1-Q_t)\right]\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_t(0) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} &
\end{aligned} \tag{56}$$

Yielding:

$$\log Y_t = \delta_t(0) + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 + \frac{\epsilon}{\epsilon-1}\log\left(Q_t\exp\left\{\frac{\epsilon-1}{\epsilon}(\delta_t(1)-\delta_t(0))\right\} + (1-Q_t)\right) \tag{57}$$

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

$$\delta_t(1) - \delta_t(0) = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}\left(\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon} - \gamma\right)\right)(1-\kappa)(1-\rho)(\mu_O - \mu_P) = \alpha\delta^{OP} \tag{58}$$

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

$$
\begin{aligned}
\delta_t(0) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \Bigg[\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left(\mu_\gamma + (1-\kappa)((1-\rho)\mu_P + \rho\log\theta_{t-1}\right) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \left(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}}\right) \\
+ \left(\frac{1}{\epsilon}-\gamma\right)\left(a_0 + a_1(1-\kappa)((1-\rho)\mu_P+\rho\log\theta_{t-1})\right) + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s} \\
+ \left[\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon}-\gamma\right)\right]\kappa\log\theta_t + \left(\frac{1}{\epsilon}-\gamma\right)\left(a_2\log\theta_{t-1} + f(Q_t)\right)\Bigg]
\end{aligned} \tag{59}
$$

By matching coefficients between Equations 57 and Equation 44, we obtain a_0, a_1, a_2, and f.

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

$$
a_1 = \frac{\left[\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon}-\gamma\right)\right]\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \tag{60}
$$

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:

$$
a_1 = \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1-\frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} = \frac{1}{1-\kappa\omega}\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \tag{61}
$$

which is in terms of primitive parameters and is moreover positive.

Second, we match coefficients on $\log\theta_{t-1}$ to obtain an equation for a_2:

$$
a_2 = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\left[\left(\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon}-\gamma\right)a_1\right)(1-\kappa)\rho + \left(\frac{1}{\epsilon}-\gamma\right)a_2\right] \tag{62}
$$

This implies that:

$$
\begin{aligned}
a_2 &= \frac{1}{1-\omega}\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon}-\gamma\right)a_1\right](1-\kappa)\rho \\
&= \frac{1}{1-\omega}\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon}-\gamma\right)\frac{1}{1-\kappa\omega}\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\right](1-\kappa)\rho
\end{aligned} \tag{63}
$$

which is in terms of primitive parameters.

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

$$f(Q) = \frac{\frac{1}{\epsilon} - \gamma}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} f(Q) + \frac{\epsilon}{\epsilon - 1} \log\left(1 + Q\left[\exp\left\{\frac{\epsilon - 1}{\epsilon}\alpha\delta^{OP}\right\} - 1\right]\right) \tag{64}$$

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:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \log\left(1 + Q\left[\exp\left\{\frac{\epsilon - 1}{\epsilon}\alpha\delta^{OP}\right\} - 1\right]\right) \tag{65}$$

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:

$$a_0 = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left[\log\left(\frac{1 - \frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left(\mu_\gamma + (1-\kappa)(1-\rho)\mu_P\right) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) \right.$$
$$\left. + \left(\frac{1}{\epsilon} - \gamma\right)\left(a_0 + a_1(1-\kappa)(1-\rho)\mu_P\right) + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon} - \gamma\right)^2 \sigma^2_{\theta|s}\right] + \frac{1}{2}\frac{\epsilon - 1}{\epsilon}\hat{\sigma}^2 \tag{66}$$

Solving this equation yields:

$$a_0 = \frac{1}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \left[\frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}\left[\log\left(\frac{1 - \frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left(\mu_\gamma + (1-\kappa)(1-\rho)\mu_P\right) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) \right.\right.$$
$$\left.\left. + \left(\frac{1}{\epsilon} - \gamma\right) a_1(1-\kappa)(1-\rho)\mu_P + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon} - \gamma\right)^2 \sigma^2_{\theta|s}\right] + \frac{1}{2}\frac{\epsilon - 1}{\epsilon}\hat{\sigma}^2\right] \tag{67}$$

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, a_2 and f uniquely and verified that there exists a unique quasi-loglinear equilibrium. Finally, to obtain the formula for the best reply of agents, simply substitute a_0, a_1, a_2 and f into Equation 52 and label the coefficients as in the claim. □

A.2 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-loglinear 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 $e_{it} \in [-M, M]$.

Lemma 1. *When fundamentals are bounded, there exists a unique equilibrium*

Proof. To this end, we can recast any equilibrium function $\log Y(\theta, \theta_{-1}, Q)$ as one that solves the fixed point in Equation 12. 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:

$$V_M(g)(\theta, \theta_{-1}, Q) = \frac{\epsilon}{\epsilon - 1} \log \mathbb{E}_{(\theta,\theta_{-1},Q)} \left[\exp \left\{ \frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\log \left(\frac{1 - \frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}} \right) \right. \right. \right.$$
$$\left. \left. \left. - \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ -\frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] + \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \right) \right\} \right] \tag{68}$$

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:

$$\log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \geq \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) h \right\} \right] \tag{69}$$

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:

$$\log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) (g + a) \right\} \right] = \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] + \left(\frac{1}{\epsilon} - \gamma \right) a \tag{70}$$

And so:

$$
\begin{aligned}
V_M(g+a)(\theta,\theta_{-1},Q) &= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_{(\theta,\theta_{-1},Q)}\left[\exp\left\{\frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\left(\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right)\right.\right.\right. \\
&\left.\left.\left. -\log \mathbb{E}_{(s,Q)}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right]+\log \mathbb{E}_{(s,Q)}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)g\right\}\right]+\left(\frac{1}{\epsilon}-\gamma\right)a\right)\right\}\right] \\
&= V_M(g)(\theta,\theta_{-1},Q)+\omega a
\end{aligned}
\tag{71}
$$

where $\omega \in [0,1)$ by assumption. Note that the modulus of contraction ω is precisely the claimed strategic complementarity parameter in Equation 8. 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-loglinear equilibrium in the unbounded fundamentals case is an appropriately-defined ε-equilibrium for any $\varepsilon > 0$. Let the unique quasi-loglinear equilibrium we have guessed and verified be $\log Y^*$. We say that g is a ε-equilibrium if

$$||g - V_M(g)||_p < \varepsilon \tag{72}$$

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 2. *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 68. 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:

$$\lim_{M\to\infty} ||\log Y^* - V_M(\log Y^*)||_p = 0 \tag{73}$$

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

$$||\log Y^* - V_M(\log Y^*)||_p < \varepsilon \quad \forall M \in \mathbb{N} : M > \bar{M} \tag{74}$$

Completing the proof. □

A.3 Proof of Theorem 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 convex and compact set to itself. 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 $T_\theta(1) = P_O(a_o + (a_1 + a_2)\log\theta + f(1), 1, 0) = 1$ and $Q = 0$ is a steady state for θ if and only if $T_\theta(0) = P_P(a_0 + (a_1 + a_2)\log\theta, 0, 0) = 0$. Thus, $Q = 1$ is a steady state if and only if $P_O^{-1}(1;1) \leq a_0 + (a_1 + a_2)\log\theta + f(1)$ and $Q = 0$ is a steady state if and only if $P_P^{-1}(0;0) \geq a_0 + (a_1 + a_2)\log\theta$. 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:

$$\begin{aligned} T_\theta'(Q) = {} & P_O(a_0 + (a_1 + a_2)\log\theta + f(Q), Q, 0) - P_P(a_0 + (a_1 + a_2)\log\theta + f(Q), Q, 0) \\ & + Q\frac{\mathrm{d}}{\mathrm{d}Q}P_O(a_0 + (a_1 + a_2)\log\theta + f(Q), Q, 0) \\ & + (1 - Q)\frac{\mathrm{d}}{\mathrm{d}Q}P_P(a_0 + (a_1 + a_2)\log\theta + f(Q), Q, 0) \end{aligned} \tag{75}$$

Thus, for $\theta < \theta_P$ and $Q = 0$:

$$\begin{aligned} T_\theta'(0) &= P_O(a_0 + (a_1 + a_2)\log\theta, 0, 0) - P_P(a_0 + (a_1 + a_2)\log\theta, 0, 0) \\ &\quad + \frac{\mathrm{d}}{\mathrm{d}Q}P_P(a_0 + (a_1 + a_2)\log\theta + f(Q), Q, 0)\mid_{Q=0} \\ &= P_O(a_0 + (a_1 + a_2)\log\theta, 0, 0) \end{aligned} \tag{76}$$

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 $P_O(a_0 + (a_1 + a_2)\log\theta, 0, 0) < 1$. Moreover, for $\theta < \theta_P$, we have that: $P_O(a_0 + (a_1 + a_2)\log\theta, 0, 0) \leq P_O(a_0 + (a_1 + a_2)\log\theta_P, 0, 0) = P_O(P_P^{-1}(0;0), 0, 0)$. Thus, a sufficient condition for $T_\theta'(0) < 1$ for $\theta < \theta_P$ is that $P_O(P_P^{-1}(0;0), 0, 0) < 1$.

For $\theta > \theta_O$ and $Q = 1$, we have that:

$$\begin{aligned} T'_\theta(1) &= P_O(a_0 + (a_1 + a_2)\log\theta + f(1), 1, 0) - P_P(a_0 + (a_1 + a_2)\log\theta + f(1), 1, 0) \\ &\quad + \frac{\mathrm{d}}{\mathrm{d}Q} P_O(a_0 + (a_1 + a_2)\log\theta + f(1), 1, 0) \mid_{Q=1} \\ &= 1 - P_P(a_0 + (a_1 + a_2)\log\theta + f(1), 1, 0) \end{aligned} \tag{77}$$

where the second equality follows by observing that $P_O = 1$ and both $\frac{\partial P_O}{\partial \log Y}$ and $\frac{\partial P_O}{\partial Q}$ are zero for $\theta > \theta_O$. Hence, we have that $T'_\theta(1) < 1$ when $P_P(a_0 + (a_1 + a_2)\log\theta + f(1), 1, 0) > 0$. For $\theta > \theta_O$ we have that $P_P(a_0 + (a_1 + a_2)\log\theta + f(1), 1, 0) \geq P_P(a_0 + (a_1 + a_2)\log\theta_O + f(1), 1) = P_P(P_O^{-1}(1, 1), 1, 0)$. Thus, a sufficient condition for $T'_\theta(1) < 1$ for $\theta > \theta_O$ is that $P_P(P_O^{-1}(1, 1), 1, 0) > 0$. □

A.4 Proof of Corollary 5

Proof. From Equation 52 in the proof of Proposition 1, we have that the log production of firm i at time t is described in the unique quasi-log-linear equilibrium by:

$$\begin{aligned} \log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \Bigg[& \log\left(\frac{1 - \frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left[\log\gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right] \\ & - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) + + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma\right)^2 \sigma^2_{\theta|s} \\ & + \left(\frac{1}{\epsilon} - \gamma\right)\left[a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + a_2 \log\theta_{t-1} + f(Q_t)\right] \Bigg] \end{aligned} \tag{78}$$

We substitute this expression into the production function to obtain an equation for hiring $\log L_{it} = \frac{1}{\alpha}(\log x_{it} - \log\theta_{it})$. Subtracting lagged labor from both sides yields Equation 25. □

B Additional Theoretical Results and Extensions

This appendix covers several additional results and model extensions. First, we provide formal results on the model's impulse response functions and its propensity to undergo boom-bust cycles (B.1). Second, we theoretically characterize and quantify the normative implications of fluctuations (B.2). Third, we study equilibrium dynamics under a benchmark model of Bayesian model updating and contrast these predictions with those obtained in our main analysis (B.3). Fourth, fifth, sixth, and seventh we extend the baseline model to respectively incorporate a continuum of different levels of optimism (B.4), models about idiosyncratic fundamentals (B.5), multi-dimensional models (B.6), and model updating that depends on idiosyncratic fundamentals (B.7). In each case, we characterize equilibrium dynamics and show how our main theoretical insights extend. Eighth, we show how endogenous cycles and chaotic dynamics can obtain when agents are contrarian and implement an empirical test for their presence (B.8).

B.1 Impulse Responses and Stochastic Fluctuations

This Appendix generalizes and formalizes the observations about contagious business cycle dynamics from Section 3.3.

First, we define two important types 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 1 respectively illustrate examples of SSC-A and SSC-B transition maps.

Lemma 3 (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. 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 $\{T_\theta^n(Q_0)\}_{n\in\mathbb{N}}$. For a steady state Q^*, its basin of attraction is:

$$\mathcal{B}_\theta(Q^*) = \left\{ Q_0 \in [0,1] : \lim_{n\to\infty} T_\theta^n(Q_0) = Q^* \right\} \tag{79}$$

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

$$T_\theta^{n+1}(Q_0) = T_\theta \circ T_\theta^n(Q_0) \geq T_\theta \circ T_\theta^{n-1}(Q_0) = T_\theta^n(Q_0) \tag{80}$$

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. 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 $[0,1] \setminus \mathcal{B}_\theta(\hat{Q}_\theta)$, 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:

$$T_\theta^{n+1}(Q_0) = T_\theta \circ T_\theta^n(Q_0) \leq T_\theta \circ T_\theta^{n-1}(Q_0) = T_\theta^n(Q_0) \tag{81}$$

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 $Q_0^\infty = \lim_{n\to\infty} T_\theta^n(Q_0)$. Thus, we have that $Q_0^\infty = 0$. Hence, $[0, \hat{Q}_\theta) \subseteq \mathcal{B}_\theta(0)$. Second, consider $Q_0 = \hat{Q}_\theta$. 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 $Q_0 \in (\hat{Q}_\theta, 1]$. 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. □

In the SSC-A case there is a unique, (almost) globally stable steady state (left panel of Figure 1). 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 1). These two classes delineate two qualitatively different regimes for models dynamics: one with stable model convergence around a long-run steady state (SSC-A) and one with a strong role for initial conditions and hysteresis (SSC-B).

We now study how the economy responds to deterministic and stochastic fundamental and optimism shocks. For this analysis, we restrict attention to the SSC class, noting that this is an assumption solely on primitives.[20]

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 $\theta = 1$:

$$\theta_t = \begin{cases} 1, & t = 0, \\ \hat{\theta}, & t = 1, \\ 1, & t \geq 2. \end{cases} \tag{82}$$

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

[20]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.

the IRFs in the shock size $\hat{\theta}$. For simplicity, we focus on the case with i.i.d. productivity shocks in which $\rho = 0$.

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 1 (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:*

$$\log Y_t = \begin{cases} a_0 + f(\hat{Q}_1), & t = 0, \\ a_0 + a_1 \log\hat{\theta} + f(\hat{Q}_1), & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \qquad Q_t = \begin{cases} \hat{Q}_1, & t \leq 1, \\ Q_2, & t = 2, \\ T_1(\log Y_{t-1}, Q_{t-1}), & t \geq 3. \end{cases} \tag{83}$$

Moreover, $Q_2 = \hat{Q}_1 P_O(a_0 + a_1\log\hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) + (1-\hat{Q}_1)P_P(a_0 + a_1\log\hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) > \hat{Q}_1$*,* 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. By Proposition 1 and substituting the form of the shock process from Equation 82, we obtain the formula for the output IRF. For the fraction of optimists, we see that:

$$\begin{aligned} Q_2 &= \hat{Q}_1 P_O(a_0 + a_1\log\hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) + (1-\hat{Q}_1)P_P(a_0 + a_1\log\hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) \\ &> \hat{Q}_1 P_O(a_0 + f(\hat{Q}_1), \hat{Q}_1, 0) + (1-\hat{Q}_1)P_P(a_0 + f(\hat{Q}_1), \hat{Q}_1, 0) = \hat{Q}_1 \end{aligned} \tag{84}$$

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 3 as we are in the SSC-A case. The hump shape is obtained if $\log Y_1 \leq \log Y_2$. This corresponds to

$$\log Y_1 = a_0 + a_1\log\hat{\theta} + f(\hat{Q}_1) \leq a_0 + f(Q_2) = \log Y_2 \tag{85}$$

which rearranges to the desired expression. □

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 2 (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:

$$\log Y_t = \begin{cases} a_0, & t = 0, \\ a_0 + a_1 \log \hat{\theta}, & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \qquad Q_t = \begin{cases} 0, & t \leq 1, \\ P_P(a_0 + a_1 \log \hat{\theta}, 0, 0), & t = 2, \\ T_1(\log Y_{t-1}, Q_{t-1}), & t \geq 3. \end{cases} \tag{86}$$

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 $\hat{\theta} < \exp\{f(P_P(a_0 + a_1 \log \hat{\theta}, 0, 0))/a_1\}$.

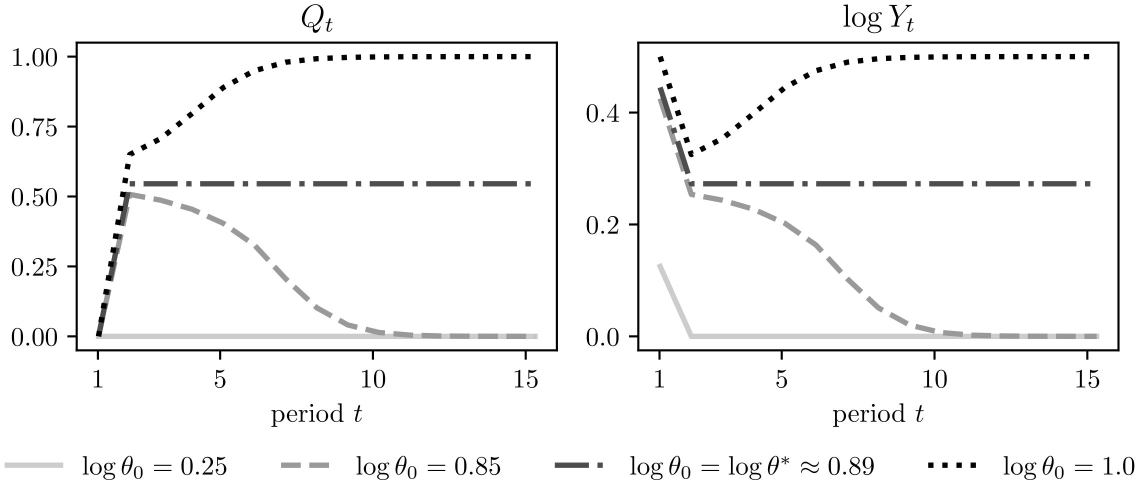
Proof. We first derive the IRF functions. The formula for the output IRF follows Proposition 1. For the IRF for the fraction of optimists, we simply observe that $Q_0 = Q_1 = 0$ and $Q_2 = P_P(a_0 + a_1 \log \hat{\theta}, 0, 0)$, and that $Q_t = T_1(Q_{t-1})$ for $t \geq 3$ by iterating forward.

We now describe the properties of the IRFs as a function of the size of the initial shock $\hat{\theta}$. First, observe that $Q_2 = P_P(a_0 + a_1 \log \hat{\theta}, 0, 0)$. Thus, we have that $Q_2 = 0$ 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 3, 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. □

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,

Figure 10: 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 LAC updating. The four initial conditions correspond to the four cases of Proposition 2.

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 2 and illustrated numerically in Figure 10. First, if a shock is small and no agent is moved toward optimism, the shock has a one-period impact on aggregate output. Second, if some agents are moved to optimism by the transitory boost to output but this fraction lies below the criticality threshold, then output steadily declines back to its pessimistic steady-state level as optimism was not sufficiently great to be self-fulfilling. Third, in the knife-edge case, optimism moves to a new (unstable) steady state and permanently increases output. Fourth, when enough agents are moved to optimism by the initial boost to output, then the economy converges to the fully optimistic steady state and optimism is completely self-fulfilling.

The impulse responses to optimism shocks are identical to those described above. One can take the formulas in Propositions 1 and 2 from $t \geq 2$ and set Q_2 equal to the value of Q that obtains following the optimism shock ε. It follows that the qualitative nature of the impulse response to an optimism shock is identical to that of a fundamental shock.

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 as it is hit by fundamental and optimism shocks. For simplicity, we once again restrict to the case of i.i.d. fundamentals, in which $\rho = 0$.

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:

$$T_{PO} = \mathbb{E}\left[\min\{\tau \in \mathbb{N} : Q_\tau = 1\} | Q_0 = 0\right], \; T_{OP} = \mathbb{E}\left[\min\{\tau \in \mathbb{N} : Q_\tau = 0\} | Q_0 = 1\right] \tag{87}$$

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 models under consideration, and that of the optimism shocks G.

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

Proposition 3 (Period of Boom-Bust Cycles)**.** *The expected regime-switching times satisfy the following inequalities:*

$$\begin{aligned} T_{PO} &\leq \frac{1}{1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^\dagger(1;0,\varepsilon) - a_0}{a_1}\right\}\right)\right]} \\ T_{OP} &\leq \frac{1}{\mathbb{E}_G\left[H\left(\exp\left\{\frac{P_O^\dagger(0;1,\varepsilon) - a_0 - f(1)}{a_1}\right\}\right)\right]} \end{aligned} \tag{88}$$

where $P_P^\dagger(x; Q, \varepsilon) = \inf\{Y : P_P(Y, Q, \varepsilon) = x\}$ *and* $P_O^\dagger(x; Q, \varepsilon) = \sup\{Y : P_O(Y, Q, \varepsilon) = x\}$*. Moreover, when* $P_O^\dagger(0; 1, 0) - P_P^\dagger(1; 0, 0) \leq f(1)$*, these bounds are tight in the sense that they are attained for some processes for fundamentals and optimism shocks* (H, G)*.*

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

$$\overline{Q}_{t+1} = \mathbb{I}[T(\overline{Q}_t, \nu_t) = 1] \tag{89}$$

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 $t = 1$:

$$\overline{Q}_1 = \mathbb{I}[T(0, \nu_0) = 1] \leq T(0, \nu_0) = Q_1 \tag{90}$$

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

$$\overline{Q}_t = \mathbb{I}[T(\overline{Q}_{t-1}, \nu_{t-1}) = 1] \leq \mathbb{I}[T(Q_{t-1}, \nu_{t-1}) = 1] \leq T(Q_{t-1}, \nu_{t-1}) = Q_t \tag{91}$$

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

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

$$\overline{\tau}_{PO} = \min\{t \in \mathbb{N} : \overline{Q}_t = 1, \overline{Q}_0 = 0\} \geq \min\{t \in \mathbb{N} : Q_t = 1, Q_0 = 0\} = \tau_{PO} \tag{92}$$

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:

$$\mathcal{G}_{PO} = \{(0^{(n-1)}, 1)\} : n \geq 1\} \tag{93}$$

Moreover, conditional on $\overline{Q}_{t-1} = 0$, the distribution of $\overline{Q}_t$ is independent of $\{\nu_s\}_{s\leq t-1}$. 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:

$$\begin{aligned}\mathbb{P}[Q_{t+1} = 1 | Q_t = 0] &= \mathbb{P}\left[P_P(a_0 + a_1 \log\theta_t, 0, \varepsilon_t) = 1\right] \\ &= \mathbb{P}\left[\theta_t \geq \exp\left\{\frac{P_P^\dagger(1; 0, \varepsilon_t) - a_0}{a_1}\right\}\right] \\ &= 1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^\dagger(1; 0, \varepsilon) - a_0}{a_1}\right\}\right)\right]\end{aligned} \tag{94}$$

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

$$\tau_{PO} \prec_{FOSD} \overline{\tau}_{PO} \sim \text{Geo}\left(1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^\dagger(1; 0, \varepsilon) - a_0}{a_1}\right\}\right)\right]\right) \tag{95}$$

An immediate corollary is that:

$$T_{PO} = \mathbb{E}[\tau_{PO}] \leq \mathbb{E}[\overline{\tau}_{PO}] = \frac{1}{1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^\dagger(1;0,\varepsilon) - a_0}{a_1}\right\}\right)\right]} \tag{96}$$

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

$$\underline{Q}_{t+1} = \mathbb{I}[T(\underline{Q}_t, \nu_t) \neq 0] \tag{97}$$

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 $\{\nu_t\}_{t\in\mathbb{N}}$. And so, we have that if $\underline{Q}_t$ has reached 0 then so too has Q_t. The possible sample paths in this case are:

$$\mathcal{G}_{OP} = \{(1^{(n-1)}, 0)\} : n \geq 1\} \tag{98}$$

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

$$\begin{aligned}\mathbb{P}[Q_{t+1} = 0 | Q_t = 1] &= \mathbb{P}\left[\theta_t \leq \exp\left\{\frac{P_O^\dagger(0;1,\varepsilon_t) - a_0 - f(1)}{a_1}\right\}\right] \\ &= \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_O^\dagger(0;1,\varepsilon) - a_0 - f(1)}{a_1}\right\}\right)\right]\end{aligned} \tag{99}$$

And so we have:

$$T_{OP} \leq \frac{1}{\mathbb{E}_G\left[H\left(\exp\left\{\frac{P_O^\dagger(0;1,\varepsilon) - a_0 - f(1)}{a_1}\right\}\right)\right]} \tag{100}$$

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:

$$\Theta^* = \left(-\infty, \exp\left\{\frac{P_O^\dagger(0;1,0) - a_0 - f(1)}{a_1}\right\}\right] \cup \left[\exp\left\{\frac{P_P^\dagger(1;0,0) - a_0}{a_1}\right\}, \infty\right) \tag{101}$$

and suppose that θ takes values only in this set, where the two sub-intervals are disjoint as $P_O^\dagger(0;1,0) - P_P^\dagger(1;0,0) \leq f(1)$. Moreover, suppose that optimism shocks equal zero with probability one. 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 $\mu_A < \exp\left\{\frac{P_O^\dagger(0;1,0) - a_0 - f(1)}{a_1}\right\}$ and $\mu_B > \exp\left\{\frac{P_P^\dagger(1;0,0) - a_0}{a_1}\right\}$. Taking the limit as $\sigma \to 0$, the support of H converges to being contained within Θ^*. □

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, G) such that the fictitious processes coincide always with the true processes.[21]

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

B.2 Welfare Implications

In this appendix, we derive the normative implications of contagious models 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 4 (Models and Welfare)**.** *For any path of aggregate optimism $\mathbf{Q} = \{Q_t\}_{t=0}^\infty$, aggregate welfare is given by*

$$\begin{aligned}\mathcal{U}(\mathbf{Q}) = U_C^* \sum_{t=0}^\infty \beta^t \exp\left\{(1-\gamma)f(Q_t)\right\} \\ - U_L^* \sum_{t=0}^\infty \beta^t \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \exp\left\{(1+\psi)d_3 f(Q_t)\right\}\end{aligned} \tag{102}$$

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

$$\frac{U_C^*}{U_L^*} \times \frac{\exp\left\{(1-\gamma)f(1)\right\} - 1}{\exp\left\{(1+\psi)(d_2 + d_3 f(1))\right\} - 1} > 1 \tag{103}$$

[21] 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, G) for which the bound is attained by taking weighted averages of the optimistic and pessimistic models, making the uncertainty under each sufficiently small, and eliminating optimism shocks.

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

$$\tau^* = \exp\left\{(1-\omega)\left(\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}\right)f(1)\right\} - 1 \tag{104}$$

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

$$\mathcal{U}(\mathbf{Q}) = \sum_{t=0}^{\infty}\beta^t\left(\mathbb{E}_H\left[\frac{C_t(Q_t,\theta_t)^{1-\gamma}}{1-\gamma}\right] - \mathbb{E}_H\left[\int_{[0,1]}\frac{L_{it}(\gamma_i,s_{it},Q_t)^{1+\psi}}{1+\psi}\mathrm{d}i\right]\right) \tag{105}$$

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:

$$\begin{aligned}
\mathbb{E}_H\left[\frac{C_t^{1-\gamma}(Q_t,\theta_t)}{1-\gamma}\right] &= \mathbb{E}_H\left[\frac{1}{1-\gamma}\exp\left\{(1-\gamma)\log Y(Q_t,\theta)\right\}\right] \\
&= \mathbb{E}_H\left[\frac{1}{1-\gamma}\exp\left\{(1-\gamma)\left(a_0 + a_1\log\theta + f(Q_t)\right)\right\}\right] \\
&= \frac{1}{1-\gamma}\exp\left\{(1-\gamma)\left(a_0 + a_1\mu_H + f(Q_t)\right) + \frac{1}{2}a_1^2\sigma_H^2\right\} \\
&= \frac{1}{1-\gamma}\exp\left\{(1-\gamma)\left(a_0 + a_1\mu_H\right) + \frac{1}{2}a_1^2\sigma_H^2\right\}\exp\left\{(1-\gamma)f(Q_t)\right\} \\
&= U_C^*\exp\left\{(1-\gamma)f(Q_t)\right\}
\end{aligned} \tag{106}$$

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

$$L_{it} = d_1\log\theta_t + d_2\lambda_{it} + d_3 f(Q_t) + v_{it} \tag{107}$$

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

$$\begin{aligned}
&\int_{[0,1]}\frac{L_{it}(\gamma_i,s_{it},Q_t)^{1+\psi}}{1+\psi}\mathrm{d}i \\
&= \frac{1}{1+\psi}\left(Q_t\exp\{(1+\psi)d_2\} + (1-Q_t)\right) \\
&\quad\times\exp\left\{(1+\psi)(d_1\log\theta + \mu_v + d_3 f(Q_t)) + \frac{1}{2}(1+\psi)^2\sigma_v^2\right\}
\end{aligned} \tag{108}$$

Hence, the expectation over θ is given by:

$$
\begin{aligned}
&\mathbb{E}_H\left[\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1+\psi} \mathrm{d}i\right] \\
&= \frac{1}{1+\psi}\left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \\
&\quad \times \exp\left\{(1+\psi)d_3 f(Q_t)\right\} \exp\left\{(1+\psi)(d_1\mu_H + \mu_v) + \frac{1}{2}(1+\psi)^2(\sigma_v^2 + d_1^2\sigma_H^2)\right\} \\
&= U_L^*\left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \exp\left\{(1+\psi)d_3 f(Q_t)\right\}
\end{aligned}
\tag{109}
$$

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

$$
\begin{aligned}
\mathcal{U}(\mathbf{Q}) = U_C^* &\sum_{t=0}^{\infty} \beta^t \exp\left\{(1-\gamma) f(Q_t)\right\} \\
&- U_L^* \sum_{t=0}^{\infty} \beta^t \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \exp\left\{(1+\psi)d_3 f(Q_t)\right\}
\end{aligned}
\tag{110}
$$

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:

$$
\begin{aligned}
\log x_{it} = &\frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) - \log \mathbb{E}_{it}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right] \right. \\
&\left. + \log \mathbb{E}_{it}\left[\exp\left\{\left(\frac{1}{\epsilon} - \gamma\right)\log Y_t\right\}\right] \right) + \Xi(\tau)
\end{aligned}
\tag{111}
$$

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-loglinear equilibrium, where:

$$
\log Y(\theta, \tau) = a_0 + a_1 \log\theta + \frac{1}{1-\omega}\Xi(\tau) \tag{112}
$$

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

$$
\log x_{it}(\tau) = \log x_{it}(0) + \frac{1}{1-\omega}\Xi(\tau) \tag{113}
$$

Which implies that:

$$\log L_{it}(\tau) = \log L_{it}(0) + \frac{1}{\alpha}\frac{1}{1-\omega}\Xi(\tau) \tag{114}$$

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

$$\begin{aligned}\mathcal{U}(\tau) = U_C^* \sum_{t=0}^{\infty} \beta^t \exp\left\{(1-\gamma)\frac{1}{1-\omega}\Xi(\tau)\right\} \\ - U_L^* \sum_{t=0}^{\infty} \beta^t \exp\left\{(1+\psi)d_3\frac{1}{1-\omega}\Xi(\tau)\right\}\end{aligned} \tag{115}$$

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

$$\mathcal{U}(1) = \mathcal{U}(\tau^*) \tag{116}$$

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

$$\tau^* = \exp\left\{(1-\omega)\left(\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}\right)f(1)\right\} - 1 \tag{117}$$

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, intermediate 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 4 can be directly applied in our numerical calibration from Section 6 to calculate the welfare effects of optimism without approximation. We calculate the average payoff of the representative household under three scenarios. The first corresponds to the calibrated model dynamics in simulation, under the assumption that the pessimistic model is correctly specified.[22] 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

[22]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 contagious 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 contagious optimism suggests that, in our macroeconomic calibration, losses from inducing misallocation are more than compensated by level increases in output.

B.3 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. For simplicity, we focus on the case of i.i.d. shocks ($\rho = 0$). Formally, this corresponds to the following law of motion for Q_t:

$$Q_{t+1} = \int_{[0,1]} \mathbb{P}_i[\mu = \mu_O | \{\log Y_j, Q_j\}_{j=0}^t] \, \mathrm{d}i \tag{118}$$

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 5 (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 i.i.d., 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:*

$$\log Y_t = \begin{cases} a_0 + a_1 \log\theta_t & \text{if } a < 0, \\ a_0 + a_1 \log\theta_t + f(1) & \text{if } a > 0. \end{cases} \tag{119}$$

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:

$$\theta_0 = \frac{\log Y_0 - a_0 - f(Q_0)}{a_1} \tag{120}$$

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 $\{Y_j\}_{j \le t}$ (and not $\{Q_j\}_{j \le t}$) by computing:

$$\theta_t = \frac{\log Y_t - a_0 - f(Q_t)}{a_1} \tag{121}$$

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

$$\lambda_{i,t+1} = \mathbb{P}_i[\mu = \mu_O | \{\theta_j\}_{j=1}^t] = \lambda_{i,t+1} = \mathbb{P}_i[\mu = \mu_O | \{Y_j\}_{j=1}^t] \tag{122}$$

By application of Bayes rule, we obtain:

$$\lambda_{i,t+1} = \mathbb{P}[\mu = \mu_O | \theta_t, \lambda_{i,t}] = \frac{f_O(\theta_t)\lambda_{i,t}}{f_O(\theta_t)\lambda_{i,t} + f_P(\theta_t)(1 - \lambda_{i,t})} \tag{123}$$

which implies that:

$$\begin{aligned} \frac{\lambda_{i,t+1}}{1 - \lambda_{i,t+1}} &= \frac{f(\log \theta_t | \mu = \mu_O)}{f(\log \theta_t | \mu = \mu_P)} \frac{\lambda_{i,t}}{1 - \lambda_{i,t}} \\ &= \exp\left\{ \frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{\sigma^2} \right\} \frac{\lambda_{i,t}}{1 - \lambda_{i,t}} \end{aligned} \tag{124}$$

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:

$$\begin{aligned} \Omega_{i,t+1} &= \Omega_{i,t} + \frac{(\log \theta_t - \mu_P)^2 - (\log \theta_t - \mu_O)^2}{\sigma^2} \\ &= \Omega_{it} + a + \xi_t \end{aligned} \tag{125}$$

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:

$$\frac{\Omega_{i,t}}{t} = \frac{1}{t}\Omega_{i,0} + \frac{t-1}{t} a + \frac{1}{t} \sum_{i=1}^{t} \xi_i \to^{a.s.} a \tag{126}$$

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. □

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 A1), and moreover, in the long run, rules out the features of macroeconomic dynamics that we derive in Section 3 as consequences of the endogenous evolution of optimism.

B.4 Continuous Models

Our main analysis featured two levels of optimism. However, much of our analysis generalizes to a setting with a continuum of levels of optimism. For expositional simplicity, in this section, we abstract from optimism shocks and assume that productivity is i.i.d ($\rho = 0$). The model is as in Section 2, but now $\mu \in [\mu_P, \mu_O]$ and the distribution of models 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 models 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 6 (Equilibrium Characterization with Continuous Models)**.** *There exists a quasi-loglinear equilibrium:*

$$\log Y(\log \theta_t, Q_t) = a_0 + a_1 \log \theta_t + f(Q_t) \tag{127}$$

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

$$\mathrm{d}Q_{t+1}(\mu') = \int_{\mu_P}^{\mu_O} P_{\mu'}(\mu, a_0 + a_1 \log \theta_t + f(Q_t), Q_t) \mathrm{d}Q_t(\mu) \tag{128}$$

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:

$$\log Y_t = \delta_t(\mu_P) + \frac{1}{2}\frac{\epsilon - 1}{\epsilon}\hat{\sigma}^2 + \frac{\epsilon}{\epsilon - 1} \log \left(\int_{\mu_P}^{\mu_O} \exp\left\{ \frac{\epsilon - 1}{\epsilon}(\delta_t(\tilde{\mu}) - \delta_t(\mu_P)) \right\} \mathrm{d}Q_t(\tilde{\mu}) \right) \tag{129}$$

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:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \log\left(\int_{\mu_P}^{\mu_O} \exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{\tilde{\mu}-\mu_P}{\mu_O-\mu_P}\right\} \mathrm{d}Q(\tilde{\mu})\right) \tag{130}$$

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 6, 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 models as:

$$K_Q(\tau) = \log\left(\mathbb{E}_Q[\exp\{\tau\tilde{\mu}\}]\right) \tag{131}$$

We therefore have that $\log\left(\mathbb{E}_Q[\exp\{\tau(\tilde{\mu}-z)\}]\right) = K_Q(\tau) - \tau z$. It follows by Equation 130 that:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\omega}\left[K_Q\left(\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{1}{\mu_O-\mu_P}\right) - \frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{\mu_P}{\mu_O-\mu_P}\right] \tag{132}$$

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

$$K_Q(\tau) = \mu_Q\tau + O(\tau^2) \tag{133}$$

We therefore have that:

$$f(Q) = \frac{1}{1-\omega}\alpha\delta^{OP}\frac{\mu_Q-\mu_P}{\mu_O-\mu_P} + O\left(\left(\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{1}{\mu_O-\mu_P}\right)^2\right) \tag{134}$$

We now can express the static, general equilibrium effects in terms of mean of the model 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 model 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 contagiousness 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:

$$\begin{aligned} \mathrm{d}Q_{t+1}(\mu') &= \int_{\mu_P}^{\mu_O} P_{\mu'}(a_0 + a_1 \log\theta_t + f(\mu_{Q,t}), \mu_{Q,t})\mathrm{d}Q_t(\mu) \\ &= P_{\mu'}(a_0 + a_1 \log\theta_t + f(\mu_{Q,t}), \mu_{Q,t}) \end{aligned} \tag{135}$$

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

$$\mu_{Q,t+1} = T(\mu_{Q,t}, \theta_t) = \int_{\mu_P}^{\mu_O} \mu' P_{\mu'}(a_0 + a_1 \log\theta_t + f(\mu_{Q,t}), \mu_{Q,t})\mathrm{d}\mu' \tag{136}$$

Which is simply a continuous state analog of the difference equation expressed in Corollary 3 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:

$$\begin{aligned} \hat{P}^{-1}(x; \mu_Q) &= \sup\{Y : P(Y, Q) = \delta_x\} \\ \check{P}^{-1}(x; \mu_Q) &= \inf\{Y : P(Y, Q) = \delta_x\} \end{aligned} \tag{137}$$

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

$$\theta_O = \exp\left\{\frac{\check{P}^{-1}(\mu_O; \mu_O) - a_0 - f(1)}{a_1}\right\},\ \theta_P = \exp\left\{\frac{\hat{P}^{-1}(\mu_P; \mu_P) - a_0 - f(1)}{a_1}\right\} \tag{138}$$

The following result establishes that these thresholds characterize extremal multiplicity:

Proposition 7 (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*

$$\check{P}^{-1}(\mu_O; \mu_O) - \hat{P}^{-1}(\mu_P; \mu_P) \leq f(1) \tag{139}$$

Proof. This follows exactly the same steps as the proofs of Proposition 2 and Corollary 4, replacing the appropriate inverses defined above. $\square$

Thus, the same conditions that give rise to multiplicity with binary models 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 7 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.5 Models About Idiosyncratic Fundamentals

In the main analysis, we assumed that models described properties of aggregate fundamentals. In this section, we characterize equilibrium dynamics when models describe properties of idiosyncratic fundamentals. For expositional simplicity, we suppose that productivity shocks are i.i.d. (or $\rho = 0$). 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 $\log \tilde{\theta}_{jt} \sim N(0, \sigma_{\tilde{\theta}}^2)$ for all $j \neq i$. However, agents disagree about the mean of their own idiosyncratic productivity: optimistic agents believe that $\log \tilde{\theta}_{it} \sim N(\mu_O, \sigma_{\tilde{\theta}}^2)$ while pessimistic agents believe that $\log \tilde{\theta}_{it} \sim N(\mu_P, \sigma_{\tilde{\theta}}^2)$. The rest of the model is identical.

In this context, dynamics are identical conditional on the static relationship between output and models. Moreover, the static relationship between output and models 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 8 (Equilibrium Characterization with Models About Idiosyncratic Fundamentals)**.** *There exists a unique equilibrium such that:*

$$\log Y(\log \theta_t, Q_t) = \tilde{a}_0 + a_1 \log \theta_t + \tilde{f}(Q_t) \tag{140}$$

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*

$$\tilde{f}(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \log\left(1 + Q\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\tilde{\delta}^{OP}\right\} - 1\right]\right) \tag{141}$$

where $\tilde{\delta}^{OP}$ is defined in Equation 142.

Proof. The proof follows exactly the steps of the proof of Proposition 1 where the aggregate model 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 49 differs in that $\mu_{it} = \log \gamma_i + \kappa s_{it}$. Tracking this through to Equation 53, 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 58, at which point we have that the exact formula for δ^{OP} changes and is now given by:

$$\alpha\tilde{\delta}^{OP} = \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}(1-\kappa)(\mu_O - \mu_P) \tag{142}$$

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 models 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.6 Multi-Dimensional Models

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

$$\log\theta_t = (1-\rho)\mu + \rho\log\theta_{t-1} + \sigma\nu_t \tag{143}$$

with $\nu_t \sim N(0,1)$ and i.i.d.. Models now correspond to a vector of (μ, ρ, σ), indexing the mean, persistence and variance of the process for fundamentals. The set of models 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 model at a time, and let $Q_{t,k}$ be the fraction of agents who believe model $(\mu_k, \rho_k, \sigma_k)$ at time t. Finally, we assume that agents face the same signal-to-noise ratio κ, regardless of the model that they hold.[23] Together, these assumptions ensure that agents' posteriors are normal and place a common weight on models 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 9 (Equilibrium Characterization with Multi-Dimensional Models and Persistence)**.** *There exists a quasi-loglinear equilibrium:*

$$\log Y(\log\theta_t, \log\theta_{t-1}, Q_t) = a_0 + a_1\log\theta_t + a_2\log\theta_{t-1} + f(Q_t, \theta_{t-1}) \tag{144}$$

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

$$Q_{t+1,k} = \sum_{k'\in\mathcal{K}} Q_{t,k'} P_{k'}(k, a_0 + a_1\log\theta_t + a_2\log\theta_{t-1} + f(Q_t,\theta_{t-1}), Q_t) \tag{145}$$

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:

$$\log Y(\log\theta_t, \log\theta_{t-1}, Q_t) = a_0 + a_1\log\theta_t + a_2\log\theta_{t-1} + f(Q_t, \theta_{t-1}) \tag{146}$$

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 model weights. Let E denote the standard basis for $\mathbb{R}^K$ with k-th basis vector denoted by

$$e_k = \{\underbrace{0,\ldots,0}_{k-1}, 1, \underbrace{0,\ldots,0}_{K-k}\} \tag{147}$$

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

We have that $\lambda_{it} = e_k$ for some $k \leq K$. Under this model loading, we have that agent's posteriors are given by:

$$\log \theta_{it} | \lambda_{it}, s_{it} \sim N\left(\log \gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \tag{148}$$

with:

$$\begin{aligned} \mu(e_k, \theta_{t-1}) &= (1-\rho_k)\mu_k + \rho_k \log \theta_{t-1} \\ \sigma^2_{\theta|s}(e_k) &= \frac{1}{\frac{1}{\sigma_k^2} + \frac{1}{\sigma^2_{\epsilon,k}}} \quad \kappa = \frac{1}{1 + \frac{\sigma^2_{\varepsilon,k}}{\sigma^2_k}} \end{aligned} \tag{149}$$

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:

$$\log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right] = -\frac{1+\psi}{\alpha}\left(\log \gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \left(\sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \tag{150}$$

$$\begin{aligned} \log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right] &= \left(\frac{1}{\epsilon} - \gamma\right)\left(a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1})\right) \\ &+ \frac{1}{2}\left(\frac{1}{\epsilon} - \gamma\right)^2 a_1^2 \sigma^2_{\theta|s}(\lambda_{it}) \end{aligned} \tag{151}$$

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

$$\begin{aligned} \log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \Bigg[& \log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left[\log \gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right] \\ & - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \left(\sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \\ & + \left(\frac{1}{\epsilon} - \gamma\right)\left[a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + a_2 \log \theta_{t-1} + f(Q_t, \theta_{t-1})\right] \\ & + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma\right)^2 \sigma^2_{\theta|s}(\lambda_{it}) \Bigg] \end{aligned} \tag{152}$$

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 54 and:

$$\begin{aligned}\delta_t(e_k) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\ &+ \frac{1+\psi}{\alpha}\left[\log\gamma_i + \kappa\log\theta_t + (1-\kappa)\mu(e_k,\theta_{t-1})\right] \\ &- \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}(e_k)+\sigma^2_{\tilde{\theta}}\right) \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1\left(\kappa\log\theta_t + (1-\kappa)\mu(e_k,\theta_{t-1})\right) + a_2\log\theta_{t-1} + f(Q_t,\theta_{t-1})\right] \\ &+ \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}(e_k)\Bigg]\end{aligned} \tag{153}$$

for all $k \leq K$. Aggregating these best replies, using Equation 55, we obtain that:

$$\begin{aligned}\log Y_t &= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right] \\ &= \frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(e_k) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}\right) \\ &= \delta_t(e_1) + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 + \frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\left(\delta_t(e_k)-\delta_t(e_1)\right)\right\}\right)\end{aligned} \tag{154}$$

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:

$$f(Q,\theta_{t-1}) = \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}f(Q,\theta_{t-1}) + \frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1})\right\}\right) \tag{155}$$

and so:

$$f(Q,\theta_{t-1}) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1})\right\}\right) \tag{156}$$

Completing the proof. □

In the multidimensional models case with persistence, the past value of fundamentals interacts non-linearly with the cross-sectional model distribution in affecting aggregate output. However, without more structure, the properties of the dynamics generated by this multi-dimensional system are essentially unrestricted.

B.7 Persistent Idiosyncratic Shocks and Belief Updating

We now extend the analysis from Section B.6 to the case where agents' idiosyncratic states drive model 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:

$$\log \tilde{\theta}_{it} = \rho_{\tilde{\theta}} \log \tilde{\theta}_{i,t-1} + \zeta_{it} \tag{157}$$

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 models Q_t is now insufficient for describing aggregate output. This is because models 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 models 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 models given $\tilde{\theta}$ as $\check{Q}_{t,k|\tilde{\theta}} = \frac{\check{Q}_{t,k}(\tilde{\theta})}{f_{\tilde{\theta}}(\tilde{\theta})}$.

Proposition 10 (Equilibrium Characterization with Multi-Dimensional Models, Aggregate and Idiosyncratic Persistence, and Idiosyncratic Model Updating)**.** *There exists a quasi-loglinear equilibrium:*

$$\log Y(\log \theta_t, \log \theta_{t-1}, \check{Q}_t) = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(\check{Q}_t, \theta_{t-1}) \tag{158}$$

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

Proof. This proof follows closely that of Proposition 9. Under Model loading λ_{it}, we have that the agent's posterior regarding $\log \theta_{it}$ is given by:

$$\log \theta_{it} | \tilde{\theta}_{it-1}, \lambda_{it}, s_{it} \sim N\left(\log \gamma_i + \rho_{\tilde{\theta}} \log \tilde{\theta}_{it-1} + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_\xi\right) \tag{159}$$

where $\mu(\lambda_{it}, \theta_{t-1})$, κ, and $\sigma^2_{\theta|s}(\lambda_{it})$ are as in Proposition 9. Then substitute $\log \gamma_i + \rho_{\tilde{\theta}} \tilde{\theta}_{it-1}$ for $\log \gamma_i$ and follow the Proof of Proposition 9 until the aggregation step (Equation 154).

We now instead have that:

$$
\begin{aligned}
\log Y_t &= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\tilde{\theta}_{it-1},\lambda_{it}\right]\right] \\
&= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_t(e_k,\tilde{\theta}_{it-1})+\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\right)^2\hat{\sigma}^2\right\}\right] \\
&= \frac{\epsilon}{\epsilon-1}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(e_k,\tilde{\theta})+\frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right) \qquad (160) \\
&= \delta_t(e_1,1)+\frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 \\
&\quad+\frac{\epsilon}{\epsilon-1}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\epsilon}\left(\delta_t(e_k,\tilde{\theta})-\delta_t(e_1,1)\right)\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right)
\end{aligned}
$$

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 $k\leq K$. 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 9, we then have that:

$$
f(\check{Q},\theta_{t-1})=\frac{\frac{\epsilon}{\epsilon-1}}{1-\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1},\tilde{\theta})\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right) \qquad (161)
$$

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:

$$
f(\check{Q})=\frac{\frac{\epsilon}{\epsilon-1}}{1-\omega}\log\left(\mathbb{E}_{\tilde{\theta}}\left[\check{Q}_{t|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\}+(1-\check{Q}_{t|\tilde{\theta}})\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\}\right]\right) \qquad (162)
$$

where:

$$
\begin{aligned}
\delta_{OP}(\tilde{\theta}) &= \alpha\delta_{OP}+\frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}}\log\tilde{\theta} \\
\delta_{PP}(\tilde{\theta}) &= \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}}\log\tilde{\theta}
\end{aligned} \qquad (163)
$$

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:

$$
\begin{aligned}
&\check{Q}_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\} + (1-\check{Q}_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\} \\
&= Q_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\left(\alpha\delta_{OP} + \xi\log\tilde{\theta}\right)\right\} + (1-Q_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\} \\
&= Q_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\}\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\}-1\right] + \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\}
\end{aligned}
\tag{164}
$$

Taking the expectation of the relevant terms, we obtain:

$$
\begin{aligned}
&\mathbb{E}_{\tilde{\theta}}\left[\check{Q}_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\} + (1-\check{Q}_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\}\right] \\
&= \left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\}-1\right]\exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}Q_t \\
&\quad + \mathrm{Cov}_t\left(Q_{t|\tilde{\theta}}, \tilde{\theta}^{\frac{\epsilon-1}{\epsilon}\xi}\right) + \exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}
\end{aligned}
\tag{165}
$$

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

$$
\begin{aligned}
f(\check{Q}_t) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\omega}\log\Bigg(&\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\}-1\right]\exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}Q_t \\
&+ \mathrm{Cov}_t\left(Q_{t|\tilde{\theta}}, \tilde{\theta}^{\frac{\epsilon-1}{\epsilon}\xi}\right) + \exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2\frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\}\Bigg)
\end{aligned}
\tag{166}
$$

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

$$
\log\hat{\theta}_{it} = \sum_{\tau=1995}^{2019}\beta_\tau\cdot(\mathrm{opt}_{i\tau}\cdot\mathbb{I}[\tau=t]) + \chi_{j(i),t} + \gamma_i + \varepsilon_{it}
\tag{167}
$$

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 A11. 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 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 model updating in response to idiosyncratic conditions is not, at least in our data, quantitatively significant.

B.8 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 3). 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 in the absence of fundamental and aggregate shocks. We do so in a model with "contrarian" agents whose updating contradicts recent data and/or consensus. Our analysis of endogenous models 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 $\{Q, T(Q), \ldots, T^{k-1}(Q)\}$ 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,[24] in equilibrium, the induced probabilities that optimists and pessimists respectively become optimists are quadratic and given by:[25]

$$\tilde{P}_O(Q) = a_O + b_O Q - cQ^2\,,\ \tilde{P}_P(Q) = a_P + b_P Q - cQ^2 \tag{168}$$

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 contagiousness and associativeness (through the subsumed equilibrium map), and c captures any non-linearity.

The following result describes the potential dynamics:

Proposition 11. *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

$$\begin{aligned} T(Q) &= Q(a_O + b_O Q - cQ^2) + (1-Q)(a_P + b_P Q - cQ^2) \\ &= a_P + (a_O - a_P + b_P)Q - (c + b_P - b_O)Q^2 \end{aligned} \tag{169}$$

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

$$\eta = 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)} \tag{170}$$

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

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

[25] This can be microfounded in a generalization our earlier LAC model 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:

$$T^k(Q) = \left(h^{-1} \circ T'^k \circ h\right)(Q) \tag{171}$$

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:

$$\hat{T}(\hat{h}(Q)) = (k + \hat{\alpha}^2) + 2\hat{\alpha}\hat{\beta}Q + \hat{\beta}^2 Q^2 \tag{172}$$

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

$$\hat{h}^{-1}(\hat{T}(\hat{h}(Q))) = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}} + 2\hat{\alpha}Q + \hat{\beta}Q^2 \tag{173}$$

   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:

$$\omega_0 = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}},\ \omega_1 = 2\hat{\alpha},\ \omega_2 = -\hat{\beta} \tag{174}$$

   We therefore have that:

$$k = \hat{\beta}\omega_0 + \hat{\alpha}(1 - \hat{\alpha}) = -\omega_2\omega_0 + \frac{\omega_1}{2}\left(1 - \frac{\omega_1}{2}\right) \tag{175}$$

   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:

$$\check{T}(\check{h}(Q)) = \eta\left(\check{\alpha}(1 - \check{\alpha}) + \check{\beta}(1 - 2\check{\alpha})Q - \check{\beta}^2 Q^2\right) \tag{176}$$

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

$$\check{h}^{-1}(\check{T}(\check{h}(Q))) = \frac{\eta\check{\alpha}(1 - \check{\alpha}) - \check{\alpha}}{\check{\beta}} + \eta(1 - 2\check{\alpha})Q - \eta\check{\beta}Q^2 \tag{177}$$

Matching coefficients, we find:

$$k = \frac{\eta\check{\alpha}(1-\check{\alpha}) - \check{\alpha}}{\check{\beta}},\ 0 = \eta(1-2\check{\alpha}),\ 1 = -\eta\check{\beta} \tag{178}$$

We therefore obtain that:

$$k = \eta(\check{\alpha} - \eta(1-\check{\alpha})) = \frac{\eta}{2}\left(1 - \frac{\eta}{2}\right) \tag{179}$$

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

$$T = \hat{h}^{-1} \circ \check{h}^{-1} \circ \check{T} \circ \check{h} \circ \hat{h} \tag{180}$$

with

$$\begin{aligned}\eta &= 1 + \sqrt{1 - 4\left(-\omega_2\omega_0 + \frac{\omega_1}{2}\left(1 - \frac{\omega_1}{2}\right)\right)} = 1 + \sqrt{(\omega_1 - 1)^2 + 4\omega_2\omega_0} \\ &= 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)}\end{aligned} \tag{181}$$

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:

$$\begin{aligned}&\max_{Q\in[0,1]} \tilde{P}_i(Q) = \max\left\{a_i, a_i + b_i - c, \left(a_i + \frac{b_i^2}{4c}\right)\mathbb{I}[0 \leq b_i \leq 2c]\right\} \leq 1 \\ &\min_{Q\in[0,1]} \tilde{P}_i(Q) = \min\{a_i, a_i + b_i - c\} \geq 0\end{aligned} \tag{182}$$

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

$$\begin{aligned} & \max_{(a_O,a_P,b_O,b_P)\in\mathbb{R}^4} (a_O - a_P + b_P - 1)^2 + 4a_P(b_P - b_O) \\ & \text{s.t.} \max\{a_O, a_O + b_O\} \leq 1, \max\{a_P, a_P + b_P\} \leq 1 \\ & \min\{a_O, a_O + b_O\} \geq 0, \min\{a_P, a_P + b_P\} \geq 0 \end{aligned} \tag{183}$$

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 11 shows how to translate an updating rule of the form of Equation 168 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

$$\begin{aligned} \text{opt}_{it} = \alpha_1 \, \text{opt}_{i,t-1} + \beta_1 \text{opt}_{i,t-1} \cdot \overline{\text{opt}}_{i,t-1} + \\ \beta_2(1 - \text{opt}_{i,t-1}) \cdot \overline{\text{opt}}_{i,t-1} + \tau \left(\overline{\text{opt}}_{i,t-1}\right)^2 + \gamma_i + \varepsilon_{it} \end{aligned} \tag{184}$$

where γ_i is a firm fixed effect. This model allows the effects of contagiousness to depend on agents' previous state. In the mapping to Equation 168, $\alpha = a_P$, $\alpha_1 = a_O - a_P$, $\beta_1 = b_O$, $\beta_2 = b_P$, 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 170:

$$\hat{\eta} = 1 + \sqrt{(\hat{\alpha}_1 + \hat{\beta}_2 - 1)^2 + 4\hat{\alpha}_1(\hat{\tau} + \hat{\beta}_2 - \hat{\beta}_1)} \tag{185}$$

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 11. 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 dynamics obtain.

Our estimates are presented in Table A17. 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 contagious 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 et al., 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).

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.[26] 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 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

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

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 dataset is Compustat Annual Fundamentals. Our main variables of interest are defined in Table A18. 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 Table A18. 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`).[27] As in Ottonello and Winberry (2020) and Flynn and Sastry (2024), 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 (2024).

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;
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

[27] A small difference from Flynn and Sastry (2024) 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.

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 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:

$$\text{Share}_{M,j'} = \frac{\sum_{i:j(i)=j'} \sum_t \text{MaterialExpenditure}_{it}}{\sum_{i:j(i)=j'} \sum_t \text{Sales}_{it}} \tag{186}$$

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

$$\begin{aligned} \alpha_{M,j'} &= \mu \cdot \text{Share}_{M,j'} \\ \alpha_{K,j'} &= 1 - \alpha_{M,j'} \end{aligned} \tag{187}$$

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

$$\begin{aligned} \alpha_{M,j'} &= 1 \\ \alpha_{K,j'} &= 0 \end{aligned} \tag{188}$$

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

$$\log \tilde{\theta}_{it} = \log \text{Sales}_{it} - \frac{1}{\mu} \left(\alpha_{M,j(i)} \cdot \log \text{MatExp}_{it} + \alpha_{K,j(i)} \cdot \log \text{CapStock}_{it} \right) \tag{189}$$

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

$$\log \hat{\theta}_{it} = \log \tilde{\theta}_{it} - \chi_{j(i),t} \tag{190}$$

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

$$Y_i = \theta_i M_i^{\alpha_j} K_i^{1-\alpha_j} \tag{191}$$

They act as a monopolist facing the demand curve

$$p_i = Y_i^{-\frac{1}{\epsilon}} \tag{192}$$

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

$$\max_{M_i, K_i} \left\{ (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} - q_j M_i - r_j K_i \right\} \tag{193}$$

We first justify our formulas for the input shares (Equation 187). 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

$$\begin{aligned} q_j &= M_i^{-1} \alpha_j \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \\ r_j &= K_i^{-1} (1-\alpha_j) \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \end{aligned} \tag{194}$$

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

$$\begin{aligned} \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{q_j M_i}{p_i Y_i} \\ 1 - \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{r_j K_i}{p_i Y_i} \end{aligned} \tag{195}$$

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

$$Y = \left(\left(\frac{\epsilon - 1}{\epsilon} \right) \theta_i \left(\frac{\alpha_j}{q_j} \right)^{\alpha} \left(\frac{1 - \alpha_j}{r_j} \right)^{1-\alpha_j} \right)^{\epsilon} \Rightarrow p = \left(\frac{\epsilon}{\epsilon - 1} \right) \theta_i^{-1} \left(\frac{q_j}{\alpha_j} \right)^{\alpha_j} \left(\frac{r_j}{1 - \alpha_j} \right)^{1-\alpha_j} \tag{196}$$

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:

$$\alpha = \mu \frac{q_j M_i}{p_i Y_i} \tag{197}$$

Finally, we now apply Equations 189 and 190 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

$$\log \tilde{\theta}_i = \left(1 - \frac{1}{\epsilon}\right)(\log \theta_i - \alpha \log q_j - (1 - \alpha) \log p_{K,j}) \tag{198}$$

We finally observe that the industry-level means are

$$\chi_j = \left(1 - \frac{1}{\epsilon}\right)\left(\log \bar{\theta}_j - \alpha \log q_j - (1 - \alpha) \log p_{K,j}\right) \tag{199}$$

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

$$\log \hat{\theta}_i = \left(1 - \frac{1}{\epsilon}\right)(\log \theta_i) \tag{200}$$

or our measurement captures physical TFP, up to scale.

E Additional Empirical Results

E.1 A Test for Coefficient Stability

Here, we study the bias that may arise from omitted variables in our estimation of the effect of optimism on hiring, or δ^{OP} in Section 5.1, Equation 29, and Table 1. In particular, we apply the method of Oster (2019) to bound bias in the estimate of δ^{OP} under external assumptions about selection on unobservable variables and to calculate an extent of unobservable selection that could be consistent with a point estimate $\delta^{OP} = 0$ that corresponds to our null hypothesis (*i.e.,* "optimism is irrelevant for hiring"). We find that our results are highly robust by this criterion.

Set-up and Review of Methods. To review, our estimating equation is

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \tag{201}$$

Hiring and optimism are constructed as described in Section 4, at the level of firms and fiscal years. We treat firm and industry-by-time fixed effects as baseline controls that are necessary for interpreting the regression.[28] As our main "discretionary" controls, we consider current and past TFP and lagged labor—that is, $X_{it} = \{\log \hat{\theta}_{it}, \log \hat{\theta}_{i,t-1}, \log L_{i,t-1}\}$. Under our baseline model, these controls help increase precision, as they are in principle observable variables that explain hiring (Corollary 5). Thus, in this Appendix, we will study the regression model in which the fixed effects are partialed out of both the outcome, main regression, and controls, as indicated below with the $\perp$ superscript:

$$\Delta \log L_{it}^{\perp} = \delta^{OP} \text{opt}_{it}^{\perp} + \tau' X_{it}^{\perp} + \varepsilon_{it}^{\perp} \tag{202}$$

The essence of the method proposed by Oster (2019), who builds on the approach of Altonji, Elder, and Taber (2005), is to extrapolate the change in the coefficient in interest upon the addition of control variables, taking into account the better fit (*i.e.,* additional R^2) from adding the new regressors. To exemplify the logic, consider a case in which we first estimated Equation 202 without controls, obtaining a coefficient estimate of $\hat{\delta}_{NC}^{OP}$ and an R^2 of $\hat{R}_{NC}^2$, and then estimated the same equation with controls, obtaining a coefficient estimate of $\hat{\delta}_C^{OP}$ and an R^2 of $\hat{R}_C^2$. Both estimates are restricted to a common sample, for comparability. If $\hat{R}_C^2 = 1$, then (up to estimation error) we might presume that $\hat{\delta}_C^{OP} - \hat{\delta}_{NC}^{OP}$ estimates the entirety of the theoretically possible omitted variables bias, as there is no

[28] The latter, in particular, controls for the effect of fundamentals on hiring in our macroeconomic model. We leverage this interpretation of the *biased* estimate of δ^{OP} from a regression lacking this fixed effect in Appendix F.4.

remaining unmodeled variation in hiring. If $\hat{R}^2_C < 1$ and $\hat{R}^2_C - \hat{R}^2_{NC}$ is small (*i.e.,* the controls did not greatly improve fit), then we might presume that the residual still contains unobserved variables that could contribute toward more bias—in other words, the observed omitted variables bias $\hat{\delta}^{OP}_C - \hat{\delta}^{OP}_{NC}$ is only a small fraction of what is possible.

To formalize this idea, Oster (2019) introduces two auxiliary parameters: λ (the *proportional degree of selection*, called δ in the original paper), which controls the relative effect of observed and unobserved controls on the outcome, and $\bar{R}^2$, which is the maximum achievable fit of the regression with all (possibly bias-inducing) controls, presumed in the example above to be 1. Conditional on $\bar{R}^2$, Oster (2019) proposes an intuitively reasonable (and, in special cases and under specific asymptotic arguments, consistent) estimator for the degree of selection required to induce a zero coefficient, $\hat{\lambda}^*$. Conditional on both $\bar{R}^2$ and λ, Oster (2019) also proposes a bias-corrected coefficient estimator, which is $\hat{\delta}^*_{OP}$ in our language.

The key parameter that the researcher has to specify for the first calculation is $\bar{R}^2$: the proportion of variance in the outcome variable (hiring, net of firm and sector-by-time fixed effects) that can be explained by factors that correlate with the variable of interest (optimism) and explain the outcome variable. As the main source of omitted variation that could influence optimism and hiring is news about fundamentals, we benchmark $\hat{\bar{R}}^2$ by estimating a regression in which we include our base control set $X_{it} = \{\log \hat{\theta}_{it}, \log \hat{\theta}_{i,t-1}, \log L_{i,t-1}\}$ and control for two years of future fundamentals and labor choice, or

$$Z_{it} = \{\log \hat{\theta}_{i,t+1}, \log \hat{\theta}_{i,t+2}, \log L_{i,t+1}, \log L_{i,t+2}\}$$

This yields $\hat{\bar{R}}^2 = 0.459$. Oster (2019) also suggests as a benchmark that $\bar{R}^2$ could be taken as three times the R^2 in the controlled regression. We also report robustness to $\bar{R}^2_{\text{II}} = 0.387$, three times the value of $R^2 = 0.129$ that we find in the controlled regression. Thus, our baseline value of $\hat{\bar{R}}^2 = 0.459$ is more demanding than that suggested by Oster (2019). We finally construct the bias-corrected coefficients assuming $\lambda = 1$, or equal selection on unobservables and observables, for both values of $\bar{R}^2$.

Results. We report the results of this exercise in Table A2. Under our baseline value of $\hat{\bar{R}}^2 = 0.459$, we find that the degree of selection required to induced a zero coefficient is $\hat{\lambda}^* = 1.69$. This is well above the value of $\hat{\lambda}^* = 1$ that Oster (2019) suggests is likely to be conservative. Under the "three times R^2" benchmark, we obtain that $\hat{\lambda}^* = 2.15$. In both cases, we are robust to there being more selection on unobservables than on observables. According to Oster (2019), approximately 50% of the published top-journal articles in their sample are not robust to this extent of selection.

E.2 Alternative Empirical Strategy: CEO Change Event Studies

To further isolate variation in the beliefs held by firms that is unrelated to fundamentals, we study the effects on hiring of changes in beliefs induced by plausibly exogenous managerial turnover.

Data. To obtain plausibly exogenous variation in beliefs held at the firm level, we will examine the year-to-year change in firm-level beliefs stemming from plausibly exogenous CEO changes. To do this, we use the dataset of categorized CEO exits compiled by Gentry et al. (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

$$\Delta \log L_{it} = \delta^{CEO}\text{opt}_{it} + \psi\,\text{opt}_{i,t-1} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it} \tag{203}$$

This differs from our baseline Equation 29 by including parametric controls for lagged values of the model loadings, but removing a persistent firm fixed effect.[29] If the studied CEO changes are truly exogenous, as we have suggested, then the model loadings of the new CEO are, conditional on the model 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' beliefs.

We present our results in Table A19. We obtain estimates of δ^{CEO} that are quantitatively

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

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

$$\begin{aligned}\Delta \log L_{it} = \delta^{\text{NoChange}}\text{opt}_{it} + \delta^{\text{Change}}(\text{opt}_{it} \times \text{ChangeCEO}_{it}) + \alpha^{\text{Change}}\text{ChangeCEO}_{it} \\ + \psi\, \text{opt}_{i,t-1} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it}\end{aligned} \tag{204}$$

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.

Contagiousness 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

$$\overline{\text{opt}}^{\text{ceo}}_{j(i),t-1} = \frac{1}{|M_{j(i),t}|} \sum_{k \in M^c_{j(i),t}} \text{opt}_{k,t-1} \tag{205}$$

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 34. We present the corresponding estimates in Table A20. We find similar point estimates under IV and OLS, although the IV estimates are significantly noisier.

E.3 Measuring Contagiousness via Granular Instrumental Variables

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

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

$$\text{opt}_{it} = \tau' X_{it} + \chi_{j(i),t} + \gamma_i + u_{it} \tag{206}$$

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 5.1, 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:

$$\overline{\text{opt}}_t^{g,sw} = \sum_i \frac{\text{sales}_{it}}{\sum_i \text{sales}_{it}} \hat{u}_{it} \tag{207}$$

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 $\hat{u}_{it}$:

$$\overline{\text{opt}}_t^{g,sw} = \sum_{i':j(i)=j(i'),i'\neq i} \frac{\text{sales}_{i't}}{\sum_i \text{sales}_{i't}} \hat{u}_{i't} \tag{208}$$

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 33, but with one of the sales-weighted variables $Z_t \in \{\overline{\text{opt}}_t^{sw}, \overline{\text{opt}}_t^{g,sw}\}$:

$$\text{opt}_{it} = u \ \text{opt}_{i,t-1} + s \ Z_{t-1} + r \ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it} \tag{209}$$

The coefficient s measures contagiousness with respect to the sales-weighted measures of optimism. We estimate Equation 209 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

$$\text{opt}_{it} = u_{\text{ind}} \ \text{opt}_{i,t-1} + s_{\text{ind}} \ Z_{j(i),t-1} + r_{\text{ind}} \ \Delta \log Y_{j(i),t-1} + \gamma_i + \chi_t + \varepsilon_{it} \tag{210}$$

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 A21. First, studying aggregate contagiousness, 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 contagiousness via the sales-weighted measure (column 4). We find imprecise estimates, centered around 0, for contagiousness 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 Estimation

In this appendix, we provide complete details on the estimation of the model.

F.1 Normalizations

We begin by making two economically irrelevant normalizations to ease the interpretation of the results. First, we set $a_0 = 0$. As we are not concerned with the level of output in the model, this is a harmless normalization. Second, we normalize the updating rules so that an economy with no productivity shocks and no optimism shocks has an equal fraction of optimists and pessimists. As we have estimated optimism in the data as being above or below the time-series average level of optimism, this is also harmless normalization. More specifically, we update the LAC transition probabilities by introducing a parameter C_P:

$$\begin{aligned} P_O^H(\log Y, Q, \varepsilon) &= [\frac{u}{2} + r \log Y + sQ + C_P + \varepsilon]_0^1 \\ P_P^H(\log Y, Q, \varepsilon) &= [-\frac{u}{2} + r \log Y + sQ + C_P + \varepsilon]_0^1 \end{aligned} \tag{211}$$

And we set C_P such that an economy with neutral fundamentals ($\log \theta_t = \log \theta_{t-1} = 0$), equal optimists and pessimists ($Q = 1/2$), and no optimism shocks ($\varepsilon = 0$) continues to have equal optimists and pessimists. Specifically, this implies $C_P = \frac{1-s}{2}$.

F.2 Estimation Methodology

To calibrate the model, we proceed in four steps.

1. *Setting macro parameters.* We first set $(\epsilon, \gamma, \psi, \alpha)$. In Section 6.1 and Table 5, 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.4 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 $f(Q_t)$. We take our estimate of δ^{OP} from column 1 in Table 1. This regression identifies δ^{OP} for the reasons described in Corollary 5.
3. *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).[30] We apply our estimated function f to our measured time series of optimism to get an estimated optimism component of output. we then calculate

   $$\log \hat{Y}_t^f = \log \hat{Y}_t - \hat{f}(\hat{Q}_t) \tag{212}$$

   (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

   $$\log \hat{Y}_t^f - 0.086 \log \hat{Y}_t^f = .0078(\zeta_t + .32\, \nu_{t-1}) \tag{213}$$

   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 $a_1\sigma = 0.0078$ and $a_2\sigma = 0.32$. There is a unique such pair, reported in Table 5, which also is therefore the maximum likelihood estimate of (κ, σ).
4. *Calibrating the updating rule $(u, r, s, \sigma_\varepsilon^2)$.* The coefficients of the LAC updating model are estimated in column 1 of Table 3. Conditional on the previous calibration, we set σ_ε^2 so that within model Q_t has the same standard deviation as the aggregate optimism time series, which is 0.0533.

F.3 Estimation Details for the Multi-Dimensional Model

We introduce two strategies to measure granular topics. The first is a partially supervised method that detects firms' discussion of the nine *Perennial Economic Narratives* described by Shiller (2020). The second is an unsupervised Latent Dirichlet Allocation model (Blei et al., 2003), which flexibly identifies clusters of topics discussed by firms. We then describe

[30] 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.

how we combine this measurement with LASSO regressions to discipline the key parameters of the constellation model.

Narrative Identification of Topics. In his book *Narrative Economics*, Robert Shiller identifies a set of nine *Perennial Economic Narratives* that recur throughout American history. These are: 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. We quantify US firms' adoption of these narratives by measuring the similarity of the firms' language with the language Shiller uses to describe each narrative. This method "narratively identifies narratives" because it uses prior knowledge from Shiller's historical study to inform our approach.

Formally, we use a "tf-idf" method related to prior work by Hassan, Hollander, Van Lent, and Tahoun (2019) and Flynn and Sastry (2024). For each narrative k, we first compute the term-frequency-inverse-document-frequency (tf-idf) score to obtain a set of words most indicative of that narrative:

$$\text{tf-idf}(w)_k = \text{tf}(w)_k \times \log\left(\frac{1}{\text{df}(w)}\right) \tag{214}$$

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 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 $\mathcal{W}_k$. We print the twenty most common words in each $\mathcal{W}_k$ in Table A14.

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

$$\widehat{\text{Shiller}}_{it}^{k} = \sum_{w \in \mathcal{W}_k} \text{tf}(w)_{it} \tag{215}$$

We then compute a binary measure of narrative adoption by comparing to the in-sample median: $\text{Shiller}_{it}^k = \mathbb{I}[\widehat{\text{Shiller}}_{it}^{k} > \text{med}(\widehat{\text{Shiller}}_{it}^{k})]$. In Figure A9, we plot the raw time series for the aggregate variable corresponding to each chapter's narrative.

Unsupervised Recovery of Narratives via LDA. While "narrative identification" may help us focus on an *ex ante* reasonable set of topics, this method will invariably miss other topics—for example, those that pertain more heavily to our sample period than to the broader sweep of US economic history studied by Shiller. To identify topics without relying on external references, we apply Latent Dirichlet Allocation (LDA), a hierarchical Bayesian model in which documents are constructed by combining a low-dimensional, latent set of topics (Blei et al., 2003). The topics themselves are characterized 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.[31] We estimate a model with 100 topics. In Table A15, we print the top ten terms associated with each of our estimated topics. Given the estimated LDA, we construct the document-level topic score as the posterior probability of that topic in the estimated document-specific topic distribution $\hat{p}$:

$$\widehat{\text{topic}}_{it}^{k} = \hat{p}(k|d_{it}) \tag{216}$$

We then compute a binary measure of topic discussion by comparing to the in-sample median: $\text{topic}_{it}^{k} = \mathbb{I}[\widehat{\text{topic}}_{it}^{k} > \text{med}(\widehat{\text{topic}}_{it}^{k})]$.

F.4 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. By Corollary 5, we first recall that firms' hiring can be written in equilibrium as

$$\Delta \log L_{it} = \tilde{c}_{0,i} + \tilde{c}_{10} \log \theta_t + \tilde{c}_{11} \log \theta_{t-1} + \tilde{c}_2 f(Q_t) + \tilde{c}_3 \log \theta_{it} + \tilde{c}_4 \log L_{i,t-1} + \delta^{OP} \lambda_{it} + \zeta_{it} \tag{217}$$

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

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

$$\Delta \log L_{it} = \gamma_i + \chi_{j(i),t} + \delta^{OP} \text{opt}_{it} + z_{it} \tag{218}$$

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 model-implied hypothesis that $\mathbb{E}[z_{it}\text{opt}_{it}] = 0$,

[31] 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.

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:

$$\Delta \log L_{it} = \gamma_i + \delta^{OP}\text{opt}_{it} + \tilde{c}_{10} \log \theta_t + \tilde{c}_{11} \log \theta_{t-1} + \tilde{z}_{it} \tag{219}$$

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 219. 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:

$$\begin{aligned} \mathbb{E}[\hat{\delta}|X] &= \delta^{OP} + \left((X'X)^{-1}\mathbb{E}[X'Q|X] \frac{\alpha\omega}{1-\omega}\delta^{OP} \right)_1 \\ &= \delta^{OP} \left(1 + \frac{\alpha\omega}{1-\omega} \left((X'X)^{-1}\mathbb{E}[X'Q|X] \right)_1 \right) \end{aligned} \tag{220}$$

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

$$(X'X)^{-1}\mathbb{E}[X'Q|X] = \mathbb{E}\left[(X'X)^{-1}X'Q|X \right] \tag{221}$$

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

$$Q_t = \gamma_i + \beta_O^Q \text{opt}_{it} + \beta_\theta^Q \log \theta_t + \beta_{\theta_{-1}}^Q \log \theta_{t-1} + \varepsilon_t \tag{222}$$

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:

$$\mathbb{E}[\hat{\delta} \mid X] = \delta^{OP} \left(1 + \frac{\alpha\omega}{1-\omega} \right) \tag{223}$$

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 strong and are why we do not adopt this strategy for our main quantitative analysis. Nevertheless, we will find similar results, as we now describe.

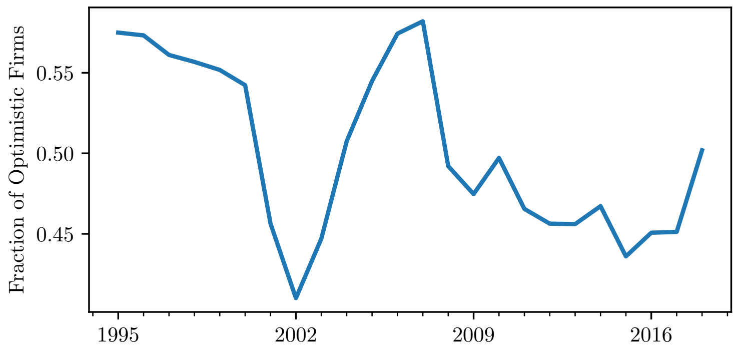
Empirical Application and Results. To operationalize this in practice, we compare estimates of Equation 218 and 219. For the latter, we proxy TFP using the cyclical component of both capacity adjusted and capacity un-adjusted TFP using the data of Fernald (2014).[32] We moreover maintain the assumption of $\alpha = 1$, or constant returns to scale, to map our estimates back to implied multipliers.

Our results are reported in Table A23, 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 A16, 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.

[32]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.

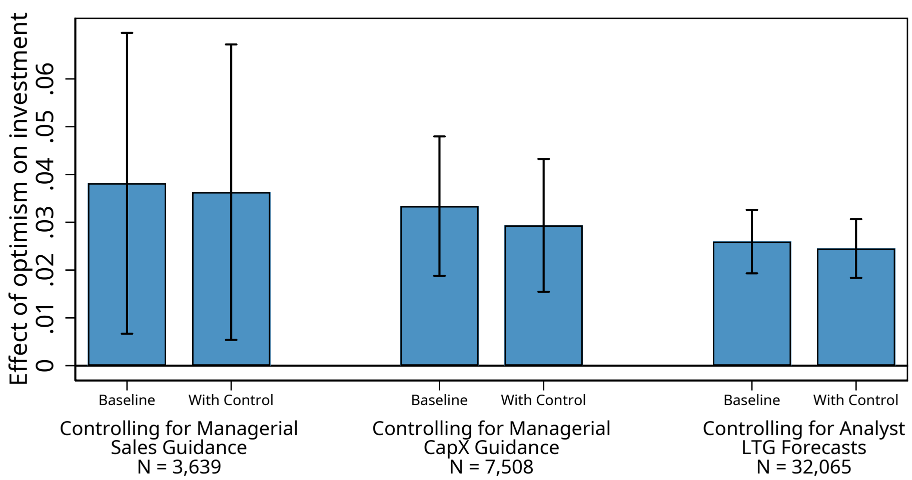
G Additional Figures and Tables

Figure A1: The Time Series of Optimism



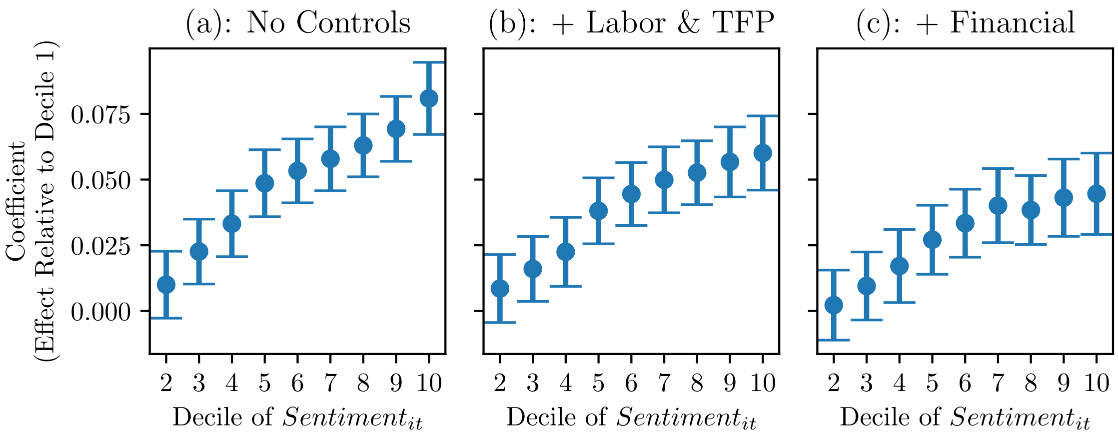
Notes: The plotted variable is the fraction of optimistic firms in each fiscal year. By construction, half of the firm-year observations in our sample are coded as optimistic. Section 4.2 describes our measurement strategy in full detail.

Figure A2: Language Matters for Capital Investment Conditional on Beliefs



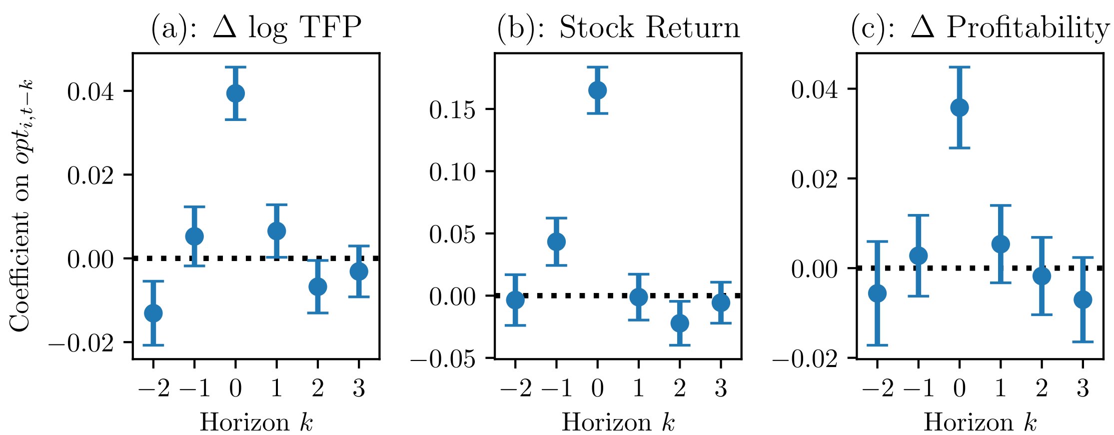
Notes: The regression model is Equation 29, the outcome is the change in firms' log capital stock from year $t-1$ to t, the main regressor is a binary indicator for optimism, and all specifications include firm and industry-by-time fixed effects. In each panel, we add a different control variable measuring beliefs: managerial guidance for sales growth (log of guidance value minus log of last year's sales), managerial guidance for capital expenditures growth (log of guidance value minus log of last year's capital expenditures), and analysts' long-term growth forecasts (both contemporaneous and first lag). The two bars show the coefficient on optimism on a common sample without and with the controls, respectively. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals based on standard errors clustered by firm ID.

Figure A3: Net Sentiment and Hiring



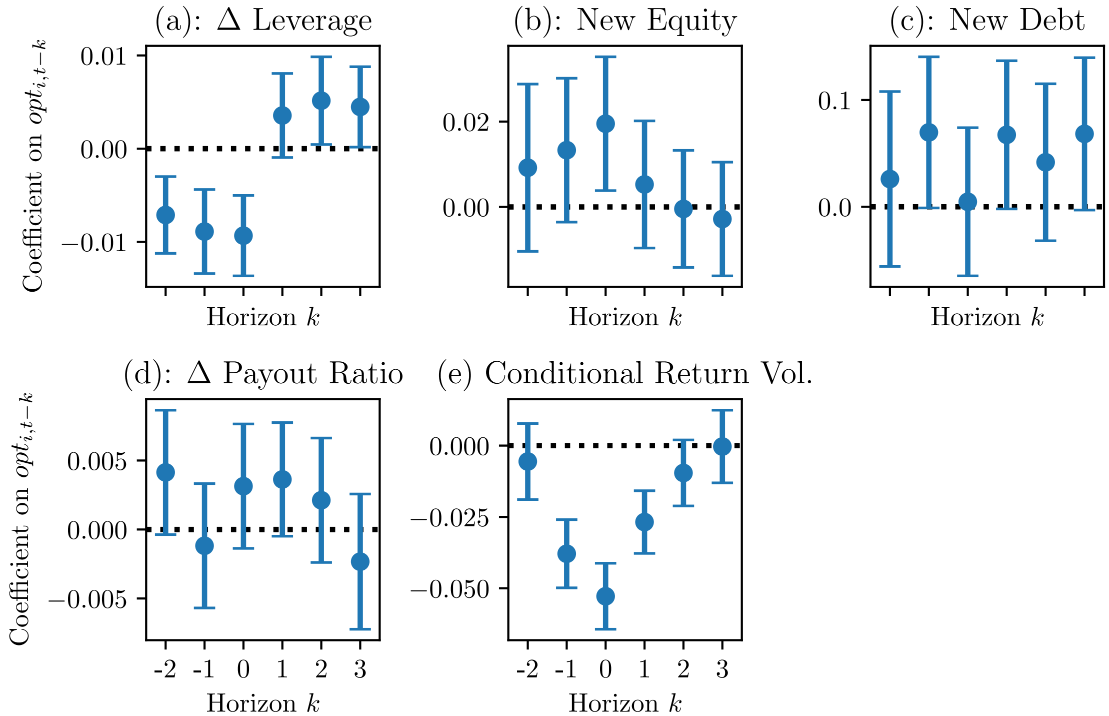
Notes: In each panel, we show estimates from the regression $\Delta \log L_{it} = \sum_{q=1}^{10} \beta_q \cdot (\text{sentiment}_{iqt}) + \tau' X_{it} + \gamma_i + \chi_{j(i),t} + \epsilon_{it}$, where sentiment$_{iqt}$ indicates decile q of the continuous sentiment variable. Panel (a) estimates this 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 A4: Dynamic Relationship between Optimism and Firm Fundamentals, Conference-Call Measurement



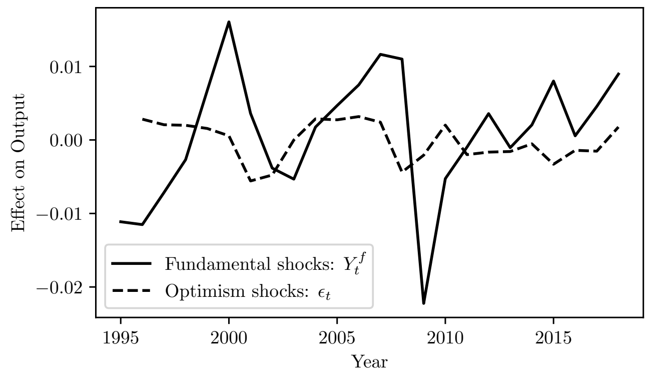
Notes: The regression model is Equation 31 (as in Figure 5), but measuring optimism from sales and earnings conference calls. Each coefficient is estimated from a separate projection regression. The outcomes are (a) the log change in TFP, calculated as described in Appendix D.2, (b) the log stock return inclusive of dividends over the fiscal year, 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 A5: Dynamic Relationship Between Optimism and Financial Variables



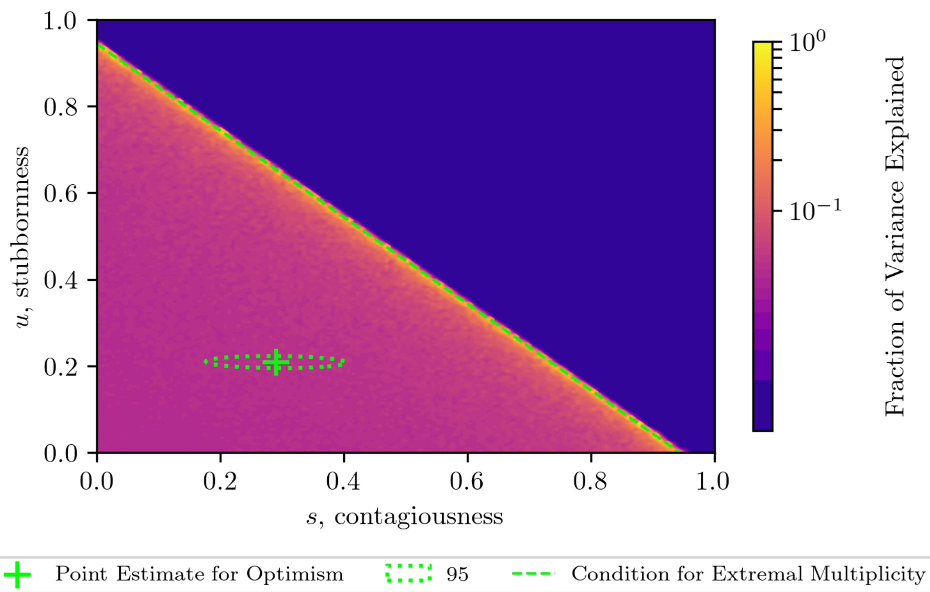
Notes: The regression model is Equation 31 (as in Figure 5), but with financial fundamentals as outcomes. Each dot shows the coefficient on binary optimism from a separate projection regression. The outcome variables are: (a) the fiscal-year-to-fiscal-year difference in leverage, which is total debt (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; (d) total dividends divided by earnings before interest and taxes (EBIT); and (e) squared stock returns (volatility). 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 A6: Fundamental and Optimism Shocks That Explain US GDP



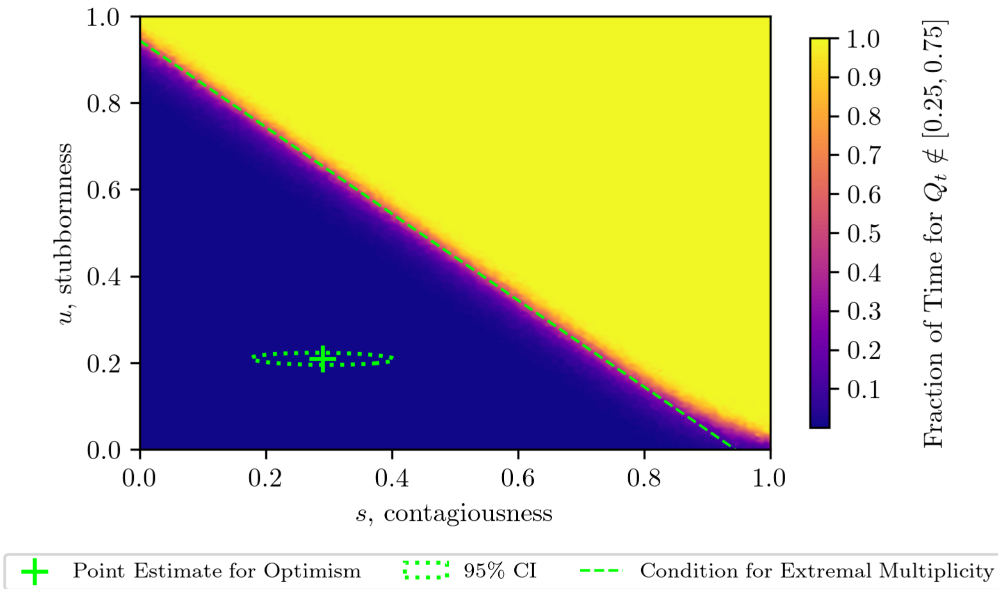
Notes: This figure shows the shocks that rationalize movements in optimism and detrended real GDP in recent US history, as analyzed in Section 6.2. The solid line is the exogenous process for fundamental output and the dashed line is the sequence of optimism shocks. The dashed line is rescaled by $\delta^{OP}(1-\omega)^{-1}$ to be, up to linear approximation of f, in units of output.

Figure A7: Variance Decomposition for Different Values of Stubbornness and Contagiousness, No Optimism Shocks



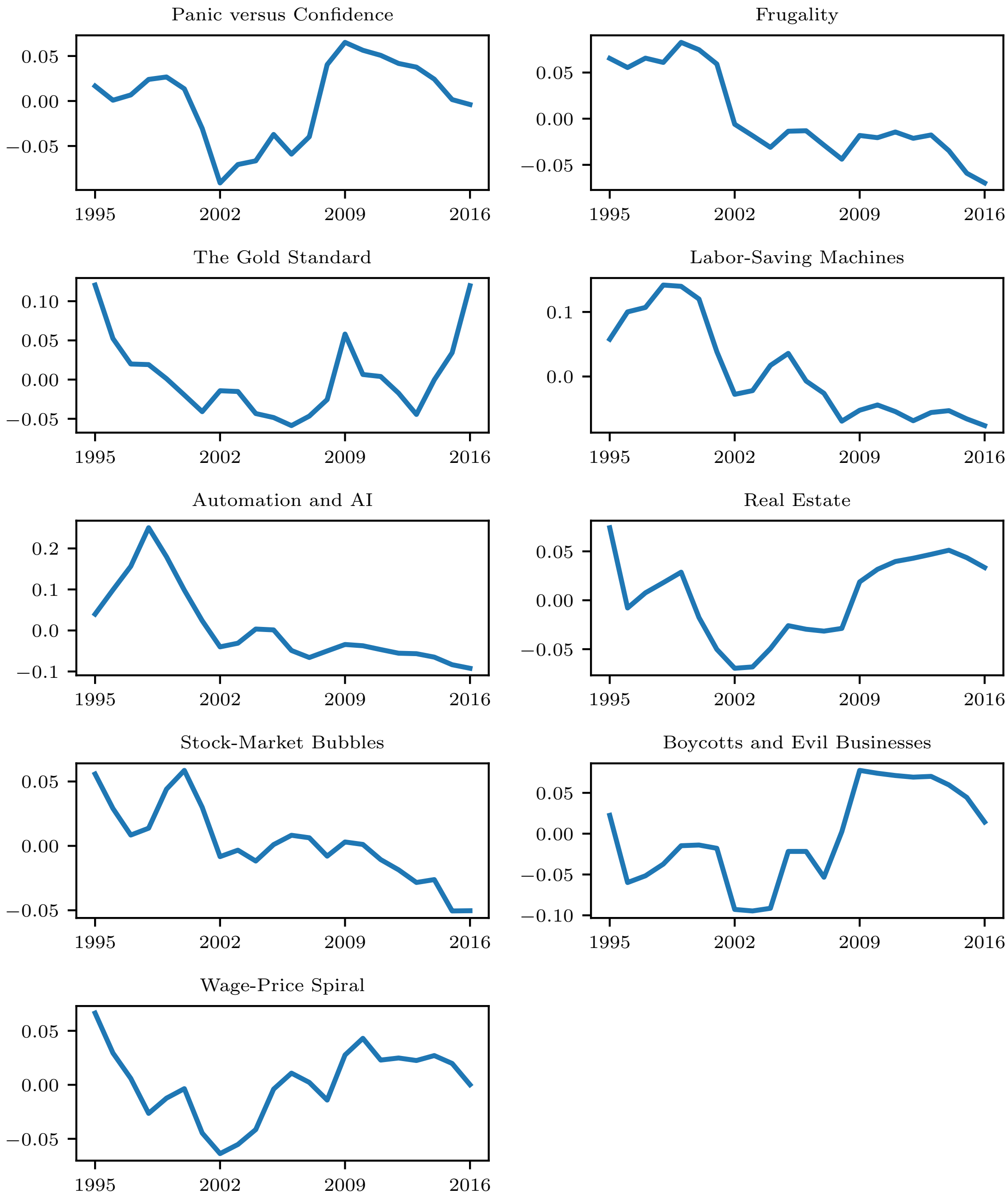
Notes: This Figure replicates Figure 8, with a different color bar scale, in the variant model with no exogenous shocks to optimism. Calculations vary u and s, holding fixed all other parameters at their calibrated values. The shading corresponds to the fraction of variance explained by optimism, or Share of Variance Explained$_0$ defined in Equation 38. The plus is our calibrated value of (u, s), corresponding to a variance share 4.7%, and the dotted line is the boundary of a 95% confidence set. The dashed line is the condition of extremal multiplicity from Corollary 4 and Equation 22.

Figure A8: Tendency Toward Extremal Optimism



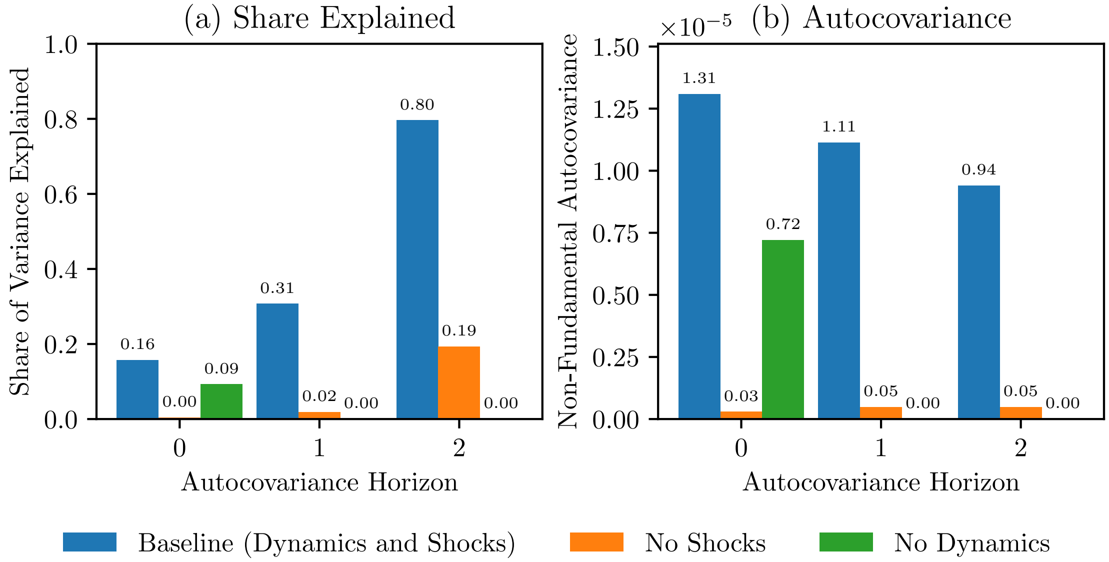
Notes: This Figure plots, in color, the fraction of time that optimism Q_t lies outside of the range $[0.25, 0.75]$ and therefore concentrates at extreme values. Calculations vary u and s, holding fixed all other parameters at their calibrated values. The plus is our calibrated value of (u, s), corresponding to an extremal share of 0%, and the dotted line is the boundary of a 95% confidence set. The dashed line is the condition of extremal multiplicity from Corollary 4 and Equation 22.

Figure A9: Time Series for Shiller's Perennial Economic Narratives



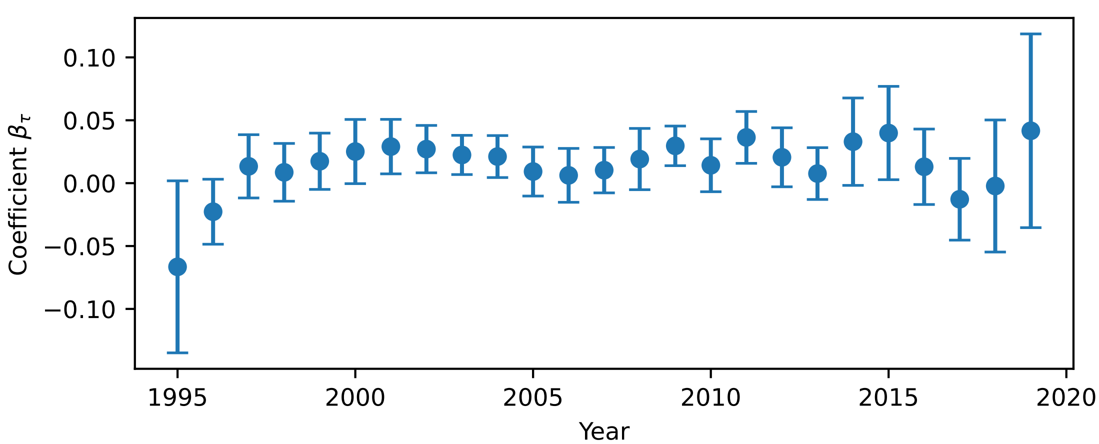
Notes: Each panel plots the time-series average of the narrative variable defined for the corresponding chapter of Shiller (2020)'s *Narrative Economics*. The units are cross-sectional averages of z-score transformed variables.

Figure A10: Optimism and Output Variance in the Constellations Model



Notes: This figure recreates Figure 7 in the model with constellations. The left panel plots the fraction of variance, one-year autocovariance, and two-year autocovariance explained by endogenous optimism in model simulations. The right panel plots the total non-fundamental autocovariance. In each figure, we plot results under three model scenarios: the baseline model with optimism shocks and optimism dynamics (blue), a variant model with no shocks, or $\sigma^2_{\varepsilon,k} = 0$ for all k (orange), and a variant model with shocks but no dynamics for model spread, or $u^k = r^k = s^k = 0$ for all k (green).

Figure A11: Time-Varying Relationship Between Optimism and TFP



Notes: Each dot is a coefficient β_τ estimated from Equation 167, 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 4.2.

Table A2: Robustness to Assumptions About Unobserved Selection When Estimating the Effect of Optimism on Hiring

Panel A: Regression Estimates

	(1)	(2)
	Outcome is $\Delta L_{it}^{\perp}$	
$\text{opt}_{it}^{\perp}$	0.0373	0.0305
Controls		✓
N	39,298	39,298
R^2	0.005	0.129

Panel B: Oster (2019) Statistics

	(1)	(2)
	$\bar{R}^2$ is	
	$\hat{\bar{R}}^2 = 0.459$	$\bar{R}^2_{\Pi} = 0.387$
λ^* ($\delta^{OP} = 0$)	1.691	2.151
δ^*_{OP} ($\lambda = 1$)	0.0126	0.0165

Notes: This table summarizes the coefficient stability test described in Appendix E.1. Panel A shows estimates of Equation 202, with and without controls for current and lagged log TFP and lagged log labor. The estimate in column 1 differs from that in column 1 of Table 1 due to restricting to a common sample in columns 1 and 2. The R^2 values are for the model after partialing out fixed effects, and hence correspond with unreported "within-R^2" values in Table 1. Panel B prints the two statistics of Oster (2019). In column 1, we set $\bar{R}^2$ equal to our estimated value of 0.459, calculated as described in the text from an "over-controlled" regression of current hiring on lagged controls and future hiring and productivity. In column 2, we use $\bar{R}^2$ given by three times the R^2 in the controlled hiring regression. The first row (λ^* ($\delta^{OP} = 0$)) reports the degree of proportional selection that would generate a null coefficient. The second row ($\delta^*_{OP}(\lambda = 1)$) is the bias corrected effect assuming that unobservable controls have the same proportional effect as observable controls.

Table A3: 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 29. 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 A4: 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 29 and the outcome is the log change in firms' employment from year $t-1$ to t. The main regressor is a binary indicator for optimism, defined in Section 4.2. In all specifications, we trim the 1% and 99% tails of the outcome variable. In column 5, the regression model is Equation 30, the outcome is the log change in firms' employment from year t to $t+1$, and control variables are dated $t+1$.

Table A5: 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 29, using $\text{opt}_{i,t-1}$ as an instrument for opt_{it}. Specifically, the structural equation is

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

the endogenous variable is opt_{it} and the excluded instrument is $\text{opt}_{i,t-1}$. 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 A6: 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 A7: The Effect of Optimism on All Inputs

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome is					
	$\Delta \log L_{it}$		$\Delta \log M_{it}$		$\Delta \log K_{it}$	
opt_{it}	0.0355	0.0305	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 A8: The Effect of Optimism on Stock Prices, High-Frequency Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome is stock return on					
	Filing Day		Prior Four Days		Next Four Days	
opt_{it}	0.000145	-0.000142	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 A9: Optimism and Managerial Optimism Relative to Analysts

	(1)	(2)
	Outcome is $\text{GuidanceOptExAnte}_{i,t+1}$	
opt_{it}	0.0351	0.00517
	(0.0197)	(0.0309)
Ind.-by-time FE	✓	✓
Lag labor		✓
Current and lag TFP		✓
N	3,821	2,190
R^2	0.143	0.178

Notes: The regression model is a variant of Equation 32 with a different outcome variable. The outcome, GuidanceOptExAnte, is a binary indicators for whether is an indicator of whether managers' sales guidance exceeds the analyst consensus. Standard errors are two-way clustered by firm ID and industry-year.

Table A10: Optimism is Contagious 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 3 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 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 A11: Optimism is Contagious and Associative, NYSE Peer Set Model

	(1)	(2)
	Outcome is opt_{it}	
Own lag, $\text{opt}_{i,t-1}$	0.214	0.135
	(0.0080)	(0.0166)
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	52,258	8,514
R^2	0.501	0.501

Notes: The regression model is Equation 34. Industry and peer average optimism are leave-one-out averages of the optimism variable over the respective sets of firms. We 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. 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. 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. 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 statistically significant (estimate 0.243, standard error 0.075).

Table A12: Sentiment is Contagious 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 33 for column 1, and a variant of Equation 34 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 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 A13: Sentiment is Contagious and Associative, Controlling for Past and Future Outcomes

	(1)	(2)	(3)	(4)	(5)
	Outcome is sentiment_{it}				
Aggregate lag, $\overline{\text{sentiment}}_{t-1}$	0.253	0.385	0.410		
	(0.0519)	(0.0651)	(0.1103)		
Ind. lag, $\overline{\text{sentiment}}_{j(i),t-1}$				0.175	0.151
				(0.0360)	(0.0409)
Time FE				✓	✓
Firm FE	✓	✓	✓	✓	✓
Own lag, $\text{opt}_{i,t-1}$	✓	✓	✓	✓	✓
$(\Delta \log Y_{t+k})_{k=-2}^{2}$		✓	✓		
$(\Delta \log Y_{j(i),t+k})_{k=-2}^{2}$			✓		✓
N	63,881	48,889	37,643	51,555	37,643
R^2	0.568	0.578	0.599	0.599	0.601

Notes: The regression model is a variant of Equation 35 for column 1-3, and an analogous variant of industry-level specification for columns 4 and 5 (*i.e.*, Equation 34 with past and future controls), with the continuous variable sentiment_{it} (and averages thereof) substituted for binary optimism. Columns 1 and 4 correspond, respectively, with columns 1 and 3 of Table A12. The added control variables are two leads, two lags, and the contemporaneous value of: real GDP growth (columns 2-3), and industry-level output growth (columns 3 and 5). 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 A14: 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 4.2.

Table A15: The Ten Most Common Words for Each Selected Topic

Topic 1		Topic 2		Topic 3		Topic 4		Topic 5		Topic 6		Topic 7		Topic 8	
lease	0.047	solid	0.791	foreign	0.097	borrower	0.034	plan	0.066	advertising	0.029	insurance	0.082	derivative	0.078
tenant	0.042	scheme	0.02	currency	0.067	agent	0.029	participant	0.031	retail	0.028	loss	0.031	value	0.05
landlord	0.03	line	0.009	income	0.045	lender	0.022	employee	0.02	brand	0.018	income	0.018	fair	0.048
lessee	0.017	asset	0.008	tax	0.038	agreement	0.02	committee	0.015	credit	0.018	investment	0.017	rate	0.039
rent	0.016	income	0.008	exchange	0.035	loan	0.02	employer	0.014	consumer	0.017	fix	0.016	interest	0.038
lessor	0.014	debt	0.007	comprehensive	0.023	credit	0.018	make	0.013	distribution	0.016	policy	0.015	asset	0.038
property	0.012	tax	0.007	translation	0.023	bank	0.013	account	0.013	card	0.015	business	0.015	hedge	0.025
term	0.011	cash	0.006	loss	0.021	administrative	0.012	provide	0.011	marketing	0.015	life	0.014	gain	0.022
day	0.009	credit	0.006	gain	0.018	interest	0.012	payment	0.01	food	0.013	premium	0.013	credit	0.019
provide	0.008	loss	0.005	financial	0.017	make	0.011	amount	0.01	store	0.013	write	0.012	financial	0.019
Topic 9		**Topic 10**		**Topic 11**		**Topic 12**		**Topic 13**		**Topic 14**		**Topic 15**		**Topic 16**	
benefit	0.089	stock	0.036	international	0.068	fund	0.059	financial	0.041	corporation	0.119	million	0.036	trustee	0.02
plan	0.08	common	0.033	united	0.065	investment	0.046	income	0.039	board	0.032	debt	0.031	seller	0.016
asset	0.06	financial	0.033	group	0.052	asset	0.032	cash	0.024	meeting	0.02	due	0.023	respect	0.014
pension	0.04	cash	0.022	global	0.031	trading	0.03	consolidated	0.02	stock	0.02	earning	0.022	indenture	0.013
define	0.033	asset	0.019	canada	0.022	value	0.026	approximately	0.018	director	0.016	percent	0.022	holder	0.011
cost	0.031	accounting	0.014	limited	0.022	management	0.022	asset	0.015	president	0.015	segment	0.018	notice	0.011
value	0.023	business	0.013	reference	0.021	market	0.02	statement	0.012	financial	0.013	interest	0.018	provide	0.011
tax	0.022	item	0.012	incorporate	0.017	capital	0.019	share	0.012	officer	0.012	include	0.017	interest	0.011
obligation	0.018	equity	0.011	us	0.013	income	0.017	accounting	0.012	business	0.011	relate	0.015	person	0.01
income	0.018	loss	0.011	sa	0.013	fee	0.015	tax	0.012	vote	0.01	information	0.015	purchaser	0.01
Topic 17		**Topic 18**		**Topic 19**		**Topic 20**		**Topic 21**		**Topic 22**		**Topic 23**		**Topic 24**	
agreement	0.071	type	0.058	stock	0.152	stock	0.049	gaming	0.035	double	0.405	exhibit	0.042	member	0.499
party	0.018	accounting	0.042	common	0.086	compensation	0.039	service	0.029	solid	0.214	incorporate	0.03	scheme	0.125
provide	0.014	lease	0.039	price	0.037	tax	0.039	network	0.022	income	0.022	reference	0.03	line	0.036
termination	0.011	topic	0.038	exercise	0.036	share	0.028	wireless	0.021	scheme	0.018	item	0.026	amount	0.027
write	0.01	asset	0.037	option	0.036	income	0.023	local	0.019	cash	0.016	registrant	0.023	abstract	0.026
employee	0.009	codification	0.034	purchase	0.034	average	0.019	cable	0.015	loss	0.014	exchange	0.023	asset	0.017
set	0.009	publisher	0.034	agreement	0.03	expense	0.018	provide	0.014	tax	0.014	pursuant	0.019	balance	0.015
notice	0.008	equipment	0.031	share	0.027	asset	0.016	equipment	0.013	balance	0.009	annual	0.018	datum	0.014
information	0.008	balance	0.026	value	0.019	outstanding	0.016	access	0.013	asset	0.007	bank	0.017	type	0.014
day	0.008	definition	0.022	warrant	0.017	weight	0.015	video	0.012	receivable	0.007	financial	0.017	value	0.013
Topic 25		**Topic 26**		**Topic 27**		**Topic 28**		**Topic 29**		**Topic 30**					
medical	0.176	june	0.136	executive	0.072	reorganization	0.048	court	0.038	technology	0.018				
health	0.142	march	0.123	compensation	0.03	bankruptcy	0.047	settlement	0.027	revenue	0.017				
care	0.123	note	0.089	employment	0.025	plan	0.044	district	0.021	development	0.015				
provide	0.028	agreement	0.057	officer	0.024	predecessor	0.036	certain	0.019	business	0.013				
management	0.027	august	0.05	board	0.024	successor	0.027	litigation	0.016	customer	0.012				
system	0.027	financial	0.026	committee	0.02	chapter	0.021	action	0.016	stock	0.012				
federal	0.024	interest	0.024	director	0.019	asset	0.019	complaint	0.012	product	0.012				
program	0.023	item	0.016	chief	0.017	court	0.018	damage	0.011	support	0.009				
insurance	0.022	payable	0.015	president	0.017	cash	0.016	approximately	0.011	market	0.009				
service	0.02	due	0.014	annual	0.015	certain	0.014	case	0.01	service	0.008				

Notes: The ten most common words (lemmatized bigrams) in example topics estimated by LDA and selected by our LASSO procedure as relevant for hiring (see Section 5.1). Weights correspond to relative importance for scoring the document. The LDA model and our estimation procedure are described in Section 4.2.

Table A16: 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	0.192	0.335	0.316	0.181
High ψ	1.0	0.0	2.5	2.6	0.133	1.154	0.175	0.359	0.186	0.106
High γ	1.0	1.0	0.4	2.6	-0.784	0.560	0.041	0.184	0.090	0.052
Empirical Multiplier	1.0	0.0	1.15	2.6	0.250	1.333	0.167	0.329	0.215	0.123
Calibrated Multiplier	1.0	0.0	0.845	2.6	0.313	1.455	0.168	0.324	0.235	0.134
High ϵ	1.0	0.0	0.21	5.0	0.490	1.962	0.109	0.240	0.317	0.181
Decreasing RtS	0.75	0.0	0.05	2.6	0.490	1.962	0.125	0.238	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, $\hat{c}_Q(0)$, and at a one-year horizon, $\hat{c}_Q(1)$, by optimism. The second two are the fraction of output losses in the 2000-02 downturn and 2007-09 downturn explained by fluctuations in 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 Becko et al. (2024). Calibrated multiplier adjusts ψ to match our own calculation of the multiplier in our setting in Appendix F.4. 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 A17: 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 184. η is a function of the regression coefficients defined in Equation 185, and interpretable in the model of cycles and chaos in Appendix B.8. Standard errors are two-way clustered by firm ID and industry-year. The standard error for η is calculated using the delta method.

Table A18: 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 A19: 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 203 for columns 1-3, and Equation 204 for column 4. The outcome is the log change in firms' employment. opt_{it} is a binary indicator for the optimism, defined in Section 4.2. 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 A20: The Contagiousness 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 A21: Optimism is Contagious 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 33 and 34, respectively modeling the spread of optimism at the aggregate and industry level, using granular identification of spillovers (contagiousness). $\overline{\text{opt}}^{sw}_{t-1}$ 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.3. 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 A22: Calibration Parameters for the Constellation Model

Name	ζ	u	r	s	M	Variance
Lease, Tenant, Landlord...	-0.135	0.063	-0.342	0.820	-0.113	0.003
Solid, Scheme, Line...	0.055	0.434	-2.304	0.678	0.103	0.049
Foreign, Currency, Income...	0.021	0.373	-0.180	0.514	-0.114	0.005
Borrower, Agent, Lender...	-0.066	0.064	0.019	0.747	-0.189	0.010
Plan, Participant, Employee...	-0.023	0.020	0.450	0.871	-0.109	0.012
Advertising, Retail, Brand...	0.056	0.324	0.078	0.594	-0.082	0.005
Insurance, Loss, Income...	-0.039	0.250	-0.132	0.651	-0.099	0.002
Derivative, Value, Fair...	0.042	0.410	0.099	0.407	-0.183	0.011
Benefit, Plan, Asset...	0.097	0.335	-0.500	0.568	-0.100	0.001
Stock, Common, Financial...	-0.041	0.230	-0.212	0.285	-0.484	0.002
International, United, Group...	0.032	0.321	1.369	0.729	0.053	0.020
Fund, Investment, Asset...	0.054	0.219	0.365	0.837	0.057	0.001
Financial, Income, Cash...	0.085	0.084	1.018	0.921	0.011	0.121
Corporation, Board, Meeting...	0.050	0.201	0.812	0.783	-0.012	0.117
Million, Debt, Due...	0.026	0.307	0.138	0.405	-0.288	0.002
Trustee, Seller, Respect...	-0.079	-0.006	-0.165	1.002	-0.003	0.015
Agreement, Party, Provide...	-0.117	0.039	-0.067	0.864	-0.097	0.021
Type, Accounting, Lease...	-0.049	0.371	0.592	0.600	-0.031	0.150
Stock, Common, Price...	0.043	0.198	1.020	0.945	0.146	0.031
Stock, Compensation, Tax...	0.023	0.274	-0.671	0.686	-0.041	0.071
Gaming, Service, Network...	0.042	0.375	0.137	0.444	-0.181	0.004
Double, Solid, Income...	0.032	0.450	-2.022	0.684	0.129	0.046
Exhibit, Incorporate, Reference...	0.033	0.187	0.139	0.802	-0.011	0.094
Member, Scheme, Line...	0.035	0.470	-0.655	0.537	0.006	0.011
Medical, Health, Care...	0.056	0.361	0.026	0.522	-0.116	0.001
June, March, Note...	0.040	0.242	0.312	0.663	-0.094	0.040
Executive, Compensation, Employee...	0.024	0.163	0.880	0.894	0.058	0.041
Reorganization, Bankruptcy, Plan...	-0.085	0.357	-0.119	0.206	-0.436	0.000
Court, Settlement, District...	-0.104	0.363	0.363	0.560	-0.079	0.010
Technology, Revenue, Development...	0.091	0.299	0.674	0.559	-0.138	0.013
Panic versus Confidence	0.017	0.223	-0.143	0.428	-0.349	0.003
The Gold Standard	0.017	0.204	0.724	0.955	0.159	0.005
Labor-Saving Machines	0.024	0.212	0.239	0.278	-0.511	0.001
Automation and AI	0.029	0.214	0.196	0.148	-0.638	0.001
Real Estate	0.022	0.206	-0.130	0.552	-0.243	0.002
Stock-Market Bubbles	0.012	0.221	0.126	0.472	-0.306	0.000
Boycotts and Evil Businesses	0.042	0.168	0.043	0.640	-0.192	0.003
Wage-Price Sprials	0.021	0.221	-0.113	0.669	-0.110	0.002

Notes: This table reports the topics used in the calibration of Section 6.4. The first set of rows are LDA topics, identified by their three highest-scoring terms, and the second set of rows are chapters of Shiller (2020), identified by shortened forms of their titles. The topics are selected via post-LASSO estimation of Equation 42, and the first column reports the coefficients. The remaining columns report estimates of stubbornness, associativeness, and contagiousness; the composite statistic M; and the unconditional time-series variance of each topic. In the estimation, we re-normalize each topic to have a positive ζ.

Table A23: 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.4. 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 223. Standard errors are double-clustered by industry-year and firm ID.

References

Joseph G Altonji, Todd E Elder, and Christopher R Taber. Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184, 2005.

Marco Battaglini. Chaos and unpredictability in dynamic social problems. Technical Report 28347, National Bureau of Economic Research, 2021.

Marianne Baxter and Robert G King. Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4):575–593, 1999.

Paul Beaudry, Dana Galizia, and Franck Portier. Putting the cycle back into business cycle analysis. *American Economic Review*, 110(1):1–47, 2020.

John S Becko, Joel P Flynn, and Christina Patterson. Fiscal policy in a networked economy. Working Paper 29619, National Bureau of Economic Research, 2024.

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, pages 993–1022, 2003.

Michele Boldrin and Michael Woodford. Equilibrium models displaying endogenous fluctuations and chaos: a survey. *Journal of Monetary Economics*, 25(2):189–222, 1990.

Liuchun Deng, M Ali Khan, and Tapan Mitra. Continuous unimodal maps in economic dynamics: On easily verifiable conditions for topological chaos. *Journal of Economic Theory*, 201:105446, 2022.

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

Joel P Flynn and Karthik A Sastry. Attention cycles. Working Paper 32553, National Bureau of Economic Research, 2024.

Xavier Gabaix and Ralph SJ Koijen. Granular instrumental variables. Working Paper 28204, National Bureau of Economic Research, 2020.

Richard J Gentry, Joseph S Harrison, Timothy J Quigley, and Steven Boivie. A database of ceo turnover and dismissal in s&p 1500 firms, 2000–2018. *Strategic Management Journal*, 42(5):968–991, 2021.

Tarek A Hassan, Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun. Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics*, 134(4): 2135–2202, 2019.

Matthew Hoffman, Francis Bach, and David Blei. Online learning for latent dirichlet allocation. *Advances in Neural Information Processing Systems*, 23, 2010.

Markku Kaustia and Ville Rantala. Common analysts: method for defining peer firms. *Journal of Financial and Quantitative Analysis*, 56(5):1505–1536, 2021.

Tien-Yien Li and James A Yorke. Period three implies chaos. *The American Mathematical Monthly*, 82(10):985–992, 1975.

Tim Loughran and Bill McDonald. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65, 2011.

Bill McDonald. Software repository for accounting and finance, 2021. Accessed June 2020. Available at: https://sraf.nd.edu/.

Emily Oster. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204, 2019.

Pablo Ottonello and Thomas Winberry. Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, 88(6):2473–2502, 2020.

SEC. Investor bulletin: How to read a 10-k. Article, Securities and Exchange Commission Office of Investor Education and Advocacy, 2011. Available at: https://www.sec.gov/files/reada10k.pdf.

Robert J Shiller. *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press, 2020.

Eric W Weisstein. Logistic map. *Wolfram MathWorld*, 2001.

Iván Werning. Comment on: Is the macroeconomy locally unstable and why should we care? In *NBER Macroeconomics Annual 2016*, volume 31, pages 540–552. University of Chicago Press, 2017. doi: 10.1086/690257.