2021

Three essays in macroeconomics

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

THREE ESSAYS IN MACROECONOMICS

by

JULIO L. ORTIZ

B.S., George Washington University, 2013
M.A.P.E., Boston University, 2019

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2021

Approved by

First Reader

Stephen Terry, Ph.D.
Assistant Professor of Economics

Second Reader

Adam Guren, Ph.D.
Assistant Professor of Economics

Third Reader

Pascual Restrepo, Ph.D.
Assistant Professor of Economics

Acknowledgments

I am extraordinarily grateful to my friends and family whose support propelled me through this program. I am especially grateful to my parents Julio E. Ortiz and Luisa Ortiz for all of their sacrifices, and for instilling in me a deep appreciation of higher education. My parents, along with my sister Evelyn Choy, have always shown me unconditional love. I would also like to thank my wonderful partner, Carter Rosekrans. His love and companionship kept me grounded and allowed me to see more clearly when the stress of the program blurred my view.

I would importantly like to thank my advisors Adam Guren, Pascual Restrepo, and Stephen Terry for their advice and guidance throughout the years. They have taught me a great deal not just about economics, but about how to be an economist. I am especially indebted to my co-main advisors Adam Guren and Stephen Terry for their unwavering support and willingness to answer even the most trivial question. Reminding myself that they were rooting for me was sometimes all I needed to get through a rough patch in this process. I would also like to thank the rest of the Department of Economics at Boston University. Through the coursework, seminars, and office hours, every interaction has been incredibly valuable to me.

Julio L. Ortiz

Doctoral Candidate

Department of Economics

THREE ESSAYS IN MACROECONOMICS

JULIO L. ORTIZ

Boston University, Graduate School of Arts and Sciences, 2021

Major Professor: Stephen Terry, Ph.D.
Assistant Professor of Economics

ABSTRACT

This dissertation consists of three essays studying firm dynamics and expectation formation. The first essay quantifies a tradeoff associated with lean inventory management. The second essay makes sense of simultaneous over- and underreaction in a noisy information setting with time-varying volatility. The third essay offers a new testable implication as a way to narrow the set of models of belief formation that are consistent with survey data.

The first essay investigates just-in-time production (JIT). I first construct a new measure of JIT at the firm level through a text search. Relative to non-adopters, I document that adopters experience higher sales and smoother outcomes, however, they are also more cyclical and sensitive to weather events. Motivated by these facts, I build and structurally estimate a dynamic general equilibrium model of JIT production. Relative to a counterfactual reflecting the adoption patterns of the 1980s, firms in the estimated economy benefit from a 1% increase in firm value in normal times. Amid a COVID-like disaster, however, the estimated economy experiences a 1.6 percentage point sharper contraction.

The second essay examines the role that volatility can play in generating seemingly non-rational behavior. First, I document that the same professional forecaster over- and underreacts to distinct macroeconomic variables. I then show that such behavior

can arise in a noisy information environment with unobserved volatility and costly model adoption. In such a model, forecasters overreact to variables for which they have less precise information and underreact to variables for which they have more precise information. I provide empirical evidence in favor of this explanation and calibrate a version of this model to show that it can replicate meaningful shares of simultaneous over- and underreaction.

The third essay similarly relates to survey expectations. By way of example, I show that rational and non-rational models alike are able to deliver the same linear relationship between forecast errors and revisions. I specifically focus on a rational model of strategic interaction as well as non-rational models of overconfidence and diagnostic expectations. I propose examining the serial correlation of revisions instead as it is able to distinguish between these three models.

Contents

List of Tables

List of Figures

List of Abbreviations

ARED	....	Adaptively Rational Equilibrium Dynamics
FIRE	....	Full information rational expectations
JIT	....	Just-in-time
KF	....	Kalman Filter
MCMC	....	Markov Chain Monte Carlo
MSE	....	Mean square error
NAICS	....	North American Industry Classification System
NOAA	....	National Oceanic and Atmospheric Administration
PF	....	Particle Filter
SMM	....	Simulated Method of Moments
SNR	....	Signal-to-noise ratio
SPF	....	Survey of Professional Forecasters
TFP	....	Total Factor Productivity

Chapter 1

Spread Too Thin: The Impact of Lean Inventories

1.1 Introduction

Up to 70% of manufacturers have reportedly adopted just-in-time (JIT) production, a management philosophy that aims to minimize the time between orders.[1] Firms adopt JIT in an effort to cut costs associated with managing large material purchases and storing idle stocks. Instead, these firms commit to placing smaller more frequent orders from suppliers.[2] Consequently, lean inventory management has contributed to the 11% reduction in aggregate inventory holdings as a share of sales from 1992-2019.[3]

Do improvements in inventory management matter for macroeconomic fluctuations? Theoretically, in general equilibrium, inventories have been found to be immaterial for aggregate dynamics (Iacoviello, Schiantarelli, & Schuh, 2011; Khan & Thomas, 2007). Empirically, some find that inventory management improvements decreased aggregate volatility (Davis & Kahn, 2008) while others (Stock & Watson, 2002) find that it was broadly inconsequential.

This paper offers a new perspective on the role of lean inventories in driving aggregate fluctuations, finding that it can create macro fragility in the face of unexpected

[1] In 2015, the Compensation Data Manufacturing & Distribution Survey found that 71% of surveyed firms employ lean manufacturing. Similarly, in 2007, the Industry Week/MPI Census of Manufacturers found that 70% of respondents had implemented lean manufacturing.

[2] Ohno (1988) provides a detailed history of JIT which first started with Toyota's Kanban system.

[3] U.S. Census Bureau, Total Business: Inventories to Sales Ratio [ISRATIO], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/ISRATIO.

shocks such as COVID-19. I document evidence of this tradeoff from a novel dataset of JIT firms and quantitatively assess the role that lean production plays at the micro and macro levels in a structurally estimated heterogeneous firms model.

I first provide firm-level evidence linking the JIT adoption decision to higher firm sales and lower firm volatility. This provides motivating evidence as well as a set of moments that I use when structurally estimating the model. Within firms, JIT adoption is associated with a 16% decrease in inventory-to-sales ratios and a 9% increase in sales. In addition, JIT firms experience an 8-9% decline in employment and sales growth volatility. These empirical results, though not causal, are consistent with positive selection into adoption which subsequently yields firm-level efficiency gains as in my model.

I then exploit variation external to the firm and document that JIT adopters are exposed to the business cycle and other unexpected aggregate events. At the firm level, sales growth among JIT firms comoves more closely with GDP growth than non-JIT firms. JIT firms are estimated to be between 50-70% more cyclical than non-JIT firms. In addition, JIT adopters experience a 4 percentage point sharper drop in sales growth when their suppliers face unexpected weather disasters (Barrot & Sauvagnat, 2016). My analysis points to heightened sensitivity among JIT firms upon the realization of external shocks, indicating that an economy composed of more JIT producers is less resilient to such disturbances.

In light of these empirical facts, I build and structurally estimate a dynamic general equilibrium model of JIT production. The model features a rich distribution of firms that differ in their idiosyncratic productivity, inventory holdings, and inventory management strategy. Materials are needed for production and can be acquired subject to a stochastic fixed order cost. JIT firms draw order costs from a distribution that is first order stochastically dominated by those of non-JIT firms. Implementing

JIT requires incurring a fixed initial adoption cost and a smaller continuation cost thereafter. In a given period, firms must choose their JIT adoption status, whether to order materials, and how much to produce.

I numerically solve and structurally estimate the model via the simulated method of moments (SMM) based on data from 1980-2018. Relative to a counterfactual economy with less JIT, re-estimated from data during the 1980s, the estimated model yields a welfare gain of 0.6%.[4] In addition, the estimated model delivers a 0.2% increase in measured TFP in the steady state. Intuitively, JIT adoption leads to an economy-wide reduction in fixed order costs which enables adopters to better align their material input use with their realized productivity. As a result, measured aggregate productivity rises as firms smooth out their inventory cycles, leading to a 9% reduction in firm-level sales volatility, consistent with the decline observed in the micro data.

Whereas individual adopters benefit from JIT in normal times, the existence of leaner firms renders the economy more vulnerable to unexpected shocks. I consider an unanticipated productivity shock calibrated to match the drop in real US GDP during the onset of the COVID-19 pandemic. This unforeseen supply chain disruption mimics the nature of the COVID-19 shock, while relating more generally to my empirical evidence on weather disasters.[5]

Relative to the counterfactual economy, the JIT economy experiences a higher frequency of stock outs and a more gradual depletion of inventories. Since JIT firms store fewer materials in their plants, an unexpected spike in the price of materials makes them more susceptible to stocking out. At the same time, as the price of material inputs rises, inventories are suddenly more highly prized, with an increase

[4]For comparison, this welfare figure lies in the estimated range of welfare costs of business cycles and is similar to the welfare cost of managerial short-termism.

[5]This is also consistent with other work modeling COVID-19 as an unexpected shock (Arellano, Bai, & Mihalache, 2020; Espino, Kozlowski, Martin, & Sanchez, 2020).

in the shadow value of inventories within the firm. As a result, producers that do not fully stock out cut back on material input use in an effort to draw inventories down more slowly. The utilization of fewer material inputs in production in the JIT economy due to stock outs and hoarding leads to a sharper drop in output relative to the counterfactual model.

In short, my empirical and theoretical analysis reveals and quantifies a stark tradeoff between steady state gains and macro vulnerability. Firms benefit in normal times from pursuing a lean inventory strategy, however upon the realization of an unanticipated shock, an economy populated by more JIT firms experiences a deeper crisis than one with fewer lean producers. In this sense, inventories can serve as a stabilizing force.

Inventory investment has long been of interest to economists as a potential source of macroeconomic volatility.[6] Seminal contributions developed production smoothing models (Eichenbaum, 1984; Ramey & Vine, 2004) and (S,s) models (Caplin, 1985; Scarf, 1960) of inventory investment. Khan and Thomas (2007) elegantly models inventories in a general equilibrium environment with heterogeneous firms and business cycle shocks. The authors find that inventories play little to no role in amplifying or dampening business cycles.[7] My model is similar with the addition of idiosyncratic productivity, an endogenous JIT adoption decision, and a focus on large unanticipated shocks. Moreover Bachmann and Ma (2016) highlights the role that inventories play in a lumpy investment model and argues that inventories can also speak to the macro implications of investment with non-convex adjustment costs (Bachmann, Caballero, & Engel, 2013). I add to our understanding of inventories with a quantitative exercise emphasizing an important tradeoff between micro and

[6]See for instance Ahmed, Levin, and Wilson (2002), McConnell and Perez-Quiros (2000), McCarthy and Zakrajsek (2007), Irvine and Schuh (2005), and McMahon and Wanengkirtyo (2015).

[7]Iacoviello et al. (2011) comes to a similar conclusion albeit through a different model. On the other hand, Wen (2011) builds a stock-out avoidance model and finds that inventories are stabilizing.

macro stability amid unexpected disasters.

In addition, this paper relates to a strand of the management literature that focuses on assessing the gains to JIT. Kinney and Wempe (2002) finds that JIT adopters outperform non-adopters, primarily through profit margins. Nakamura, Sakakibara, and Schroeder (1998) as well as Roumiantsev and Netessine (2008) find similar evidence. Gao (2018) examines the role of JIT production in corporate cash hoarding. My paper provides a bridge between evidence documented in the management literature and the rich literature on inventories in macroeconomics by highlighting how JIT production matters for aggregate outcomes.

Furthermore, this paper relates to the literature on supply chain disruptions. On the empirical front, I adopt a strategy similar to Barrot and Sauvagnat (2016) to determine whether JIT producers are disproportionately exposed to unexpected weather disasters. Other empirical work has assessed how shocks propagate through a network of firms. For instance, Carvalho, Nirei, Saito, and Tahbaz-Salehi (2016) does this in the context of the 2011 Japanese earthquake. Similarly, Cachon, Randall, and Schmidt (2007) assesses empirical evidence of the bullwhip effect along the supply chain. From a theoretical perspective, my paper relates to models of heterogeneous firms, sunk costs, and supply chains. Meier (2020) models supply chain disruptions in the context of time to build. Moreover, I model the JIT adoption decision in a manner similar to Alessandria and Choi (2007) who model path dependent export decisions. My paper explicitly links supply chain disruptions to an important source of investment at the macro level, inventory accumulation.

The rest of the paper is organized as follows. Section 1.2 documents evidence that is consistent with the stabilizing effects of JIT at the firm level along with the exposure to unexpected shocks that it engenders at the macro level. Sections 1.3 and 1.4 develop the general equilibrium model of lean production. I estimate the model in Section 1.5. Section 1.6 quantifies the aforementioned micro-macro tradeoff associated with JIT, and Section 1.7 concludes.

1.2 Empirical Patterns Among JIT Firms

Before presenting the model of JIT production, I document empirical evidence that JIT adopters are more efficient and yet are more exposed to external shocks. I use this as motivating evidence for the model that I present in Section 1.3. This analysis will also provide moments and external validation to the model once I structurally estimate it.

I first gather firm-level information by making use of Compustat Fundamentals Annual data for manufacturers (NAICS 31-33) from 1980-2018. I merge these data with information on county-level weather events from the National Oceanic and Atmospheric Administration (NOAA) with specific links from Barrot and Sauvagnat (2016). In addition, I develop a new measure of JIT adoption among publicly traded manufacturers by updating and extending previous work in the literature (Gao, 2018; Kinney & Wempe, 2002). This is done through an exhaustive analysis of news reports and SEC filings. Following the literature, I search these documents for key words such as "JIT," "just-in-time," "lean manufacturing," "pull system," and "zero inventory." I then analyze each of these documents to confirm the year of adoption and to ensure that the firm in question implements JIT rather than any suppliers potentially mentioned in the announcements. In all, my dataset identifies the years in which approximately 185 Compustat manufacturers adopted JIT.[8] My final sample consists of an unbalanced panel of 5,099 unique manufacturing firms spanning the

[8] While the information on JIT adoption assuredly precludes any false positives, the limited nature of these documents across the thousands of manufacturers in Compustat leaves open the potential for false negatives in my sample. I account for the possibility of measurement error when modeling JIT by incorporating a parameter that in part governs the observed frequency of adoption. Section 1.5 discusses this in further detail.

Table 1.1: JIT Adoption and Profitability

	(1) Inventory-to-sales	(2) Sales
Adopter	-0.155*** (0.035)	0.092** (0.025)
Fixed effects	Firm, Industry × Year	Firm, Industry × Year
Observations	37,154	37,154

Notes: The table reports panel regression results from Compustat Annual Fundamentals of manufacturing firms (NAICS 31-33) based on regression (1.1). The regressor of interest is the firm-year specific adoption indicator. Standard errors are clustered at the firm level. The standard deviations of the dependent variables are 0.68 and 2.15, respectively. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

aforementioned time period.[9]

Using these data, I document four facts about JIT adopters.[10] First, JIT adoption is associated with lower inventory holdings and higher sales.[11] I estimate:

$$y_{ijt} = \gamma \text{adopter}_{ijt} + \delta_{jt} + \delta_i + \nu_{ijt}, \tag{1.1}$$

where y_{ijt} is an outcome variable for firm i belonging to 6-digit NAICS manufacturing industry j in year t. I specify the outcomes to be log inventory-to-sales ratio and log sales. The regressor of interest, adopter_{ijt}, is a time-varying indicator for whether firm i is a JIT adopter in a given year.

Table 1.1 reports the regression results. Adopters experience a 16% decrease in inventory-to-sales ratios and a 9% increase in sales. The results imply a change of -23% and 4% of one standard deviation in the outcomes, respectively. The regression results allude to the benefits of JIT in the model. Facing lower fixed order costs,

[9] Appendix A.1 provides summary statistics of the data.

[10] Appendix A.1 documents similar facts at the industry level, indicating that these patterns do not wash out with aggregation.

[11] This is consistent with Fullerton and McWatters (2001) and Cua, McKone, and Schroeder (2001).

Table 1.2: JIT Adoption and Variance of Outcomes

	(1) Employment growth variance	(2) Sales growth variance
Adopter	-0.079*	-0.087**
	(0.041)	(0.032)
Fixed effects	Firm, Industry × Year	Firm, Industry × Year
Observations	16,055	16,055

Notes: The table reports panel regression results from Compustat Annual Fundamentals of manufacturing firms (NAICS 31-33) based on regression (1.1). The regressor of interest is the firm-year adoption indicator. Standard errors are clustered at the firm level. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

adopters hold fewer inventories in favor of placing smaller more frequent orders. Upon shrinking their inventory stocks, adopters also incur fewer carrying costs. These cost reductions lead adopters to allocate more resources to production.

Second, JIT adopters experience less micro volatility. I re-estimate (1.1) where y_{ijt} now denotes a rolling 5-year standard deviation of sales growth and employment growth for firm i in industry j in year t. Table 1.2 reports the results. Adopters see an 8% decline in employment growth volatility and a 9% decline in sales growth volatility. This is consistent with the stabilizing role that JIT plays in the model. As firms smooth out their inventory cycles due to the lower fixed order costs, they moderate the variability of other outcomes as well.

I next document facts relating to firm-level exposure brought on by JIT, exploiting aggregate variation and examining sensitivity to a set of specific events such as macro fluctuations and weather disasters. The regression results accord with the model in that adopters are less insured against unanticipated disruptions, and an economy with more JIT firms is more exposed to aggregate shocks.

Third, JIT adopters tend to be more cyclical. I quantify this via regressions that

Table 1.3: JIT Adoption and Cyclicality

	Sales growth	Sales growth
GDP growth	1.017***	
	(0.070)	
Adopter × GDP growth	0.710***	0.476*
	(0.197)	(0.271)
Controls	Yes	Yes
Fixed Effect	Firm, Industry	Firm, Industry × Year
Observations	32,881	28,665

Notes: The table reports regression results from Compustat Annual Fundamentals of manufacturing firms (NAICS 31-33) based on regression (1.2). The independent variable of interest is the interaction between the adopter indicator and GDP growth. Control variables include logs of sales per worker, firm size, cash-to-assets, and cost of goods-to-sales, as well as the adoption indicator. Column (1) reports results without year fixed effects. Column (2) includes year fixed effects. Standard errors are clustered at the firm level. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% signficance.

interact adoption with GDP growth:

$$y_{ijt} = \gamma_1 \text{adopter}_{ijt} + \gamma_2 \text{GDPgrowth}_t + \gamma_3 \left[\text{adopter}_{ijt} \times \text{GDPgrowth}_t\right] + \mathbf{X}'_{ijt}\beta + \delta_i + \delta_j + \varepsilon_{jt}, \tag{1.2}$$

where $\mathbf{X}$ denotes a set of controls that include sales per worker, firm size, cash holdings, and cost of goods sold-to-sales. The coefficient γ_3 measures the extent to which firms exhibit more cyclicality. Table 1.3 reports the regression results. Based on column (1), a 1% increase in GDP growth is associated with a roughly 1% increase in sales growth among non-adopters. Adopters experience an additional sales growth increase of 0.7% above this baseline. After also controlling for industry trends, I find that adopters are about 50-70% more cyclical than non-adopters.

Fourth, JIT adopters are more sensitive to local weather events. I examine this

Table 1.4: JIT Adoption and Sensitivity to Local Disasters

	Sales growth
Adopter	0.104***
	(0.030)
Total upstream disasters	-0.032*
	(0.018)
Adopter × Total upstream disasters	-0.044**
	(0.019)
Fixed Effects	Firm, Industry × Year
Observations	1,192

Notes: The table reports weather event regressions from a sample of Compustat manufacturing firms (NAICS 31-33) based on regression (1.3). The independent variable of interest is the interaction between the adoption indicator and total number of upstream disasters. Control variables include number of upstream suppliers and sales per worker, cost of goods-to-sales, and inventory-to-sales for both the downstream firm and its average upstream supplier. Standard errors are clustered at the firm level. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

by estimating the following regression:

$$y_{ijt} = \psi_1 \text{adopter}_{ijt} + \psi_2 \text{disaster}_{ijt} + \psi_3 \left[\text{adopter}_{ijt} \times \text{disaster}_{ijt}\right] + \mathbf{X}'_{ijt}\beta + \delta_{jt} + \delta_i + \varepsilon_{ijt}. \tag{1.3}$$

The disaster_{ijt} variable denotes the log of total upstream disasters faced by firm i residing in industry j in year t. I collect information on weather disasters from NOAA and link these disasters to a firm's upstream suppliers' zip codes via the aforementioned Barrot and Sauvagnat (2016) links. The control variables specified in $\mathbf{X}$ include the number of upstream suppliers as well as a set of controls (sales per worker, cost of goods sold-to-sales, and inventory-to-sales) for the downstream firm and for the average of its upstream suppliers.

Table 1.4 reports the estimation results. Consistent with Table 1.1, adopters tend to be more profitable. Moreover, a 1% increase in the total number of disasters hitting a firm's suppliers tends to reduce downstream firm sales growth by around 3 percentage points, a roughly 16% standard deviation decrease. Adopters experience an additional 4 percentage point drop, making them more than twice as sensitive to upstream disasters as non-adopters.

Taken together, the data suggest that JIT adopters benefit from higher profits and smoother outcomes. At the same time, adoption is associated with heightened exposure to aggregate fluctuations and unanticipated shocks as proxied by local weather disasters. My model of heterogeneous firms with an endogenous JIT adoption decision can explain these patterns. The model also allows me to quantitatively assess the impact of JIT amid an unanticipated macro disaster, something that cannot be captured by firm level regressions.

1.3 A Model of Just-in-Time Production

Having illustrated the essence of the tradeoff in the data, I next build the full general equilibrium model which will provide quantitative statements about the implications of JIT. The model is similar in spirit to Khan and Thomas (2007) and Alessandria and Choi (2007), embedded with JIT and ultimately incorporating large unanticipated disasters rather than traditional business cycle shocks.

A representative household has preferences over consumption and leisure. The household supplies its labor frictionlessly to the two sectors of the economy: the intermediate goods sector and the final goods sector. A representative intermediate goods firm produces materials by using labor and capital. In addition, a continuum of heterogeneous final goods firms make use of labor and materials to produce using a decreasing returns to scale technology. Final goods producers are heterogeneous in idiosyncratic productivity, inventory stocks, and JIT adoption status. All markets are perfectly competitive.

The representative household is endowed with one unit of time in each period and values consumption and leisure according to the following preferences:[12]

$$U(C_t, N_t^h) = \log(C_t) + \phi(1 - N_t^h),$$

where $\phi > 0$ denotes the household's labor disutility. Total hours worked is denoted by N_t^h and labor is paid wage, w_t. In addition to wage income, the household earns a dividend each period from ownership of firms, D_t, and chooses savings on a one period riskless bond, B_{t+1}, given interest rate R_{t+1}. The representative household, facing no aggregate uncertainty, maximizes its utility:

$$\max_{C_t, N_t^h, B_{t+1}} \sum_{t=0}^{\infty} \beta^t U(C_t, N_t^h),$$

subject to its budget constraint which holds for all t,

$$C_t + B_{t+1} \leq R_t B_t + w_t N_t^h + D_t.$$

The parameter $\beta \in (0, 1)$ is the household's subjective discount factor.

The representative intermediate goods firm produces materials using capital K_t and labor L_t according to:

$$F(K_t, L_t) = K_t^{\alpha} L_t^{1-\alpha}.$$

Taking prices as given, the problem of the intermediate goods firm is:

$$\max_{K_t, L_t} q_t F(K_t, L_t) - w_t L_t - R_t K_t,$$

where q_t denotes the price of the intermediate good.

Finally, a continuum of final goods firms produce using materials, m_t, and labor,

[12] Rogerson (1988) microfounds these preferences in a model of indivisible labor and lotteries. These preferences provide tractability and are common in the literature, e.g. Gilchrist, Sim, and Zakrajsek (2014); Ottonello and Winberry (2020); Senga (2018).

n_t, according a decreasing returns to scale technology:

$$y_t = z_t m_t^{\theta_m} n_t^{\theta_n}, \quad \theta_n + \theta_m < 1,$$

where idiosyncratic productivity evolves as an AR(1) in logs:

$$\log(z_{t+1}) = \rho_z \log(z_t) + \sigma_z \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, 1).$$

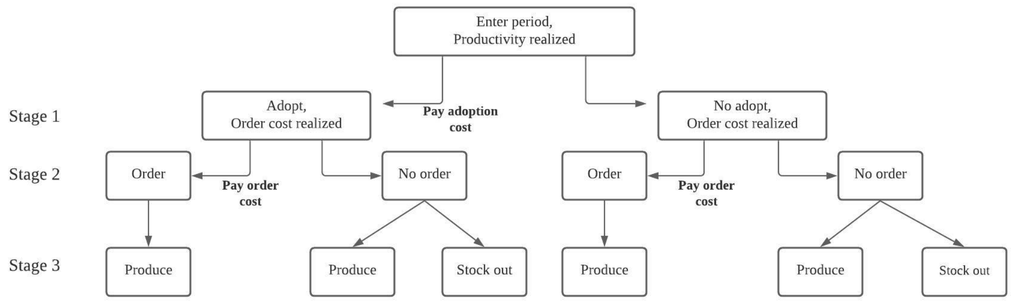
Materials are drawn from the firm's existing inventory stock, s_t, to use in production. Final goods firms procure new materials from the intermediate goods firm subject to a stochastic fixed order cost drawn from a uniform distribution.

Figure 1·1 details the final goods producers' decision-making timeline. Each period is broken into three stages. A producer enters the period with realized productivity, z_t, inventory stock, s_t, and adoption status, a_t. In the first stage, the producer decides whether or not to adopt JIT. If a producer does not enter the period as a continuing adopter, it must pay c_s in order to initially adopt. Alternatively, if the producer enters the period as an adopter, it must pay a smaller continuation cost $c_f < c_s$ in order to maintain its status as a JIT adopter.

Intuitively, adopting JIT requires that a plant repurpose its shop floor, enter into long-term contracts with suppliers to fulfill orders in a timely fashion, and possibly even purchase new technologies to share information with suppliers. The sunk setup cost, c_s, encompasses all of these one-time costs. The continuation cost, c_f, embodies smaller costs for suppliers to participate in timely delivery, costs of training labor on JIT practices and tasks, and greater attention or communication required to share information with suppliers.

In the next stage, producers learn their order cost, $\xi \sim U(0, \overline{\xi})$, and decide whether or not to place an order, o_t. JIT producers face a more favorable order cost distribution, $\overline{\xi}_A < \overline{\xi}_{NA}$. Lastly, following the adoption and the order decisions, final goods

Figure 1·1: Decisions of Final Goods Firms



Notes: The figure summarizes the order of the decisions made by final goods firms within a period.

producers decide how much to produce.

I characterize the final goods firms' problem in terms of inventory stocks rather than specific order or material input choices. In particular, if a firm enters the period with inventory stock s_t, its target inventory stock is denoted by s_t^*. This means that any orders (if placed) are defined as $o_t = s_t^* - s_t$. Following the order decision, suppose that inventory stock $\widetilde{s}_t$ is carried into the production stage. Materials used in production are then defined as $m_t = \widetilde{s}_t - s_{t+1}$ where s_{t+1} refers to the inventory stock carried forward into the next period. In what follows, I suppress the time subscript and instead denote next period variables with a prime.

Stage 1: Adoption Decision

A final goods producer begins the period with (z, s, a) and faces adoption costs $\{c_s, c_f\}$, denominated in units of labor and endogenous prices, p, q, and w. The firm first decides whether to adopt JIT. Note that the adoption status is a binary outcome. The value of adopting is:

$$V^A(z,s,a) = \max\left\{ -pwc(a) + \int V^O(z,s,1,\xi)dH(\overline{\xi}_A), \int V^O(z,s,0,\xi)dH(\overline{\xi}_{NA})\right\}, \tag{1.4}$$

where

$$c(a) = \begin{cases} c_s & \text{if no JIT } (a = 0) \\ c_f & \text{if JIT } (a = 1), \end{cases}$$

and $V^O(z, s, a, \xi)$ refers to the firm's value in the second stage. The firm's optimal adoption policy, $a'(z, s, a)$, solves this maximization.

Stage 2: Order Decision

Given the firm's order cost draw, ξ, denominated in units of labor, it then decides whether to place an order, o. If the firm is an adopter, its order cost distribution is first order stochastically dominated by those of non-adopters. The value in the second stage is[13]

$$V^O(z, s, a, \xi) = \max \left\{ -pw\xi + pqs + V^*(z, s, a, \xi), V^P(z, s, a) \right\}, \tag{1.5}$$

where the value of placing an order is

$$V^*(z, s, a, \xi) = \max_{s^* \geq s} \left[-pqs^* + V^P(z, s^*, a) \right], \tag{1.6}$$

and $V^P(z, s, a)$ is defined below. The firm's problem delivers a threshold rule for placing an order. In particular, a firm places an order if and only if the order cost draw is lower than a threshold order cost: $\xi < \xi^*(z, s, a)$ where

$$\xi^*(z, s, a) = \frac{pqs + V^*(z, s, a) - V^P(z, s, a)}{\phi}. \tag{1.7}$$

Stage 3: Production Decision

Upon making an adoption decision and choosing whether to place an order, and if so what size, the firm decides how much to produce. Suppose that a firm enters this

[13]The constraint on the order decision allows for only positive orders. In particular, the model abstracts away from inventory liquidation.

stage with inventory stock $\widetilde{s}$ such that:

$$\widetilde{s} = \begin{cases} s^*\big(z, s, a'(z, s, a)\big) & \text{if order placed} \\ s & \text{if no order placed.} \end{cases}$$

In the production stage, the firm selects labor, $n(z, \widetilde{s}, s', a)$, and materials, $(\widetilde{s} - s')$, to maximize profits. Its value function in the production stage is:

$$V^P(z, \widetilde{s}, a) = \max_{s' \in [0, \widetilde{s}]} \pi(z, \widetilde{s}, s', a) + \beta \mathbb{E}\big[V^A(z', s', a')\big] \tag{1.8}$$

where

$$\pi(z, \widetilde{s}, s', a) = p\big[zn(z, \widetilde{s}, s', a)^{\theta_n}(\widetilde{s} - s')^{\theta_m} - c_m s' - wn(z, \widetilde{s}, s', a)\big] \tag{1.9}$$

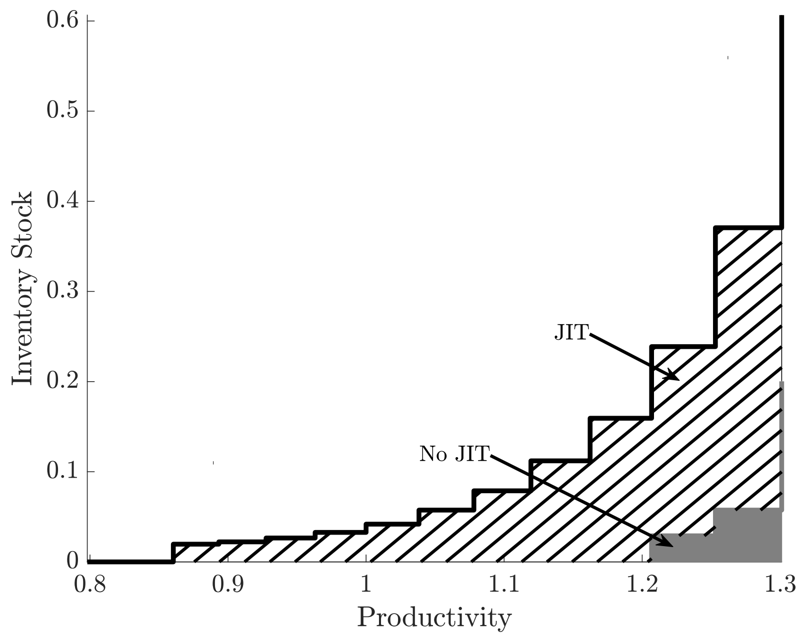
are period profits. The end of period inventory stock is denoted by s', and c_m is the carrying cost of unused input inventory, denominated in units of output.

A final goods producer is said to stock out if it enters the period with no inventories, $s = 0$, and chooses to not place an order. Without inventories the firm has no material inputs to make use of in the production stage. As a result, it foregoes production in that period, but can restart production in the future conditional on a favorable order cost draw.

1.4 Analyzing the Model

The endogenous adoption decision allows the model to replicate important features of the data, namely, higher profitability and reduced micro volatility among JIT firms. Since implementing JIT comes at a relatively large sunk cost, not all firms optimally choose to adopt JIT. Figure 1·2 plots the adoption frontiers for JIT and non-JIT producers. The black shaded area by the bottom right corner represents the region of the state space in which non-JIT firms choose to adopt JIT.

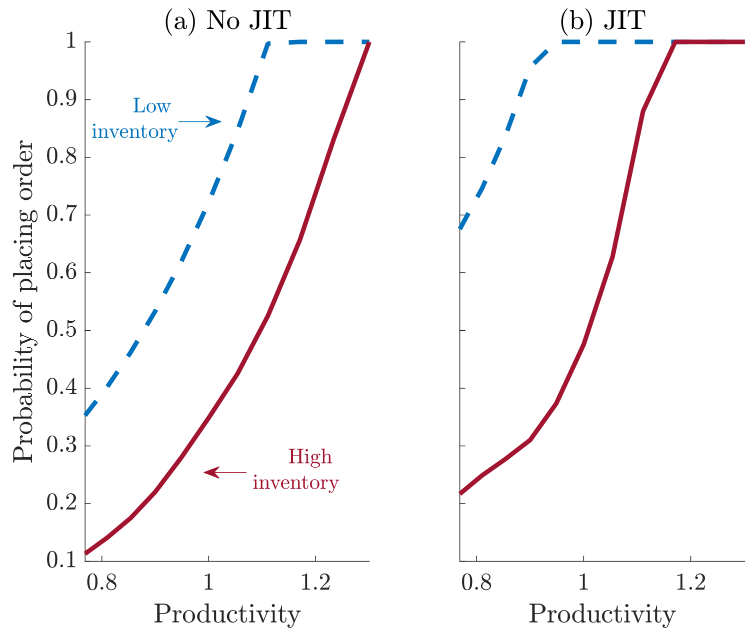
Figure 1·2: Adoption Frontiers



Notes: The figure plots the adoption frontier among JIT and non-JIT firms. The solid shaded area plots the region of the state space in which non-JIT firms select into adoption. The striped area along with the shaded area jointly denote the region of the state space in which existing JIT firms choose to remain adopters.

At the same time, a producer is likely to remain an adopter conditional on already being one. This is because the continuation cost of retaining JIT is smaller than the initial sunk cost. Hence, the endogenous adoption decision exhibits persistence. The larger blue shaded area of Figure 1·2 confirms this intuition. Only the least productive JIT producers will opt to abandon adoption. Furthermore, the scope for exiting adoption is increasing in inventory holdings. The selection detailed here could contribute to the patterns among JIT firms documented in the data. In particular, the decision to adopt JIT reflects a favorable productivity realization which, when coupled with lower average order costs, leads firms to reduce inventory stocks and incur fewer carrying costs thereby generating more output.

Figure 1·3: Order Probabilities

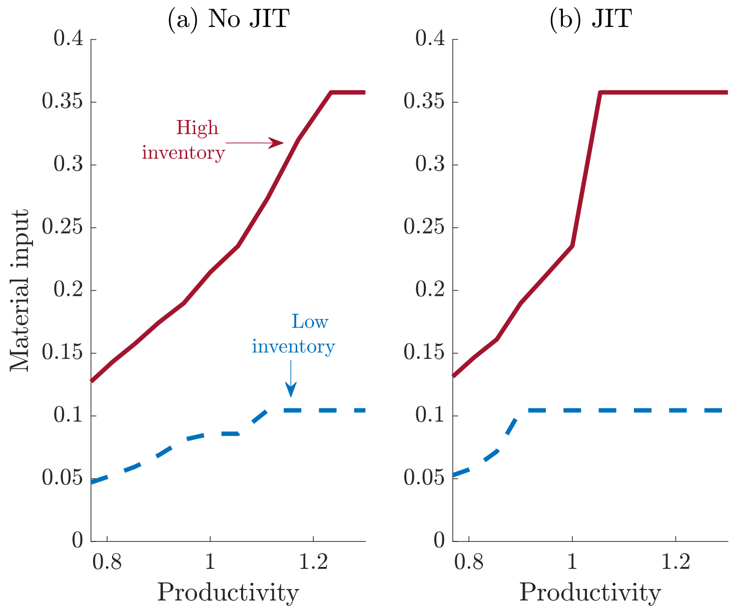


Notes: The figure plots the probability of placing an order in the order stage as a function of productivity. Panel (a) plots the probabilities among non-adopters and panel (b) plots the probabilities for adopters. The solid red line reflects a high inventory establishment in the model while the dashed blue line reflects a low inventory establishment.

Figure 1·3 shows the probability of placing an order as a function of productivity. Consistent with the decision to select into adoption, order probabilities are increasing in productivity and decreasing in inventory holdings. Moreover, the benefits of JIT adoption can be understood by comparing the two panels. Across both inventory levels, the probability of placing an order is higher for adopters since they face lower average order costs. As a result, adopters in the model place smaller and more frequent orders. This is consistent with the reduction in inventory holdings within adopters.

Figure 1·4 plots material usage as a function of productivity. Material inputs are increasing in productivity and inventory holdings. Firms with very low inventory stocks will tend to exhaust their remaining inventories regardless of their level of

Figure 1·4: Material Usage



Notes: The figure plots material usage policy functions in the production stage as a function of productivity. Panel (a) plots the policy among non-adopters and panel (b) plots the policy for adopters. The solid red line reflects a high inventory establishment in the model while the dashed blue line reflects a low inventory establishment.

productivity. Furthermore, adopters make greater use of materials when producing thereby raising output. The flat lines in these policies reflect endogenous decisions to fully utilize existing inventory stocks in production. Because adopters can restock more flexibly, due to the lower order costs, they exhaust their inventory stocks more often. As a result, production among JIT firms tends to be uninterrupted despite their lower inventory holdings. Both the order threshold and the material input policy reflect a treatment effect that allows firms to produce at lower costs which in turn raises firm sales following adoption.

A comparison of outcomes between economies that differ only in the option to adopt JIT confirms the model-implied benefits to lean production: higher sales and less volatility. Figure 1·5 visualizes simulation results from such an exercise. The figure plots a plant's simulated path in both models. The plant in each economy faces the same productivity realizations.

Upon adopting JIT, the establishment retains its status as an adopter through the rest of the simulated path despite lower productivity realizations in the latter periods. This enables the establishment to undertake production despite holding fewer inventories. The cost savings associated with JIT allow the firm to redirect its resources to production rather than order placing or inventory storage. As a result, sales are higher among JIT firms.

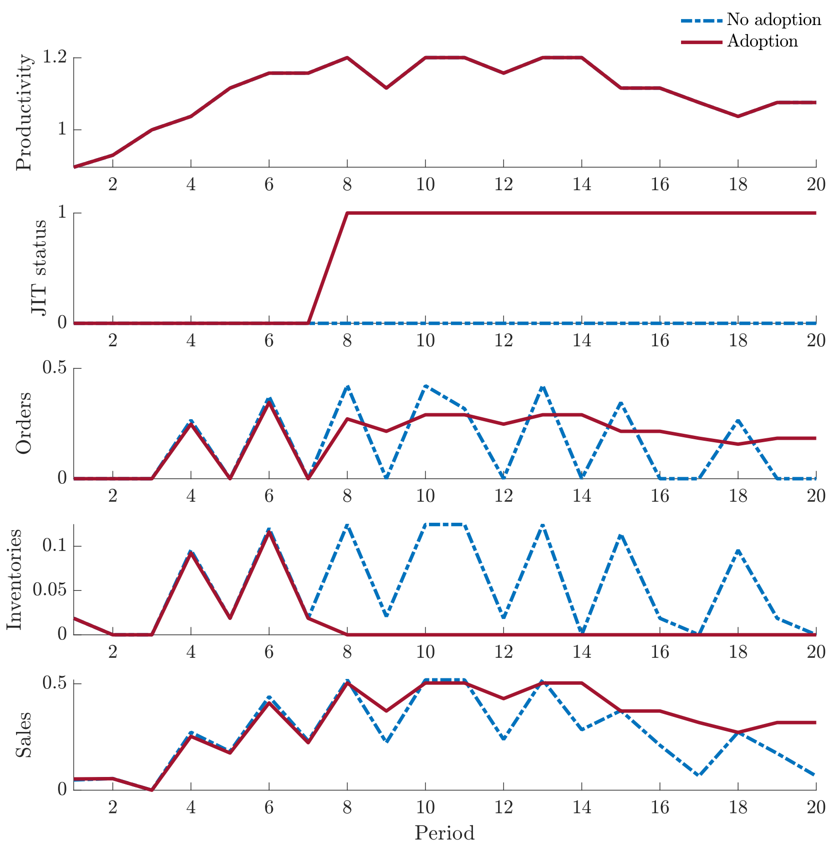
Furthermore, upon adopting JIT, the plant's simulated path for orders is smoothed considerably relative to the economy without adoption. This illustrates the insight that JIT mutes the inventory cycle. Because adopters face lower fixed order costs, their target inventory stocks are lower in the JIT model and the frequency of placing an order increases. The smoother path for orders also smooths firm sales which can explain the lower variance of outcomes among adopters in the data.

The fourth panel of the figure, however, points to the source of exposure to unexpected shocks. Since adopters hold little to no inventory stocks, an unanticipated supply chain disruption that prevents the establishment from acquiring materials in a given period will disrupt its production process. On the other hand, the same plant in the no JIT economy will on average have a larger buffer of inventories from which to draw when producing amid such an event.

1.5 Structural Estimation

I structurally estimate the model using the micro data analyzed in Section 1.2. The estimated model captures important features of the firm-level data including the adoption frequency, levels of and covariances between inventories and sales, and spikes in inventory holdings. Importantly, the estimated model allows me to quantify benefits

Figure 1·5: Adoption Mutes the Inventory Cycle



Notes: The figure plots the path of a selected establishment in the unconditional simulation. The top panel plots the (shared) path of idiosyncratic productivity across both models. The second panel plots the plant's JIT adoption status, the third panel plots orders, the fourth plots inventory holdings, and the bottom panel plots sales.

to JIT in normal times as well as the vulnerabilities that it exposes to unanticipated macro shocks.

The comprehensive search of firm financials and public statements ensures that the data on JIT adoption do not include false positives. However, information on JIT implementation is constrained to what is reported in these records. To allow for the possibility that JIT is more widespread than the empirical frequency of adoption in my sample, I use the structure of the model in order to infer patterns of adoption. I do so by defining a parameter, $\tau \in [0, 1]$, that governs the share of observed non-adopters from a simulated panel of firms.[14]

Of the 13 model parameters listed in Tables 2.5 and 1.6, I first externally calibrate five of them based on standard parameterizations in the literature. Table 2.5 details the annual calibration. The discount factor, β is set to 0.962 which is consistent with a real rate of 4%. The material share, θ_m, and the capital share, α, are set match their counterparts in the NBER-CES database for manufacturers from 1980-2011. The parameter θ_n is set to match an economy-wide labor share of 0.65. The leisure preference is calibrated so that the household works about one-third of the time.

1.5.1 Simulated Method of Moments

The parameter vector to be estimated is $\theta = \left(\rho_z \;\; \sigma_z \;\; \overline{\xi}_{NA} \;\; \overline{\xi}_A \;\; c_s \;\; c_f \;\; c_m \;\; \tau\right)'$. These parameters residing in θ govern the exogenous productivity process, the stochastic orders costs, the sunk and carrying costs, and the share of observed non-JIT firms. The model has no closed form solution, so I solve it using standard numerical dynamic programming techniques detailed in Appendix A.1. To parameterize the model, I employ SMM (Bazdresch, Kahn, & Whited, 2018; Duffie & Singleton, 1993). This is done by computing a set of targeted moments in the model and minimizing the

[14]As in my sample, a firm in the model is said to be an adopter if at least one of its establishments adopts JIT. Upon simulating a panel of firms, a share τ, are designated non-adopters irrespective of their true adoption status.

Table 1.5: External Parameterization

Description	Parameter	Value	Notes
Discount Factor	β	0.962	Real rate equal to 4%
Material share	θ_m	0.520	NBER-CES (1980-2011)
Capital share	α	0.350	NBER-CES (1980-2011)
Labor share	θ_n	0.245	Labor share equal to 0.65
Labor disutility	ϕ	2.400	Work one third of time

Notes: The table reports the five calibrated parameters for the model.

weighted distance between the empirical moments and their model-based analogs.

Specifically, I target 11 moments to estimate the eight parameters. My estimator is therefore an overidentified SMM estimator. The first targeted moment is the empirical frequency of adoption. Of the remaining ten moments, five are specific to JIT firms and five to non-JIT firms. These five moments, which are the same across both types of firms, are: the mean inventory-to-sales ratio, the covariance matrix of inventory-to-sales ratios and log sales which deliver three moments, and the frequency of positive inventory-to-sales ratio spikes, defined as instances in which the inventory-to-sales ratio exceeds 0.25.[15] I specify the asymptotically efficient choice of the weighting matrix which is the inverse of the covariance matrix of the moments.

1.5.2 Informativeness of Moments

The choice of moments is crucial for the identification of the parameters, so I discuss their informativeness in turn. While the targeted moments jointly determine the parameters to be estimated, there are nonetheless moments that are especially informative in pinning down certain parameters.

Idiosyncratic productivity persistence mostly informs the covariance between inventory-

[15]The empirical moments are listed in the third column of Table 1.7.

to-sales and log sales among adopters. An increase in ρ_z implies that a firm with a favorable productivity realization will select into adoption, reduce its inventory holdings due to the lower average fixed order costs, and experience higher sales. As a result, the covariance between inventory-to-sales and log sales among adopters becomes more negative. Moreover, idiosyncratic productivity dispersion mostly affects variances, for instance the variance of inventory-to-sales among non-adopters, as an increase in σ_z results in more dispersed outcomes among producers.

An increase in the upper support of the order cost distribution for non-adopters primarily raises the inventory-to-sales ratio for JIT firms. Intuitively, higher order costs among non-JIT producers expands the area representing the adoption frontier for non-JIT producers in Figure 1·2, leading more firms to select into adoption. These new adopters are less productive and hold more inventories which raises the overall inventory-to-sales ratio among adopters. On the other hand, an increase in the upper support of the order cost distribution for adopters primarily raises the inventory-to-sales ratio among non-JIT firms. When average fixed order costs rise among adopters, the returns to continued adoption fall leading less productive and more bloated producers abandon JIT thereby raising inventory-to-sales ratios among non-adopters.

An increase in the sunk cost of adoption primarily raises the variance of log sales among adopters since it raises the returns to continued adoption. In other words, existing adopters are willing to tolerate more dispersed productivity realizations before abandoning JIT. An increase in the continuation cost of adoption primarily reduces the scope for remaining a JIT producer. More bloated, less productive firms will therefore abandon JIT, leading to an increase in spike rates among non-adopters. On the other hand, an increase in the carrying cost of inventories leads all firms to lean out, thereby reducing spike rates. Finally, a rise in the the share of observed

Table 1.6: Estimated Parameters

Description	Parameter	Estimate
Idiosyncratic productivity persistence	ρ_z	0.878 (0.059)
Idiosyncratic productivity dispersion	σ_z	0.044 (0.017)
Order cost distribution (non-adopters)	$\overline{\xi}_{NA}$	0.483 (0.029)
Order cost distribution (adopters)	$\overline{\xi}_{A}$	0.047 (0.010)
Sunk cost of adoption	c_s	0.293 (0.087)
Continuation cost of adoption	c_f	0.110 (0.012)
Carrying cost	c_m	0.182 (0.046)
Observed share of non-adopters	τ	0.938 (0.012)

Notes: The table reports the estimated parameters with standard errors in parentheses.

non-adopters primarily reduces the frequency of adoption.

Figure A1 in Appendix A.1 outlines these key monotonic relationships between the moments and the parameters. In addition, Figure A2 helps assess the sources of identification by reporting the sensitivity of each of the eight parameters to changes in a given moment, based on Andrews, Gentzkow, and Shapiro (2017). These figures confirm the intuition laid out above.

1.5.3 Estimation Results

Table 1.6 reports the estimated parameters, all of which are precisely estimated. The estimated technology parameters, ρ_z and σ_z, are consistent with parameterizations in the literature (Hennessy & Whited, 2007; Khan & Thomas, 2008; Meier, 2020), collectively ranging from 0.68-0.89 and 0.02-0.12 respectively. My estimates imply a more persistent and less dispersed idiosyncratic productivity process than that estimated in Clementi, Castro, and Lee (2015) which is likely due to the fact that my sample consists of publicly traded Compustat manufacturers who are larger and older than the universe of manufacturers.

The upper support of the order cost distribution among non-adopters is estimated to be an order of magnitude larger than that of adopters. These order cost estimates imply that non-JIT firms place orders that are about five times larger than those of JIT firms, indicating a sizable return to adoption for those who can initiate it. Furthermore, the adoption cost estimates suggest a great deal of hysteresis in the adoption decision. In particular, firms pay a continuation cost that is slightly more than one third of the original sunk cost. Conditional on being an adopter, the probability of remaining an adopter is 91%. This estimate is similar to estimates of the sunk cost of exporting, which place the probability of remaining an exporter conditional on already being one at 87% (Alessandria & Choi, 2007). The estimated carrying cost is about 15% of the total value of sales, a non-negligible amount that prevents firms from storing too many inventories across periods. Lastly, the estimated share of observed non-adopters implies that the mass of JIT establishments in the model's steady state is about 0.15.

Given that I target 11 moments to estimate the eight parameters, the model is overidentified and will not exactly match the empirical moments. With that said, the overidentified SMM procedure fits the data well. Table 1.7 compares the 11 targeted moments generated by the model with their empirical values. Importantly, the model replicates important features between adopters and non-adopters. Relative to non-JIT firms, adopters hold fewer inventories as a share of their sales. In addition, adopters are broadly characterized by less variable outcomes and a looser association

Table 1.7: Model vs. Empirical Moments

Moment	Model	Data
Mean(inventory-sales ratio|adopter)	0.169	0.146 (0.005)
Mean(inventory-sales ratio|non-adopter)	0.191	0.194 (0.002)
z Std(inventory-sales ratio|adopter)	0.059	0.042 (0.0002)
Corr(inventory-sales ratio, log sales|adopter)	-0.185	-0.215 (0.001)
Std(log sales|adopter)	0.234	0.189 (0.014)
Std(inventory-sales ratio|non-adopter)	0.071	0.067 (0.0001)
Corr(inventory-sales ratio, log sales|non-adopter)	-0.374	-0.328 (0.0004)
Std(log sales|non-adopter)	0.277	0.263 (0.005)
Spike(inventory-sales ratio|adopter)	0.089	0.071 (0.015)
Spike(inventory-sales ratio|non-adopter)	0.217	0.223 (0.005)
Frequency of adoption	0.048	0.050 (0.005)

Notes: The table reports model-based and empirical moments with standard errors in parentheses.

between inventory-to-sales ratios and log sales. Lastly, adopters exhibit fewer spikes in inventory holdings relative to their sales.

As JIT has become more common over time, an economy with fewer adopters is a natural benchmark against which to compare the estimated model. I exploit the earlier years of my sample, 1980-1989, in order to define this counterfactual. Specifically, I hold all parameters of the estimated model fixed except for the adoption costs c_s and c_f. I estimate these two costs based on the earlier period of my sample. The resulting estimates for the adoption costs are $c_s = 0.237$ and $c_f = 0.123$, which implies a lower frequency of adoption. The model reflecting these earlier-period adoption costs will serve as my counterfactual comparison for the estimated economy throughout the discussion below.[16] [17]

1.5.4 Nontargeted Moments

To further assess the estimated model's ability to reproduce the patterns present in the data, I run empirical regressions based on a panel of simulated firms from both the estimated and counterfactual models. The results are reported in Table 1.8. The regressions in Panel A are identical to those in Table 1.1 while the regressions in Panel B are identical to those in Table 1.2.

Following adoption, the estimated model is able to successfully reproduce reductions in inventory-to-sales ratios. The OLS coefficients from both models reside within the 95% confidence interval for the point estimate in the data of -0.155. In addition, the estimated and the counterfactual models both predict an increase in sales among adopters, with the estimated model delivering a closer match to the empirical coef-

[16] Appendix A.1 fully details the subperiod estimation results and counterfactual economy parameterization. Appendix A.1 also describes an alternate counterfactual in which I re-estimate the order costs in addition to the adoption costs with no meaningful changes to the results.

[17] Intuitively, this counterfactual captures changes in the incentives to adopt JIT based on observed successes of Toyota's Kanban system, or new cohorts of managers having been trained on this type of production system.

Table 1.8: Model-Based Regressions

Panel A: Levels	Inventory-to-sales	Sales
Data	-0.155 (0.035)	0.092 (0.025)
Estimated	-0.142 (0.003)	0.098 (0.003)
Counterfactual	-0.165 (0.003)	0.123 (0.002)
Panel B: Volatility	Employment growth	Sales growth
Data	-0.079 (0.041)	-0.087 (0.032)
Estimated	-0.099 (0.007)	-0.087 (0.007)
Counterfactual	-0.117 (0.007)	-0.098 (0.007)

Notes: The table reports empirical and model-based panel regressions at the firm level from the estimated and counterfactual models with standard errors in parentheses. Panel A reports regression results as in Table 1.1. Panel B reports regression results as in Table 1.2.

ficient. Moreover, both models predict reductions in firm volatility among adopters, with the estimated model providing a closer fit to the coefficient of employment growth volatility and an exact match to sales growth volatility. With precisely estimated parameters delivering a broadly successful fit to the data, and a relevant counterfactual defined, I can now exploit this structure as a laboratory for quantitative experiments.

1.6 Quantifying the Tradeoff

Having estimated the model, I proceed to quantify the tradeoff between the long-run gains to JIT and the vulnerability to unanticipated disasters that JIT exposes. I first examine the model's steady state to characterize the benefits of lean production. I then analyze the dynamics of the estimated economy following a COVID-19 disaster.

Table 1.9: Long-Run Aggregates Across Models

Panel A: Levels			
Output	Order frequency	Order size	Price of orders
0.79	7.65	-4.68	0.72
Inventory stock	Firm value	Measured TFP	Welfare
-10.56	0.95	0.23	0.57
Panel B: Volatility			
	MP materials	Sales	Labor
	-5.01	-9.22	-8.91

Notes: The table reports steady state values of the estimated model relative to the counterfactual model, in percent deviations. Panel A reports the levels of aggregates. Panel B reports measures of firm volatility from unconditional simulation of 40,000 firms over 50 periods.

1.6.1 Steady State

A comparison between the two models points to sizable gains associated with JIT adoption. Table 1.9 reports the steady state in the estimated model relative to the counterfactual economy in percent deviations. The higher prevalence of adoption in the estimated model implies smaller, more frequent orders placed such that order demand rises.

As expected, inventory holdings fall in the estimated model. The reduction in inventories is due to a decrease in target inventory stocks across all producers.[18] Relative to the counterfactual, the estimated model delivers a roughly 11% decline in the real aggregate inventory-to-sales ratio, a figure that can account for the observed reduction in the macroeconomic time series from 1992-2019. In addition, firm value rises by about 1% in the estimated model. For reference, the literature measures firm value losses of 2% due to biases in managerial beliefs (Barrero, 2020) and 3% due to CEO turnover frictions (Taylor, 2010). Welfare in the estimated model is 0.6% higher in consumption equivalent terms, a magnitude comparable to the costs

[18]Non-JIT producers also reduce their inventory targets due to the rise in the price of orders.

of business cycles (Krusell, Mukoyama, Sahin, & Smith, 2009) and the costs of managerial short-termism (Terry, 2017).

Fixed order costs are a source of misallocation in the model. Ideally firms would like to hold no material inventories, instead placing orders and fully utilizing them when producing every period. In an effort to minimize the number of times the fixed order costs are incurred, producers hold non-zero inventories. For this reason, the estimated JIT adoption model implies a reduction in misallocation. With more adoption, a greater number of producers operate subject to lower order costs. At the aggregate level, this implies that resources are reallocated to high marginal product producers. In essence, firms place more frequent orders and therefore have the flexibility to better align their material usage with their realized micro productivity realizations. The estimated model implies that JIT adoption raises measured TFP by approximately 0.2%.

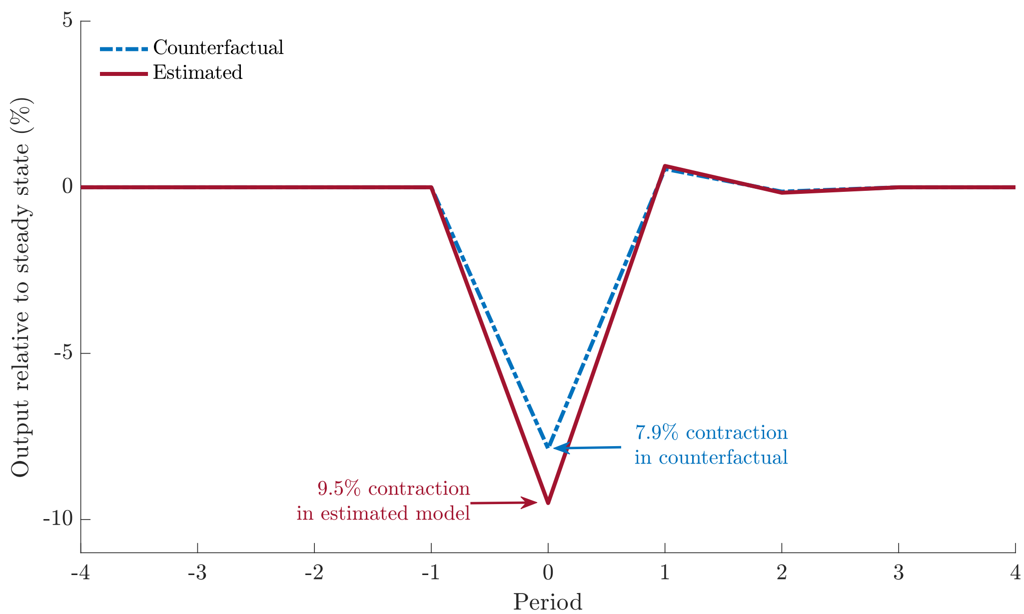
The reduction in misallocation manifests itself in lower firm volatility, consistent with Figure 1·5. Panel B of Table 1.9 reports results from an unconditional simulation of firms at the steady state. The variance in marginal product of materials falls in the estimated model relative to the counterfactual model. Furthermore, sales volatility in the estimated model falls by 9% relative to the counterfactual model. For reference, the same reduction in firm sales volatility would be achieved in a model without any JIT and a roughly 45% reduction in order costs.

1.6.2 Effects of an Unanticipated Disaster

I next show that despite enjoying higher profits and smoother firm-level outcomes, an economy populated by lean producers is more vulnerable to an unexpected disaster. To do this, I introduce aggregate productivity into the production function for intermediate goods.

$$O = AK^{\alpha}L^{1-\alpha}$$

Figure 1·6: Deeper Crisis with More Adoption

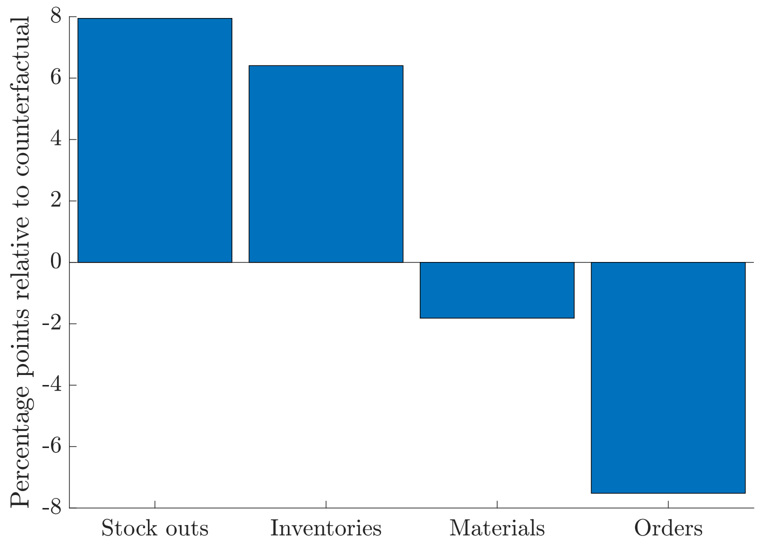


Notes: The figure plots the output response to a productivity shock that matches the 9.5% year-over-year decline in real GDP in 2020Q2.

Whereas in the steady state $A = 1$, in a disaster episode A unexpectedly falls below one. I shock this parameter so as to match the 9.5% drop in year-over-year real GDP in the second quarter of 2020. I consider a disaster duration such that the economy returns to its steady state after three years.[19] Figure 1·6 displays the endogenous output response to this unexpected disaster. In addition, Figure 1·7 reports the key differences in endogenous responses between the two models over the full disaster path. Consistent with the burgeoning literature that studies COVID-19, I model the disaster as an unanticipated event (Arellano et al., 2020; Espino et al., 2020). In Appendix A.1, I show that my quantitative results are robust to allowing

[19] As rare disasters are inherently infrequent, the number of such events is limited in short samples. Here I follow Barro and Ursua (2008) who report a mean duration of 3.5 years from a cross-country panel of disasters. I consider alternate durations in Appendix A.1 with little impact on the qualitative conclusions.

Figure 1·7: More Stock Outs and Inventory Hoarding



Notes: The figure plots the responses of key endogenous variables over the course of the simulated disaster in the estimated economy relative to the counterfactual economy (in percentage points).

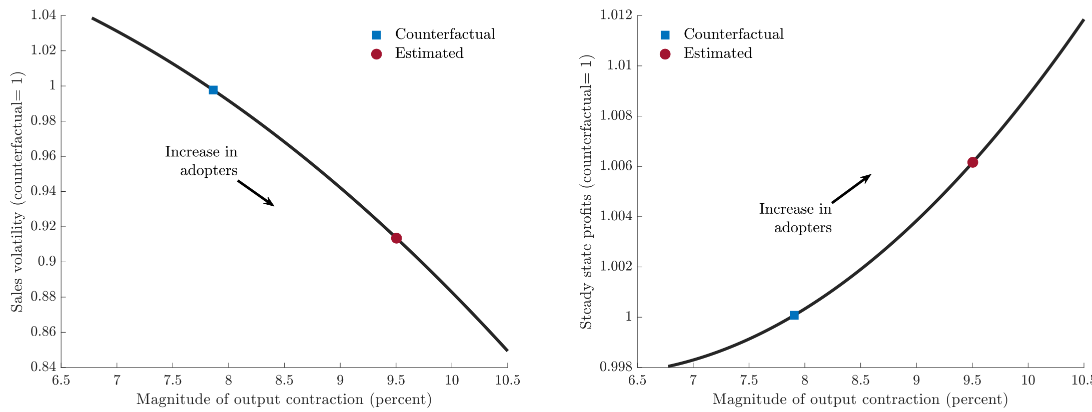
for anticipation.

Overall, the estimated model sees a 1.6 percentage point sharper output contraction on impact than the counterfactual model. During an unexpected disaster, the shadow value of inventories rises leading firms to reduce order sizes. Since firms in the JIT economy are leaner on average, the JIT economy experiences a 7.5 percentage point sharper spike in stock outs. At the same time, despite a reduction in inventory stocks in both economies, firms in the JIT economy place fewer orders and draw inventories down more slowly. As a result, these firms necessarily make use of fewer material inputs in production and sales therefore contract more sharply in the JIT model. In particular material input use falls 1.8 percentage points more over the course of the disaster in the estimated JIT economy.

A seemingly minor change to JIT adoption incentives across the two models delivers a substantial difference in the extent to which the economy falls into crisis amid a

Figure 1·8: Micro Stability vs. Macro Vulnerability

(a) Sales Volatility **(b)** Firm Profits



Notes: Panels (a) and (b) plot the magnitude of GDP contraction conditional on a 3-year disaster on the horizontal axis. Panel (a) plots an index of equilibrium firm sales volatility on the vertical axis (relative to the counterfactual economy defined above) while Panel (b) plots an index of equilibrium firm profits on the vertical axis (relative to the counterfactual economy defined above). Each point represents a different counterfactual economy, with the estimated economy denoted by the red circle and the counterfactual described in the text denoted by the blue square. The sunk cost parameters (c_s, c_f) are varied in order to generate the other counterfactual economies. The curve is a polynomial interpolation of the set of counterfactuals.

disaster. The excess output loss amounts to approximately $300 billion, a figure comparable to the funds allocated for direct cash payments to households following the passage of the CARES Act.[20] Lean inventory management therefore plays a meaningful role in determining the vulnerability of the economy to unanticipated shocks. During large unexpected disasters, inventories can in fact serve as a stabilizing force.

1.6.3 The JIT Tradeoff

Having examined the effects of lean inventory management on the economy in normal times as well as amid a COVID-19-magnitude disaster, I next trace out frontiers that illustrate the micro-macro tradeoff associated with JIT for a range of counterfactual economies. These frontiers point to an economically important tradeoff and imply that inventory management is an important source of aggregate fluctuations amid

[20] Coronavirus Aid, Relief, and Economic Security Act, H.R. 748, 116th Congress (2020).

large unexpected shocks.

Panel (a) of Figure 1·8 plots the tradeoff between firm sales volatility and the magnitude of the GDP contraction on impact for several counterfactual economies, each differing in steady state mass of JIT firms. Each point on the curve refers to a specific parameterized economy, traced out by varying the adoption costs, c_s and c_f. The red circle denotes the estimated economy and the blue square denotes the counterfactual. The panel shows that micro volatility falls with more JIT adoption, at the risk of elevated vulnerability to a shock. A 9% reduction in firm volatility comes at the cost of a 1.6 percentage point sharper GDP contraction.

Panel (b) of Figure 1·8 plots a similar tradeoff, this time comparing steady state firm profits with the magnitude of the GDP contraction. The curve slopes upward, as steady state firm profits are increasing in adoption while the extent to which the economy is vulnerable to an unanticipated shock also rises. A 0.6% increase in firm profits comes at the cost of a 1.6 percentage point sharper GDP contraction. For reference, the same increase in firm profits would arise in a model with no JIT and a 55% reduction in economy-wide order costs. The ranges of this frontier imply an economically large tradeoff between measures of micro stability or profitability and macro vulnerability.

1.7 Conclusion

At the firm level, it pays to be lean. I provide empirical evidence of the benefits of JIT inventory management among publicly traded manufacturers. Upon adopting JIT, firms hold fewer inventories, and observe higher sales and smoother outcomes. JIT firms, however, appear to be more cyclical and susceptible to disaster episodes. In a rich model of JIT production, the most productive firms adopt JIT which raises long-run firm value by 1% and reduces firm volatility by 9%. At the same time, JIT elevates firm vulnerability due to low inventory buffers. Amid an unexpected disaster, output in the estimated JIT economy contracts 1.6 percentage points more than a counterfactual economy with less JIT. Adoption, therefore, gives rise to an important and previously undocumented tradeoff which implies that inventories can indeed matter for aggregate fluctuations. Economists interested in understanding fluctuations within firms, and the responsiveness of the economy to aggregate shocks, should play close attention to both inventories and management practices.

Chapter 2

Time-Varying Volatility, Underreaction, and Overreaction

2.1 Introduction

Professional forecasts exhibit error predictability. Specifically, the covariance between ex-post errors and ex-ante revisions is non-zero and can run in either direction. A negative covariance is interpreted as an overreaction whereas a positive covariance is interpreted as an underreaction. At the same time, macroeconomic and financial time series have been found to exhibit complex dynamics such as stochastic volatility, structural breaks, and regime switching. Whereas existing models of belief formation do not generally accommodate simultaneous over- and underreactions, I show that these patterns can arise in an otherwise standard noisy information setting that incorporates unobserved time-varying volatility and heterogeneous forecasting techniques.

Error predictability in the model arises due to updating mistakes committed by forecasters. With unobserved volatility, the optimal weight to place on new information is not exactly known. In addition, there are costs associated with devising quantitative predictions. For instance, producing a forecast requires (computing) time and cognitive effort. To the extent that macroeconomic dynamics vary in their complexity, it stands to reason that forecasters tailor their models to each time series. Furthermore, subject to these costs, forecasters may select simpler misspecified models. Therefore with time-varying volatility and heterogeneous forecasting models, revisions can hold predictive power over errors, and the nature of this relationship can be variable-specific. In spite of this, my framework is compatible with rationality in the sense that reported forecasts are optimal outcomes.

Survey data has traditionally been used to test theories of expectation formation. In this paper, I make use of the Survey of Professional Forecasters (SPF) which provides a panel of multi-horizon forecasts across several macroeconomic variables.[1] In the data, over- and underreactions arise along different dimensions. First, across levels of aggregation, consensus forecasts broadly exhibit underreactions while forecaster-level predictions tend to imply overreactions. Second, across variables and forecasters, overreactions appear for some variables and underreactions for others. Both of these facts have been previously documented in the literature. This paper offers a third empirical fact: across variable, within forecaster, the same respondent appears to over- and underreact to distinct macroeconomic variables.

The presence of simultaneous over- and underreactions prompts several fundamental questions about belief formation. Are professional forecasters, presumably the most informed private agents in the economy, rational? Alternatively, do behavioral biases govern the manner in which expectations are formed? As policymakers increasingly pursue expectations-based policies such as forward guidance, taking a step toward reconciling theories of expectations formation with the data is of first-order importance.

Against this backdrop, I develop a noisy information model with unobserved time-varying volatility. Rather than obtaining an exact solution to the optimal inference problem, forecasters must approximate the posterior distribution. They may choose from a finite set of approximation methods. The available methods vary in complexity, and adopting a given method is subject to a cost that is increasing in forecasting

[1] Examples other than the SPF include the Livingston survey, the Michigan Survey of Consumers, the NY Fed Survey of Consumer Expectations, Blue Chip forecasts, the ECB Survey of Professional Forecasters, and the daily Focus Survey from the Central Bank of Brazil, among others.

model sophistication. Forecasters generate a prediction that minimizes the sum of their mean squared errors and model adoption costs.

Importantly, some forecasters adopt suboptimal forecasting models to predict variables thereby generating error predictability. I consider a stylized version of this unobserved volatility model in which forecasters can select either a suboptimal Kalman filter or an asymptotically efficient particle filter, the former being less costly to adopt than the latter.[2] I find that the underlying signal-to-noise ratio governs the extent to which over- and underreactions arise. Intuitively, the optimal weight to place on new information is increasing in the time-varying signal-to-noise ratio. Predictions based on the suboptimal model, however, erroneously update new information in a constant fashion. As a result, forecasters will tend to underreact to variables for which the average signal-to-noise ratio is high, and will overreact to variables for which the average signal-to-noise ratio is low. Put another way, there are certain features inherent to a given macroeconomic time series that explain why forecasters appear to either over- or underreact to that particular variable.

After providing simulation results that confirm the above intuition, I examine these implications in the data by exploiting the cross-section of macroeconomic variables for which forecasters report predictions in the SPF. I then parameterize the stylized model and show that it can match a quantitatively relevant share of simultaneous over- and underreactions. Taken together, my findings demonstrate that time-varying volatility, coupled with costly forecast model adoption, can rationalize important features of survey expectations data.

In their seminal paper, Coibion and Gorodnichenko (2015), henceforth CG, make sense of forecast error inefficiency while preserving the assumption of rationality. Using consensus-level data, CG show that projecting ex-post forecast errors on ex-ante

[2] The particle filter is an asymptotically efficient approximation method. I provide further details on the filter in Appendix A.2.

forecast revisions delivers an estimate of information rigidity. More recently, many studies have used forecaster-level data to test for rationality.[3] In doing so, much of this literature preserves the linearity assumption made in CG, and ultimately rejects rational expectations even under imperfect information. To make sense error predictability at the forecaster-level while also matching the CG finding of underreactions at the aggregate level, several theories of non-rational expectations have been proposed.[4] My paper relates to this strand of the literature in many ways, and provides an alternate interpretation of the errors-on-revisions coefficient.

In a recent contribution, Kohlhas and Walther (2020) also examines simultaneous over- and underreactions. The authors are able to explain overreaction to news coupled with underreaction on average with a model of asymmetric attention. Although I ground over- and underreactions from a slightly different empirical perspective, I view my paper as complementary to theirs. Whereas my model is based on heterogeneity in the underlying volatility across state variables, Kohlhas and Walther (2020) present a model of costly attention which delivers distinct signal precisions for different components of the state. In both cases, however, the underlying signal-to-noise ratio is the relevant object that varies across variables or components.[5] This paper is also related to Gabaix (n.d.) which proposes a model in which agents over- and underreact due to misperceived persistence of the data generating process. The focus of my model, however, is in how forecasters assess volatility. Nonetheless, my model can in general speak to sources of misperceived persistence, for example, unobserved structural breaks.

[3] Examples include Bordalo, Gennaioli, Ma, and Shleifer (2020), Fuhrer (2018), Dovern, Fritsche, Loungani, and Tamirisa (2015), Andrade and Bihan (2013), Broer and Kohlhas (2019), Bürgi (2016).

[4] For instance Bordalo et al. (2020) rule out rationality in favor of diagnostic expectations. Other studies such as Fuster, Hebert, and Laibson (2012a) argue in favor of models featuring misperception at long horizons. Daniel, Hirshleifer, and Subrahmanyam (1998) argues for a model of overconfidence while Broer and Kohlhas (2019) present a model of relative overconfidence.

[5] Relatedly, Broer and Kohlhas (2019) present a model of over- and underreactions. The focus in this paper is to match simultaneous over- and underreactions to endogenous public signals.

Moreover, the unobserved volatility noisy information model in this paper is in the spirit of Branch (2004), Evans and Ramey (1992) and Brock and Hommes (1997) who define adaptively rational equilibrium dynamics (ARED). Branch (2004) was the first to introduce this concept to expectations formation. My paper builds on his important insights in key ways. First, I present more complex dynamics for the state variable. Introducing nonlinearities, such as stochastic volatility, provides an even stronger justification for the use of different predictor functions. Second, I explicitly model heterogenous expectations through private information whereas in Branch (2004) predictions are assumed to be homogeneous among all who adopt a specific predictor function.[6] Taken together, my model is able to reproduce the empirical facts relating to simultaneous over- and underreactions across level of aggregation, by variable across forecasters, and by variable within forecaster.

While a discussion of nonlinearities has generally been absent in the survey expectations literature, the finance literature has previously tied nonlinear dynamics to error predictability. For instance, Lewis (1989) considers error predictability concerning dollar forecasts in the context of a structural break. Veronesi (2015) finds that over- and underreactions arise in a regime switching model of asset pricing. More recently, Lansing, LeRoy, and Ma (2020) attribute the predictability of excess returns to either volatility or deviations from rationality. To this end, my paper also relates to the literature on volatility in macroeconomics.[7]

Finally, due to the assumptions imposed on the state dynamics, this paper relates to the literature on nonlinear filtering. Several approximation methods have been devised in order to deal with nonlinearities in the evolution of a state variable. These methods include generalizations to Kalman filtering as well as importance sampling

[6] Heterogeneity in this model comes from idiosyncratic "trembles" in the reported prediction.

[7] See for instance, Justiniano and Primiceri (2008), Kim and Nelson (1999), McConnell and Perez-Quiros (2000), and Stock and Watson (2007)

algorithms, among others.[8] A strand of this literature has formalized some basic efficiency properties of particle filtering.[9]

The rest of the paper is organized as follows. Section 2.2 presents previously documented facts about error predictability at the forecaster and consensus levels, as well as a novel fact pertaining to simultaneous over- and underreactions. Section 2.3 presents the noisy information model subject to unobserved time-varying volatility. Section 2.4 introduces a stylized version of the model and provides simulation results. Section 2.5 documents empirical evidence consistent with the model. Section 2.6 parameterizes the stylized model to show that it can generate within forecaster over- and underreactions. Finally, Section 2.7 concludes.

2.2 Evidence from Survey Data

The SPF is a quarterly survey provided by the Federal Reserve Bank of Philadelphia. The survey began in 1968Q4 and provides forecasts from several forecasters across a number of macroeconomic variables over many horizons, h. The variables of interest in this paper are the forecast error and the forecast revision. To construct forecast errors from forecaster i about variable x,

$$FE^i_{t+h,t} = x_{t+h} - x^i_{t+h|t},$$

I take the difference between the realized real-time value for x at $t+h$ and the forecaster's h-step ahead prediction generated at time t. To compute forecast revisions, I exploit the term structure of forecasts generated by the survey respondents

$$FR^i_{t,t-1} = x^i_{t+h|t} - x^i_{t+h|t-1}.$$

[8]Julier and Uhlmann (2004) develop a Kalman filter for nonlinear settings while Doucet and Johansen (2009) discuss particle filtering methods.

[9]See Crisan and Doucet (2002) and Hu, Schon, and Ljung (2011)

Table 2.1: Pooled OLS Forecast Error Predictability Regressions

	Nowcast		One-Quarter Ahead		Two-Quarters Ahead	
	β_1	α_1	β_1	α_1	β_1	α_1
Estimate	-0.317***	0.569***	-0.231**	1.011***	-0.344***	0.565**
	(0.050)	(0.128)	(0.067)	(0.201)	(0.058)	(0.272)
Observations	65,070	2,323	54,067	2,309	52,220	2,295

Notes: The table reports the estimated coefficients of forecast error predictability at the current, one-, and two-quarter ahead horizons. Across all horizons, column (1), refers to the forecaster-level errors-on-revisions regression. Column (2) refers to the consensus-level errors-on-revisions regression. Standard errors for forecaster-level regressions are as in Driscoll and Kraay (1998), while Newey-West standard errors are used for aggregate-level specifications. Data used for estimation come from SPF. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

This requires making use of the h-step ahead forecasts formulated in periods t and $t-1$. In other words, I consider the fixed horizon, h and take the difference between two adjacent forecasts.

In the data, simultaneous over- and underreactions arise along different dimensions: across level of aggregation, across SPF variables pooled over forecasters, and across SPF variables within forecaster. CG present the following testable implication at the consensus-level which holds for an arbitrary horizon:

$$FE_{t+h,t} = \alpha_0 + \alpha_1 FR_{t+h,t} + \epsilon_t. \tag{2.1}$$

CG find that in the data, $\alpha_1 > 0$ for most variables which indicates that consensus forecasts underreact to new information. More recently, Bordalo et al. (2020) estimate the same regression at the forecaster-level:

$$FE^i_{t+h,t} = \beta_0 + \beta_1 FR^i_{t+h,t} + \varepsilon^i_t, \tag{2.2}$$

and find that $\beta_1 < 0$ for most macroeconomic series. The interpretation is that forecasters overreact to new information.

Table 2.2: Pooled OLS Regressions at $h = 0$, by Variable

Variable	Mnemonic	β_1	α_1
Consumer price inflation	CPI	-0.085	0.868***
Employment	EMP	-0.123	0.564***
Housing starts	HOUSING	0.063	0.359***
Industrial production	IP	-0.147*	0.513***
Nominal GDP	NGDP	-0.310***	0.421**
GDP Deflator	PGDP	-0.363***	0.350**
Real consumption	RCONSUM	-0.401***	0.098
Real federal government spending	RFEDGOV	-0.483***	0.377
Real GDP	RGDP	-0.264***	0.350**
Real nonresidential investment	RNRESIN	-0.499**	0.362
Real residential investment	RRESINV	-0.234***	0.925***
Real state/local government spending	RSLGOV	-0.660***	-0.381
3-month Treasury bill	TBILL	0.010	0.178***
10-year Treasury bond	TBOND	0.020	0.154***
Unemployment rate	UNEMP	0.082**	0.247***

Notes: The table reports the OLS coefficients from errors-on-revisions regressions across 15 macroeconomic variables reported in the Survey of Professional Forecaters. Column (3) reports the coefficient in front of the revision at the forecaster-level while column (4) reports the analogous coefficient using consensus-level data. The errors and revisions are for current period forecasts ($h = 0$). All forecasts refer to growth rates with the exception of consumer price inflation (CPI), 3-month treasury bill (TBILL), 10-year bond (TBOND), and unemployment rate (UNEMP). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

Table 2.1 reports estimates of β_1 and α_1, using data from the SPF. The estimates are obtained via OLS regressions, pooling across both forecasters and macroeconomic variables. Estimates are reported for three different horizons. Across all horizons considered, it is clear that overreactions dominate at the individual-level, while underreactions arise at the aggregate-level.

However even at the forecaster-level there is evidence of simultaneous over- and underreactions across macroeconomic variables. Table 2.2 reports variable-by-variable results for nowcasts ($h = 0$) at the forecaster- and consensus-levels. The results point to individual overreactions for most variables but underreactions for some variables such as the unemployment rate.

These findings are neither driven by entry and exit among SPF forecasters, nor different forecasters systematically reporting predictions for select macroeconomic variables. Instead, the same respondent simultaneously over- and underreacts to distinct macroeconomic variables. To show this, I estimate regression (2.2) for each forecaster i forecasting a specific variable j. This delivers an $N_i \times N_j$ matrix of estimates $\widehat{\beta}_{1,ij}$. I keep only those estimates that are significant at the 5% level. I then fix a pair of SPF variables j and k, and compute the number of forecasters such that $\widehat{\beta}_{1,ij} < 0$ and $\widehat{\beta}_{1,ik} > 0$, and normalize by the number of total forecasters reporting predictions about variables j and k. More formally, I estimate a matrix P whose elements are p_{jk} with

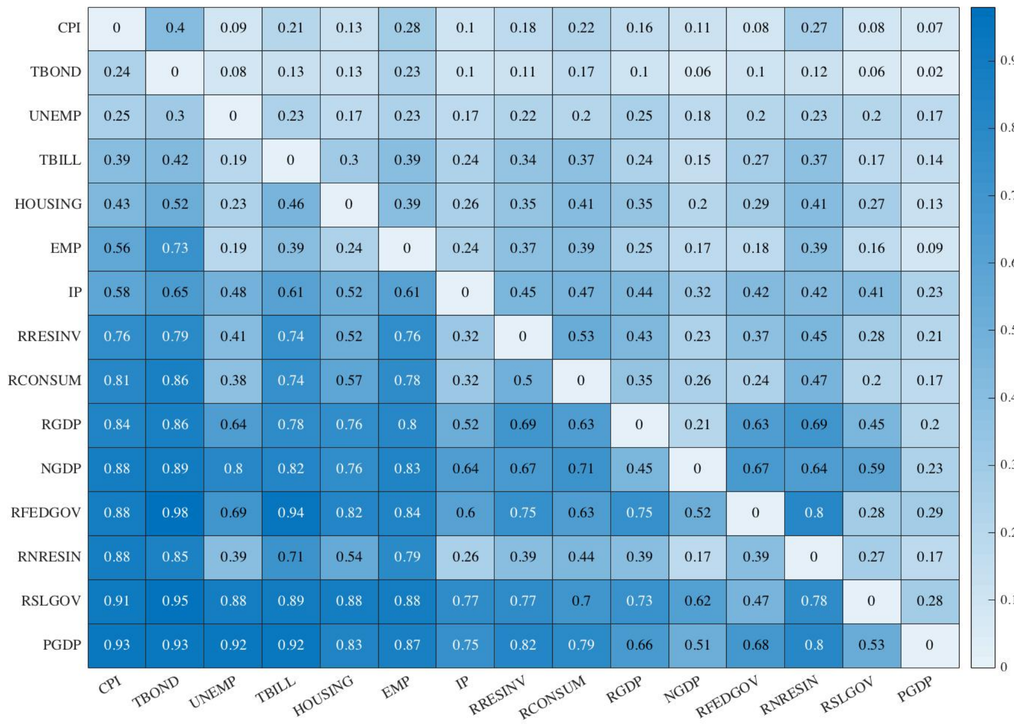
$$p_{jk} = \frac{\sum_i \mathbb{1}\left(\widehat{\beta}_{1,ij} < 0 \text{ and } \widehat{\beta}_{1,ik} > 0\right)}{\min\{N_j, N_k\}}$$

where N_x denotes the number of forecasters providing predictions of variable x and $\mathbb{1}(\cdot)$ is the indicator function. The elements of matrix P, therefore, denote the share of forecasters who simultaneously overreact to the row variable and underreact to the column variable. When p_{jk} is close to one, this means that nearly all forecasters overreact to variable j and underreact to variable k. On the other hand, when p_{jk} is close to zero, then almost no forecaster overreacts to variable j while also underreacting to variable k.

Figure 2·1 reports the results from this exercise. The heatmap verifies that a given forecaster tends to overreact to some variables and underreact to others. For instance, 84% of professional forecasters exhibit overreactions when forecasting growth in real gross domestic product (RGDP) while simultaneously underreacting to information regarding inflation based on the consumer price index (CPI).

In order to understand how individuals formulate these expectations, a theory of expectations formation must take into account that a single agent may overreact and

Figure 2·1: Frequency of Over- and Underreaction



Notes: The heatmap displays the share of forecasters who overreact to the row variable and simultaneously underreact to the column variable.

underreact to different variables. I propose a model of time-varying volatility and heterogeneity in forecasting models in order to account for this fact.

2.3 Model

I begin with a simple regime switching example to highlight the intuition. I then proceed to develop a generalized noisy information rational expectations model with unobserved time-varying volatility and heterogeneous forecasting methods. In the next section, I narrow my focus to a stylized version of this model which I later parameterize.

2.3.1 A Simple Model

Suppose that the state is described as follows:

$$s_t = \overline{s} + w_t \quad w_t \sim \mathcal{N}(0, \sigma_t^2),$$

where

$$\sigma_t^2 = \begin{cases} \sigma_L^2 & \text{with probability } q \\ \sigma_H^2 & \text{with probability } 1 - q. \end{cases}$$

The forecaster cannot directly observe the latent state s_t or its volatility σ_t^2. More specifically, the probability q is unknown. Instead, the forecaster receives a signal each period that is contaminated with noise:

$$y_t = s_t + v_t \quad v_t \sim \mathcal{N}(0, \sigma_v^2).$$

The optimal weight to place on new information (the Kalman gain) is:

$$\kappa_t = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_v^2}.$$

Without full knowledge of the probability q, the optimal weighting of new information is unknown. Suppose that the forecaster finds it too time consuming or otherwise prohibitive to ascertain q. Instead, the forecaster simply assess the state variance to be $\sigma^2 = \frac{1}{2}(\sigma_L^2 + \sigma_H^2)$. The associated Kalman gain is

$$\kappa = \frac{\sigma^2}{\sigma^2 + \sigma_v^2}.$$

The resulting weighting error can be expressed as:

$$\kappa_t - \kappa = \frac{[\sigma_t^2 - \frac{1}{2}(\sigma_L^2 + \sigma_H^2)]\sigma_v^2}{[\frac{1}{2}(\sigma_L^2 + \sigma_H^2) + \sigma_v^2](\sigma_t^2 + \sigma_v^2)}.$$

When $\sigma_t^2 = \sigma_L^2$, the weighting error is negative meaning that the forecaster puts undue weight on his signal thereby overreacting. On the other hand, when $\sigma_t^2 = \sigma_H^2$, the weighting error is positive and the forecaster underreacts to new information.

Based on this simple example, the magnitude of the overreaction depends importantly on the signal-to-noise ratio. Noisier environments deliver more negative β_1 coefficients. On the other hand, underreactions arise when q is closer to zero (i.e., the underlying signal-to-noise ratio is high). In this case, as the forecaster generates predictions, he believes the state to be less variable than it truly is thereby placing less weight on news. This mutes the effects of signal noise and creates more inertia in expectation formation than is optimal. The result is a more positive β_1 coefficient.

The weighting errors in my model stem from: (i) unobserved volatility and (ii) incentives to adopt parsimonious approximations of the volatility. With these two assumptions, over- and underreactions can arise depending on the underlying state dynamics. Assumption (ii) is crucial because if forecasters could easily observe q, then there would be no need to approximate the variance of the state. In this case, error orthogonality would hold despite the regime switching nature of the variance. However, if adopting different forecasting models comes at a cost (be it cognitive, timing, or otherwise), forecasters may find it optimal to make use of such approximations.

2.3.2 A Model of Unobserved Time-Varying Volatility

Having illustrated the basic intuition that delivers simultaneous over- and underreactions, I now turn to presenting the general model. Nonlinearities such as time-varying volatility in the underlying state complicates the forecaster's problem as he must now formulate expectations about levels and the volatilities. Supposing that there are n latent state variables and m exogenous signals, the state and observations equations are:

$$\mathbf{\bar{s}_t} = F(\mathbf{\bar{s}_{t-1}}, \mathbf{\bar{w}_t}) \qquad (2.3)$$
$$\mathbf{z_t^i} = \mathbf{C}\mathbf{\bar{s}_t} + \mathbf{D}\mathbf{v_t^i},$$

where $\mathbf{\bar{s}_t}$ is an $n \times 1$ vector, $\mathbf{z_t}$ is $m \times 1$, $\mathbf{C}$ is $m \times n$, $\mathbf{D}$ is $m \times m$ and $\mathbf{v_t^i}$ is $m \times 1$. There are no other restrictions placed on the model. In particular, $\mathbf{s_t}$ can be a vector of many different state variables, or lags of itself. Furthermore, $\mathbf{z_t^i}$ can include an arbitrary finite number of observed signals. The noise vector $\mathbf{v_t^i}$ can include private or public noise.[10]

In a linear model, $\mathbf{s_t} = \mathbf{A}\mathbf{s_{t-1}} + \mathbf{B}\mathbf{w_t}$. The crucial difference between a linear model and this one is the unobserved time-varying covariance matrix $\mathbf{B}_t$ which implies an expanded state space, $\mathbf{\bar{s}_t} = \begin{pmatrix}\mathbf{s_t} & diag(\mathbf{B_t})\end{pmatrix}'$. As a result, the error now enters *multiplicatively* into the state. This nonlinearity is modeled by the function $F(\cdot)$ which governs the evolution of the state. While the state now exhibits stochastic volatility, the shocks remain normal, and the signal structure is unchanged. Hence, the measurement equation remains linear.[11] [12]

Whereas Kalman filtering delivers an exact optimal solution in a linear Gaussian environment, it is no longer optimal in this context. The reason for this is that the Kalman filter requires one to evaluate the expected value of $\mathbf{\bar{s}_t}$ conditional on the history of signals $\mathcal{Z}_t^i = \{\mathbf{z_1^i}, \ldots, \mathbf{z_t^i}\}$. This is made intractable due to the lack of knowledge about the underlying conditional distribution. To see this more clearly, consider the scalar case where the state is s_t and there is only a private signal available to the forecaster, z_t^i. The observation equation can be expressed as a conditional

[10] I index this vector by i in general to allow for forecaster-specific signals.

[11] This could be generalized to a nonlinear measurement as well. I abstract away from this for simplicity.

[12] While I consider stochastic volatility, any nonlinearity can deliver the results presented in the paper. In particular, this model can also speak to unobserved changes in the persistence of macroeconomic time series.

likelihood, $p(z_t^i|s_t)$ and the state evolution as $p(s_{t+1}|s_t)$. The optimal filter computes $p(s_t|\mathcal{Z}_t^i)$ from a predict-update procedure

$$p(s_t|\mathcal{Z}_{t-1}^i) = \int p(s_t|s_{t-1})p(s_{t-1}|\mathcal{Z}_{t-1}^i)ds_{t-1} \quad \text{(Predict)}$$

$$p(s_t|\mathcal{Z}_t^i) = \frac{p(z_t^i|s_t)p(s_t|\mathcal{Z}_{t-1}^i)}{p(z_t^i|\mathcal{Z}_{t-1}^i)} \quad \text{(Update)}$$

where $p(z_t|Z_{t-1}^i) = \int p(z_t^i|s_t)p(s_t|Z_{t-1}^i)ds_t$.

In a linear Gaussian environment, this can be exactly computed via the Kalman recursions.[13] In a nonlinear setting, however, computing $p(s_t|\mathcal{Z}_t^i)$ is not feasible as the density cannot be obtained analytically.

In light of this, forecasters must approximate the nonlinear state. I assume that this is done by selecting from a set of costly approximation functions, $A \in \mathcal{A}$. Forecasters first select an approximation function so as to obtain an estimate of the posterior density of the underlying state. Forecasters then report their predictions, the first moment of this approximated density. Hence, the forecaster's loss function can be defined as

$$\mathcal{L} = \min_{A \in \mathcal{A}} \left[\big(\mathbf{z^i_{t+h}} - \widehat{\mathbf{z}}^{\mathbf{i,A}}_{\mathbf{t+h|t}}\big)' \big(\mathbf{z^i_{t+h}} - \widehat{\mathbf{z}}^{\mathbf{i,A}}_{\mathbf{t+h|t}}\big) + c_A^i \right], \tag{2.4}$$

where the first term is the mean square error arising from individual i's forecast which makes use of approximation function A, and the second term denotes the cost associated with adopting approximation function A.[14] I assume that these forecaster-specific costs are drawn randomly $c_A^i \sim U(0, \bar{c}_A)$.[15] This cost embodies unobserved heterogeneity among forecasters that result in the use of different forecasting models.

[13]In particular, $p(s_t|\mathcal{Z}_t^i) \sim \mathcal{N}(s_{t|t}^i, \Psi_{t|t}^i)$, where $s_{t|t}^i$ is the expected value of the posterior density and $\Psi_{t|t}^i$ is the variance.

[14]Forecasters have knowledge of the mean square error associated with each A.

[15]One could alternatively assume heterogeneous signal precision. Models featuring heterogeneous signal-to-noise ratios have been proposed in the literature, particularly to explain forecast disagreement.

After applying their approximations of the state, forecasters generate a prediction and an update according to the new information received. Since agents are formulating a forecast subject to an approximation of the state, I call these approximate predictions. An approximate prediction is defined as

$$\widehat{\mathbf{\bar{s}}}^{\mathbf{i}}_{\mathbf{t|t}} = \int \mathbf{\bar{s}_t}\, \widehat{p}(\mathbf{\bar{s}_t}|\mathcal{Z}^i_t)\, d\mathbf{\bar{s}_t}. \tag{2.5}$$

In essence, the forecaster predicts the current state according to the approximated density $\widehat{p}(\mathbf{\bar{s}_t}|\mathcal{Z}^i_t)$. In a linear Gaussian setting, the density would be obtained exactly so that $\widehat{p}(\mathbf{\bar{s}_t}|\mathcal{Z}^i_t) = p(\mathbf{\bar{s}_t}|\mathcal{Z}^i_t)$ and errors would be orthogonal.

One can express the approximate prediction as a deviation from the optimal minimum mean square error forecast

$$\widehat{\mathbf{\bar{s}}}^{i}_{t|t} = \underbrace{\mathbb{E}(\mathbf{\bar{s}}_t|\mathcal{Z}^i_t)}_{\text{Optimal}} + \underbrace{\int \mathbf{\bar{s}}_t[\widehat{\mathbf{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \mathbf{p}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}})]d\mathbf{\bar{s}_t}}_{\text{Approximation error}}. \tag{2.6}$$

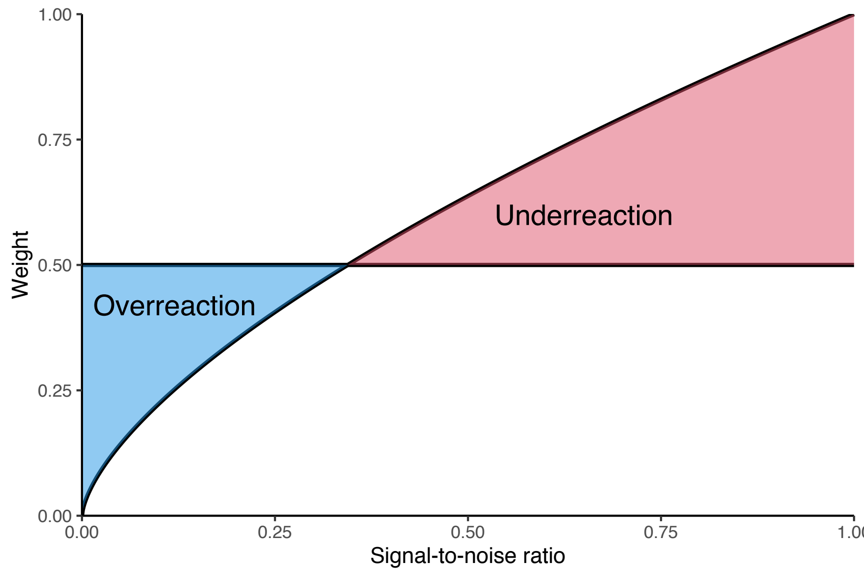
Whereas existing theories of expectation formation restrict the deviation from the optimal forecast to be either positive (overreactions) or negative (underreactions), the direction of the error here is unrestricted.

2.3.3 Scope for Over- and Underreaction

Error predictability is due to the presence of suboptimal models in the set $\mathcal{A}$. Importantly, some forecasters must select suboptimal approximations from this menu of models. A sufficient condition for the presence of over- and underreactions is that some approximation $A \in \mathcal{A}$ involves a time-invariant weight used to update new information which resides in the interval $(0, 1)$.

Intuitively, the optimal weight to place on new information is increasing in the signal-to-noise ratio. If a forecaster partially incorporates the most recent signal with a naive constant weight, then the forecaster will overweight new information when

Figure 2·2: Scope for Over- and Underreaction



Notes: The figure illustrates the optimal Kalman gain (upward sloping curve) and the suboptimal Kalman gain (horizontal line) as functions of the signal-to-noise ratio.

signals tend to be imprecise, and will underweight new information when signals tend to be more precise. Figure 2·2 illustrates this intuition.

The gaps between the curve and the horizontal line reflect weighting errors. These weighting errors are made each period among those who choose the suboptimal constant-weight forecasting model. The scope for underreaction rises with the signal-to-noise ratio.[16]

The constant weight assumption is made for the purpose of transparency. There are other forecasting techniques that could deliver misspecified time-varying weights which could deliver a similar intuition to the one displayed in Figure 2·2.

Aggregate error predictability arises because the consensus revision does not reside in any forecaster's information set. Appendix A.2 proves that $\alpha_1 \geq \beta_1$ which implies

[16] It is important to note that the optimal weight will vary along the curve according to the underlying stochastic volatility. Ultimately, the optimal weight is a complicated function of the volatility and persistence of the state as well as the noisiness of the signals, so the black curve is specific to the macroeconomic variable in question.

that underreaction can arise at the aggregate level. Aggregate underreactions arise in part due to aggregation, but also because not all forecasters adopt suboptimal models in this noisy information environment.

2.3.4 Relation to Some Theories of Expectation Formation

Amid the mounting evidence against full information rational expectations, several theories of expectations formation have been proposed in the literature. Here, I consider a few prominent theories and assess their ability to generate the empirical facts presented in Section 2.2.[17]

According to diagnostic expectations, forecasters over-weight new information according to a parameter $\theta > 0$ which ultimately governs the extent of overreaction. This parameter comes from the representativeness heuristic of Tversky and Kahneman (1974). The diagnostic nowcast is defined as:

$$x^{i,\theta}_{t|t} = x^i_{t|t} + \theta(x^i_{t|t} - x^i_{t|t-1}).$$

This theory makes use of a distorted Kalman filter called the diagnostic Kalman filter, which, as shown in Bordalo et al. (2020), is able to generate $\beta_1 < 0$ and $\alpha_1 > 0$. Because $\theta > 0$, however, this theory cannot accommodate underreactions. In other words, the sign restriction imposed on θ implies that diagnostic errors and revisions always covary negatively.

Models of overconfidence also distort the Kalman gain. The distortion stemming from overconfidence, however, is different. According to this theory, forecasters misperceive the precision of their own signals. Suppose that forecasters observe only one

[17] I do not include a formal discussion of sticky information or linear noisy information models as their shortcomings in this respect have already been documented in the literature. Appendix A.2 provides some detail on why these popular models are at odds with the data. In particular, they imply error orthogonality at the forecaster-level.

private signal:

$$z_t^i = s_t + v_t^i, \qquad v_t^i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_v^2),$$

but they perceive $\widetilde{\sigma}_v < \sigma_v$. Put another way, individual forecasters believe their own private signals to be more precise than they truly are. This results in an erroneous assessment of the noise in the system. Overconfident beliefs are recursive so that the distorted gain injects a bias to the update in each period. These beliefs are then projected forward only to be further distorted by the over-weighting of new information in the subsequent period. In other words, at an arbitrary point in time, forecasters exhibit both a non-zero ex-ante forecast error as well as a weighting error. Models of overconfidence can generate individual overreactions as well as aggregate underreactions, however, overconfidence is similarly unable to generate individual underreactions.

Strategic interaction models can also generate error predictability. For instance, strategic substitution can drive errors and revisions in opposite directions. This is because forecasters have a dual objective of not only minimizing their errors but also of distinguishing themselves from the average forecast. These models differ from the previous two in that strategic interaction models are rational. While this class of models can generate either overreaction or underreaction depending on the strategic motive assumed, it is unable to jointly deliver $\beta_1 > 0$ and $\beta_1 < 0$.

Models of noisy memory, first introduced in Azeredo da Silvera and Woodford (2019), can also generate overreactions. In a noisy memory model, forecasters do not have access to their full history of signals due to finite memory capacity. While models of rational inattention can explain individual underreactions, noisy memory may explain individual overreactions. Developing a hybrid rational inattention-noisy memory model could plausibly deliver simultaneous over- and underreactions.

Several other theories of expectations formation have been found to be inconsistent with the data. For instance, models of reputational concerns imply smoothing which can only generate underreaction. Moreover, asymmetric loss functions deliver counterfactually biased expectations, whereas the data show that professional forecasts are not unconditionally biased.

2.4 Stylized Model

To extract further insight as to how time-varying volatility can generate over- and underreactions, I consider next a stylized model of noisy information and unobserved volatility. I provide simulation results that describe the source of over- and underreactions. In the subsequent section I document cross-sectional evidence consistent with this mechanism.

2.4.1 Set Up

For simplicity, I suppose that forecasters can choose between two models: a Kalman filter (KF) and a particle filter (PF). Forecasters utilizing KF ignore the stochastic volatility and assess only the unconditional volatility of the state when formulating predictions. By ignoring time-varying volatility, forecasts based on the Kalman filter are suboptimal and generate error predictability among forecasters.

To reiterate, the Kalman gain is increasing in the signal-to-noise ratio, as illustrated in Figure 2·2. In a world with stochastic volatility, the optimal weight placed on new information varies over time. If the nature of the volatility were known in each period, then forecasters could update their predictions efficiently according to the weights traced out by the curve. If one were to ignore the time-varying volatility and filter with only the unconditional variance of the state, then he would update according to the constant Kalman gain akin to the horizontal line in the figure.

The underlying state in the stylized model is described as follows:

$$s_t = \rho s_{t-1} + e^{h_t/2} w_t, \qquad w_t \sim \mathcal{N}(0,1) \tag{2.7}$$
$$h_t = \phi_0 + \phi_1 h_{t-1} + \sigma_\eta \eta_t, \quad \eta_t \sim \mathcal{N}(0,1).$$

Furthermore, forecasters observe a contemporaneous private signal as well as a lagged public signal[18]

$$y_t^i = s_t + \sigma_v v_t^i, \qquad v_t^i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1) \tag{2.8}$$
$$x_{t-1} = s_{t-1} + \sigma_e e_{t-1}, \quad e_{t-1} \sim \mathcal{N}(0,1).$$

In addition to being unable to observe s_t, forecasters are also unable to observe h_t. As such, the innovation w_t now enters into the state multiplicatively. Moreover, let $\mathbf{z}_t^i = \begin{pmatrix} y_t^i & x_{t-1} \end{pmatrix}'$ denote the vector of signals observed by forecasters.

Each period consists of two stages. In the first stage, a forecaster observes $\mathbf{z_t^i}$ and selects an approximation function. The use of these models comes at a random cost, $c^i \sim U(0, \bar{c})$ such that the PF cost distribution has a higher upper bound relative to KF. Then in the second stage, given the predictor function and the history of signals $\mathcal{Z}_t^i$, the forecaster reports a prediction of the public signal $\widehat{x}_{t|t}^i$ which is the macroeconomic variable in question.[19][20] For simplicity, I normalize $\bar{c}_{KF} = 0$. With this in mind, forecasters minimize the following loss function:

$$\mathcal{L} = \min \Big[MSE_{KF}^i, MSE_{PF}^i + c_{PF}^i \Big], \quad c_{PF}^i \sim U(0, \bar{c}_{PF}).$$

[18] One can alternatively envision that forecasters observe a macroeconomic variable with a transitory (e_t) and persistent (s_t) component. The persistent component is what is relevant for forecasting the target variable, though it is unobserved.

[19] The public signal in the model, x, maps to the SPF variable to be forecasted. The latent state, s_t, on the other hand, is unobserved to the forecaster and the econometrician.

[20] At t, forecasters make use of their full history of signals in order to formulate a state estimate, $\widehat{\mathbf{s}}_{\mathbf{t|t}}^{\mathbf{i}}$. The first element of the forecasted state vector is $\widehat{s}_{t|t}^i$. Based on the assumption that $x_t = s_t + e_t$, it follows that $\widehat{x}_{t|t}^i = \widehat{s}_{t|t}^i$.

A forecaster will choose to make use of the more sophisticated PF if and only if

$$MSE^i_{KF} - MSE^i_{PF} \geq c^i_{PF}. \tag{2.9}$$

The lefthand side of the inequality reflects the benefit to adopting the PF, which manifests itself in a lower mean square error, whereas the righthand side denotes the relative cost to adopting the PF. The adoption cost embodies unobserved heterogeneity in model adoption. As previously described, these reduced form costs can reflect heterogeneous time constraints among professional forecasters, different levels of training or experience in forecasting techniques, institution-specific frictions that make it more difficult to adopt a particular forecasting model, etc.

Fundamentally, this is a noisy information environment in which forecasters infer the state subject to private and public signals. Therefore, as in the simple model described in the previous section, the sign of the covariance between errors and revisions depends on the underlying signal-to-noise ratio. As the signal-to-noise ratio falls, forecast revisions are increasingly driven by the noise in the system. In this case, it is as if forecasters report their predictions with measurement error since an upward revision in the reported forecast will mechanically result in a more negative forecast error. On the other hand, when the signal-to-noise ratio is high, then fluctuations in the underlying state drive the forecast revisions. As a result, an upward revision delivers a more positive forecast error.

2.4.2 Simulation Results

The simulation results reported in Table 2.3 confirm that the model is able to qualitatively explain over- and underreactions across level of aggregation and variable. Differences across level of aggregation are seen through differences in simulated individual errors-on-revisions coefficient (β_1) and its consensus analog (α_1). Moreover,

Table 2.3: Signal-to-Noise Ratio and Implied OLS Coefficients

Individual	Aggregate	SNR	β_1	α_1
Underreaction	Underreaction	1.40	0.12	0.46
Overreaction	Underreaction	0.40	-0.14	0.18

Notes: The table simulates the errors-on-revision coefficient at the forecaster-level (β_1) and the consensus-level (α_1) for two different simulated signal-to-noise ratios (SNR).

consider the difference in parameter values between models of individual underreactions and models of individual overreactions. As the signal-to-noise ratio (SNR) is reduced from 1.40 to 0.40, the sign of β_1 changes.[21] This result suggests that observed over- and underreactions at the individual level can be explained by different underlying data generating processes.

I also plot the simulated β_1 distribution across high signal-to-noise ratio and low signal-to-noise ratio parameterizations. For each of 2,000 simulations, I generate a panel of 250 forecasters over 200 periods. I then collect the errors and revisions for these forecasters and compute β_1. Figure 2·3 plots the density of β_1 across the simulations. The results confirm that the model can generate error predictability, and that over- and underreactions depend on the signal-to-noise ratio.[22]

I next turn to the SPF data to document facts consistent with this mechanism.

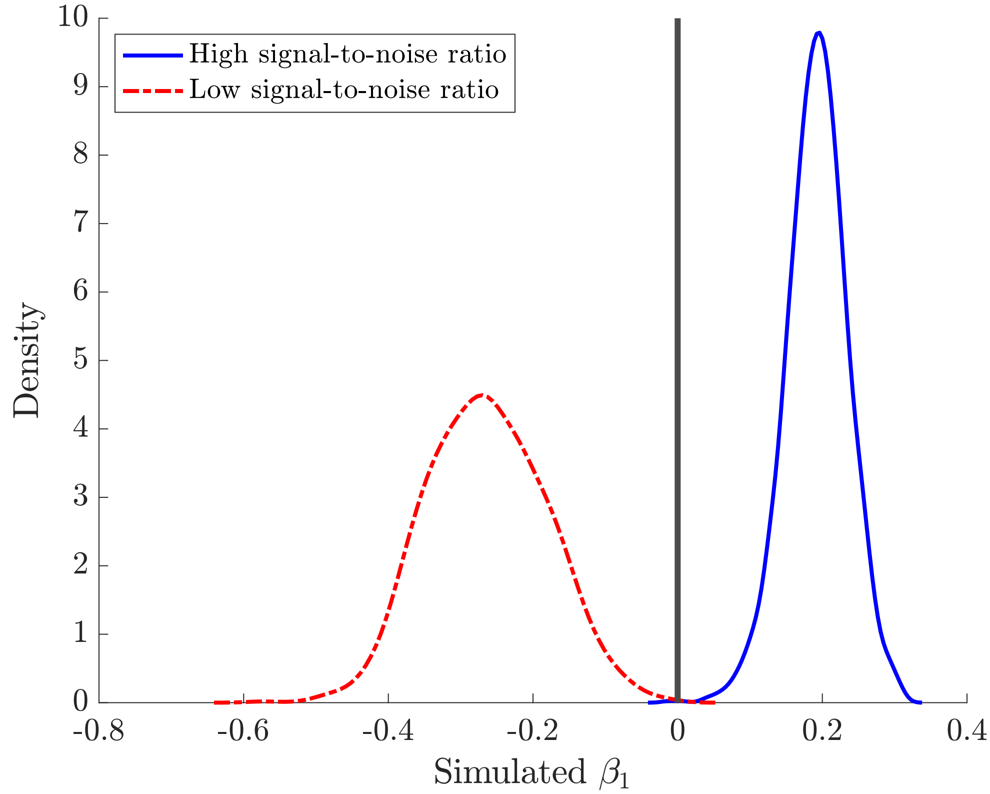
2.5 Evidence from the Survey of Professional Forecasters

This section exploits variation across the macroeconomic variables reported in the SPF in order to document evidence consistent with the idea that the signal-to-noise

[21] Although robust and reliable estimates of the SNR are currently sparse in this literature, CG provide some estimates using cross-country data which are in line with the simulated SNR values here. In addition, these simulated SNR values are similar to those that I quantify in Section 2.6.

[22] If all forecasters made use of the particle filter, however, then in large samples we would expect this distribution to be centered at zero.

Figure 2·3: Overreactions, Underreactions, and Driving Process

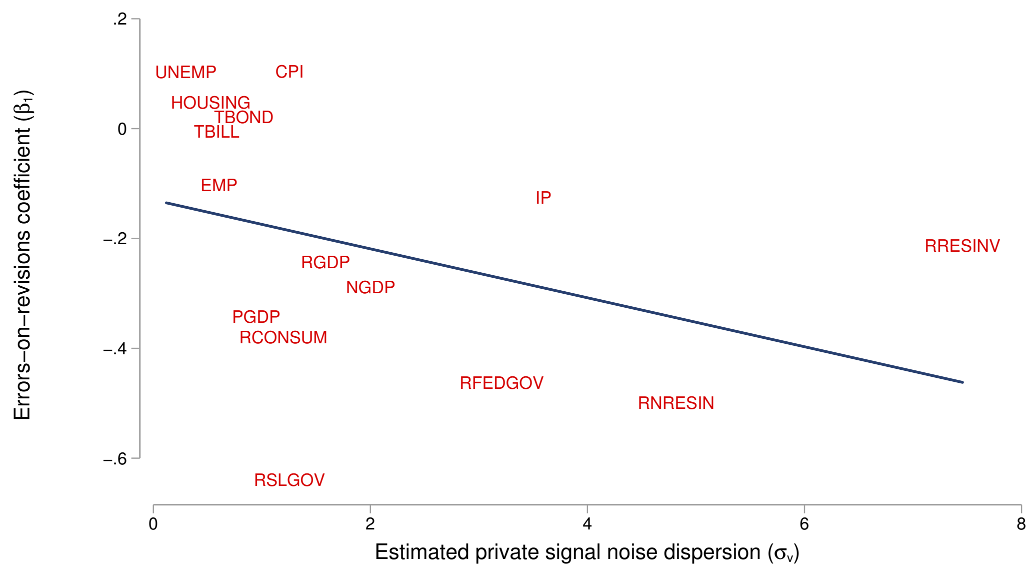


Notes: The figure plots two simulated densities of β_1 arising from a pooled individual-level errors-on-revisions regression from the stylized model. The red dashed line plots the simulation in which private signal noise variance is relatively high whereas the solid blue line plots the simulation in which private signal noise variance is relatively low.

ratio is the key driver of over- and underreactions. The next section will parameterize the model in order to speak to simultaneous over- and underreactions within forecaster.

Each of the variables is presumed to follow a specific data generating process. As a result, $\beta_1 = \beta_1(\rho, \sigma_v, \sigma_e, \phi_0, \phi_1, \sigma_\eta)$ will, in general, vary in the cross-section of SPF variables. I document four facts by measuring proxies for signal and noise. I find that variables exhibiting greater unconditional volatility tend to be variables for which we observe underreactions. On the other hand, variables that feature elevated amounts of noise are associated with observed overreactions.

Figure 2·4: Error Predictability and Revision Dispersion



Notes: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated private noise dispersion, proxied by the interquartile range of forecast revisions. Slope of fitted line is -0.045.

Testable Prediction 1: Error Predictability and Private Noise

In the stylized model, forecasters revise their predictions according to the realization of the lagged public signal as well as their contemporaneous private signal. The noise therefore feeds into the forecast revision. From the perspective of the model, the variance of private signal noise determines the amount of dispersion in revisions across forecasters. More dispersed signal noise admits more pronounced cross-sectional differences in revisions.

With this insight, I collect the pooled β_1 coefficients across SPF variables and plot these against the interquartile range of revisions across forecasters for each variable. Figure 2·4 displays the results. As the model suggests, variables exhibiting greater dispersion in revisions tend to be those for which forecasters overreact.

Testable Implication 2: Error Predictability and Public Noise

While Figure 2·4 relates β_1 to private signal noise, there is also common noise present in the model. I next turn to measure the noisiness of the public signal. Whereas the SPF variable of interest has sometimes been modeled as the latent state in the literature, it is best thought of as a lagged public signal. This is because the SPF variables are observed by all forecasters with a lag. With this in mind, the official government revisions made to these variables across different vintages can provide a partial measure of public signal noise. Assuming that the vintages following the initial real-time release of the variable eliminate some of the common noise, one can quantify these revisions over time. As a matter of notation, define x_t^I as the real-time initial data release for a given variable, and x_t^L as the last release of the variable. Then, we can define $\text{noise}_t^{\text{public}} = \text{Var}(x_t^I - x_t^L)$. I construct this variable from the first and last data vintage for all SPF variables in my sample, and then measure the dispersion of this public noise over time. Figure 2·5 relates β_1 with this measure of public signal noise. The results are consistent with the intuition of the model: variables exhibiting higher measured noise dispersion tend to deliver observed overreactions.

Testable Prediction 3: Error Predictability and Unconditional Volatility

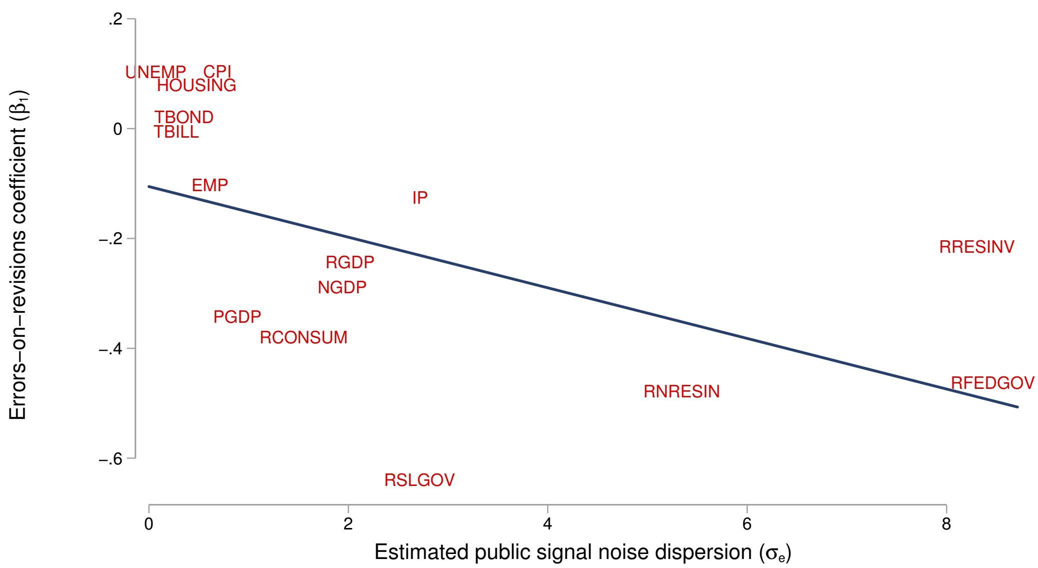
Moreover, the model predicts that with more unconditional variability in the state, there is less scope for overreaction. To test this, I proceed to estimate $\{\phi_{j,0}, \phi_{j,1}\}$ for each SPF variable, j. I then construct an estimate of unconditional volatility:[23]

$$\text{vol}_j = \exp\left(\frac{\widehat{\phi_{j,0}}}{2(1 - \widehat{\phi}_{j,1}^2)}\right).$$

Figure 2·6 relates β_1 to vol_j. The figure supports the hypothesis that variables exhibiting more variability in the state tend to provide greater scope for underreactions.

[23] I estimate the parameters of the stochastic volatiliity model, $\{\phi_0, \phi_1, \sigma_\eta\}$ using MCMC techniques (Kastner & Frühwirth-Schnatter, 2014).

Figure 2·5: Error Predictability and Public Noise



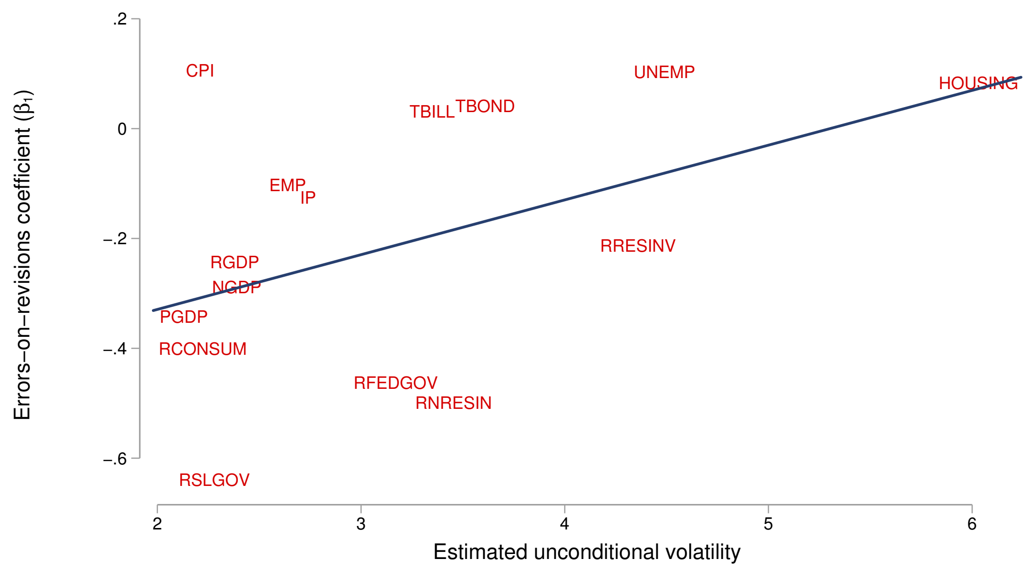
Notes: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated public noise dispersion, proxied by the standard deviation of government revisions to real-time data. Slope of fitted line is −0.046.

Furthermore, note that the variance of the state is increasing in ρ. Hence, the model predicts that more persistent variables will reduce the scope for overreactions. This is consistent with Bordalo et al. (2020) who verify this empirically.

Testable Prediction 4: Error Predictability and Release Frequency

As an additional way to measure signal precision, I consider the frequency with which these different variables are made available to the public. While professional forecasters report predictions in each quarter, some variables are made available at higher frequencies. Specifically, the SPF conducts its survey at roughly the middle of each quarter. However, some of the SPF variables are released at a monthly frequency. For instance employment statistics are released on the Friday of each month. The survey asks forecasters to provide a quarterly average of these series. Furthermore,

Figure 2·6: Error Predictability and Unconditional Volatility

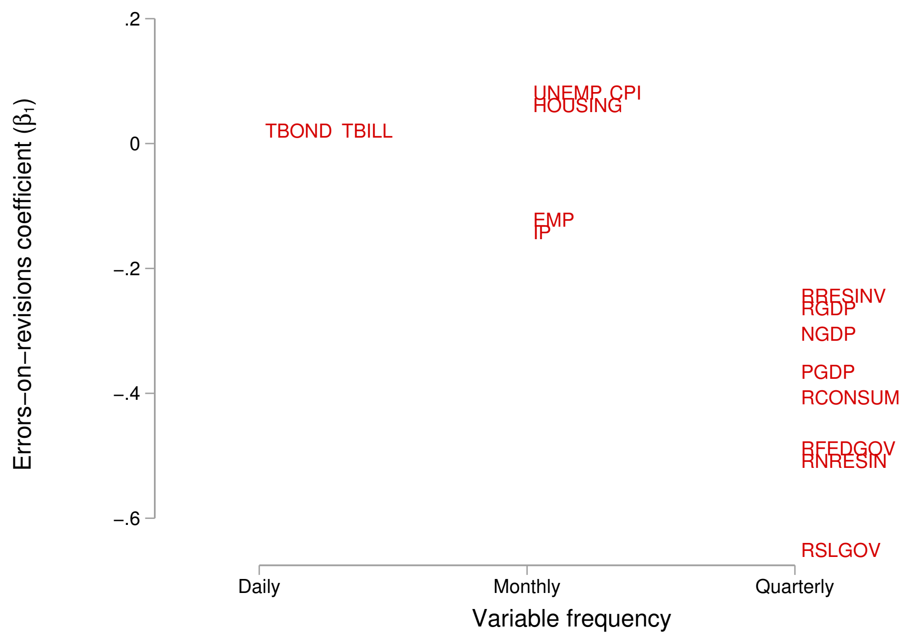


Notes: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated unconditional volatility of the state, $\exp\left(\frac{\widehat{\phi_{j,0}}}{2(1-\widehat{\phi}_{j,1}^2)}\right)$. Slope of fitted line is 0.100.

the financial time series are available at a daily frequency. As a result, forecasters have arguably more information pertaining to the eventual value of some variables in a given quarter than others. This reduces the effective noise in the lagged public signal.[24] Hence, variables available at higher frequencies should raise the scope for underreaction. Note that this does not preclude overreactive behavior in financial markets as has been readily documented. Here, I simply argue that quarterly (average of daily observations) predictions of a financial variable are better informed by the presence of daily observations through the middle of the quarter when the reported forecast is requested. On the other hand, the latest information that forecasters have for quarterly variables, such as GDP, is the previous quarter's release and an advance

[24] Alternatively, one could suppose that forecasters receive an additional informative public signal for daily or monthly SPF variables.

Figure 2·7: Error Predictability and Release Frequency



Notes: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against variable's release frequency {Daily, Monthly, Quarterly}.

estimate. Since there is additional information available for some variables and not others, and the existence of this additional information depends on the variable frequency, then it follows that there is more scope for underreaction among variables that are available at higher frequencies. Figure 2·7 confirms this.

Jointly Testing for Overreaction and Underreaction Channels

As an additional check, I formally test for these channels jointly. For the data to accord with this theory of expectations, it should be the case that an interaction of the forecast revision with each of these variables either raises or reduces the extent to which β_1 is negative in the pooled specification (column 1 of Table 2.1). To complete this exercise, I incorporate two new regressors (and all possible interactions), each capturing a source of either noise or state volatility. As a measure of noise, I select the release frequency explained above. For my measure of fundamental volatility, I take a factor analysis approach. Since the latent state and its volatility are unobservable, it is natural to consider an index of the shared variation among all SPF variables. From this exercise, I obtain a time-varying index of what I call fundamental volatility.[25]

Given these two new regressors, I modify the baseline errors-on-revisions regression (pooling across all SPF variables as in Table 2.1). Specifically, in addition to projecting errors on revisions, I specify a quarterly release indicator and the constructed fundamental volatility index. I also include interactions of each of these with the forecast revision as well as all interactions with each other. The regression results are reported in Table 2.4.

The first column of the table reproduces the first column of Table 2.1 for the relevant observations. The second column reports the fully specified regression. The relationship of interest remains the extent to which the forecast error and revision are related. The only interaction terms to enter statistically significantly are those crossed with the forecast revision. Furthermore, the signs of these two interactions are consistent with the expected signs according to my theory of expectations under unobserved time-varying volatility. In particular, noise raises the scope for overreactions as evidenced by the negative cross term between the quarterly frequency indicator and the revision. On the other hand, fundamental volatility reduces the scope for overreaction as seen by the positive coefficients in front the volatility index and the

[25]For this analysis, I drop nominal GDP since its components reside in my data set. Furthermore, I exclude CPI due to its shorter available history. For the remaining macroeconomic variables, I compute five-year rolling standard deviations and then estimate underlying principal factors. The results deliver two factors that explain roughly equal amounts of the common variance of the final vintage of SPF variables. Based on the factor loadings, I call the first factor a real residential factor, and the second a real non-residential factor (the residential factor loads highly on housing and real residential investment whereas the second factor does not). While both factors deliver the correct sign in my regression specification, I report the regression that specifies the real non-residential factor as it delivers statistically detectable results. Appendix A.2 confirms that the results are robust to window length.

Table 2.4: Modified Forecast Error Predictability Regressions

	Forecast Error (1)	(2)
Revision	-0.314***	-0.165**
	(0.0414)	(0.073)
Revision × Quarterly		-0.194**
		(0.089)
Revision × Fundamental Volatility		0.106**
		(0.041)
Observations	58,740	58,740

Notes: The table reports estimated coefficients of forecast error predictability across two specifications. The Quarterly indicator is equal to 1 if the SPF variable is released at a quarterly frequency and 0 otherwise. The Fundamental Volatility variable is a time-varying index constructed as described in the text. In addition to the interactions reported in the table, the column (2) specification includes the individual variables and their interactions as controls. Standard errors are as in Driscoll and Kraay (1998). Data used for estimation come from SPF (1964Q4-2018Q3). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

revision.

The cross-sectional correlations and regression results reported above confirm that the model mechanism is consistent with the SPF. In the next section, I parameterize the model in order to show that it is also consistent with the extent to which a forecaster simultaneously over- and underreacts to different variables as in the data.

2.6 Parameterization

The model is able to generate over- and underreactions across levels of aggregation and across variables, pooling over forecasters. In this section, I parameterize the stylized model for real GDP and unemployment in order to demonstrate that it can also generate simultaneous over- and underreactions *within* forecaster.

Specifically, I calibrate the public noise dispersion, the persistence of the latent state, and the stochastic volatility parameters $\{\sigma_e, \rho, \phi_0, \phi_1, \sigma_\eta\}$. I then find the values of $\bar{c}_{PF}$ and σ_v that minimize the distance between the model-simulated and empirical estimates of the errors-on-revisions coefficients at the forecaster and consensus-levels (β_1 and α_1).

For real GDP and unemployment respectively, I set σ_e equal to the standard deviation of the data revisions made to each variable over the sample period. The data revision is taken to be the difference between the first and final release of the data series. For the remaining parameters, I consider the revised data rather than the real-time data. Intuitively, these series should be more highly correlated with the unobserved latent state. I then estimate an AR(1) on the revised series and set ρ equal to the estimated AR(1) coefficient. Finally, I collect the squared residuals from this autoregression and estimate $\{\phi_0, \phi_1, \sigma_\eta\}$. Panel A of Table 2.5 reports the calibration for each variable.

I then parameterize σ_v and $\overline{c}_{PF}$ by minimizing the distance between the model-implied $\{\beta_1, \alpha_1\}$ from its empirical counterpart. Since I am calibrating these parameters for GDP and unemployment, the procedure amounts to searching a four-dimensional parameter space and matching four moments. For each simulation, I generate two state variables according to the dynamics described in Section 2.4. I then simulate the lagged public signal as well as the contemporaneous private signal for each variable. In every period, forecasters report a forecast for each variable according to the state dynamics, signals received, and loss function described in Section 2.5. From this simulated panel of forecasters, I construct the errors-on-revisions coefficients. I minimize the distance between the simulated and empirical OLS coefficients by making use of simulated annealing, a standard global stochastic optimization routine.

The results, reported in Panel B of Table 2.5, indicate that real GDP is characterized by more private signal noise than the unemployment rate. Furthermore, as shown in Panel C, the implied signal-to-noise ratio for real GDP is about 0.50 whereas

Table 2.5: Parameterization

Parameter	Description	Unemployment	Real GDP
Panel A: External			
Parameter	Description	Unemployment	Real GDP
ρ	State persistence	0.98	0.30
σ_e	Standard deviation of public noise	0.07	2.02
ϕ_0	Level of log variance	-0.76	0.14
ϕ_1	Persistence of log variance	0.73	0.92
σ_η	Volatility of log variance	0.69	0.39
Panel B: Internal			
Parameter	Description	Unemployment	Real GDP
σ_v	Standard deviation of private noise	0.12	2.81
$\bar{c}_{PF}$	PF cost upper bound	7.84	3.79
Panel C: Implied values			
	Description	Unemployment	Real GDP
	Signal-to-noise ratio	1.36	0.50
	Share using PF	0.53	0.82

Notes: The table reports parameterization for unemployment and real GDP. Panel A reports the external parameterization. The stochastic volatility parameters $\{\phi_0, \phi_1, \sigma_\eta\}$ are estimated according to the algorithm presented in Kastner and Frühwirth-Schnatter (2014). Panel B reports internal parameterization obtain through the minimum distance procedure described in the text. Based on this calibration, Panel C reports the implied signal-to-noise ratio and share of forecasters that utilize the PF model.

it is 1.36 for unemployment. This is consistent with the intuition of the model as well as the cross-sectional evidence in the previous section: variables that exhibit higher signal-to-noise ratios tend to be the variables for which underreactions are observed.

In addition, the cost distribution parameters indicate that costs to implementing the particle filter for real GDP are lower on average relative to the unemployment rate. The discrepancy between these two cost parameters can be attributed to the fact that mean square errors in the data are much larger in magnitude for real GDP. These costs govern in part the incentives to adopt the particle filter and, as reported in Panel C, imply that roughly 82% of forecasters optimally select to forecast with the particle filter for real GDP. On the other hand, 53% of forecasters choose the

Table 2.6: Model Fit

	Model	Data
Unemployment		
Errors-on-revisions, forecaster-level (β_1)	0.149	0.082
Errors-on-revisions, consensus (α_1)	0.212	0.247
Real GDP		
Errors-on-revisions, forecaster-level (β_1)	-0.263	-0.264
Errors-on-revisions, consensus (α_1)	0.352	0.350

Notes: The table reports empirical and model-implied moments. The calibration directly targets the errors-on-revisions moments for unemployment and real GDP at the forecaster and consensus-levels. The final row reports the share of forecasters that overreact to real GDP and simultaneously underreact to unemployment.

particle filter as the forecasting model of choice for unemployment.

Table 2.6 reports the model fit. The minimum distance procedure was able to successfully match patterns of over- and underreactions observed in the pooled forecaster-level and consensus regressions for real GDP. Though the fit for unemployment is not as tight, the model can fairly closely match the consensus-level coefficient and can qualitatively match the forecaster-level coefficient.

Lastly, I assess the calibrated model's ability to match untargeted moments. Table 2.7 reports the mean square error, standard deviation of errors, and the share of over- and underreaction. Although the model generally overstates the mean square error, it successfully matches the relative magnitudes across both variables. Furthermore, the model is broadly successful in matching the dispersion of forecast errors. Lastly, the stylized model is able to successfully match the share of simultaneous over- and underreactions. Figure 2·1 reports that about 64% of forecasters in the sample overreact to real GDP while simultaneously underreacting to unemployment. Based on the parameterization devised here, the stylized model 63% of simulated forecasters overreact real GDP and underreact to unemployment.

Table 2.7: Untargeted Moments

	Unemployment	Real GDP
Mean square error		
Model	0.107	7.150
Data	0.045	6.005
Standard deviation of forecast error		
Model	0.320	2.670
Data	0.209	2.451
Share that overreact to real GDP and underreact to unemployment		
Model	0.632	
Data	0.641	

Notes: The table reports empirical and model-implied moments. The calibration directly targets the errors-on-revisions moments for unemployment and real GDP at the forecaster and consensus-levels. The final row reports the share of forecasters that overreact to real GDP and simultaneously underreact to unemployment.

2.6.1 Implications for Information Rigidities

What does time-varying volatility coupled with noisy information imply about interpreting the coefficient α_1 as an information rigidity? Based on my model, it is apparent that α_1 does not cleanly map to the Kalman gain as it does in the scalar linear context.[26] The key intuition of Bayesian filtering, however, still holds, and the optimal weight placed on innovation errors remains a sufficient statistic for capturing the rate of learning. This weight depends on the covariances of the state estimation error and the measurement error. Quantifying the rate of learning, however, is not readily feasible from a projection of mean errors on mean revisions as has traditionally been suggested in the literature (An, Liu, & Wu, 2021; Coibion & Gorodnichenko, 2015; Dovern et al., 2015; Larsen, Thorsrud, & Zhulanova, 2020).

[26]See Appendix A.2.

In fact, from the perspective of the stylized model, the coefficient coming from errors on revisions regressions at the consensus-level may reveal misleading insights on the extent of information rigidity. Whereas a large α_1 coefficient would typically imply more information rigidities, here, α_1 is larger when the signal-to-noise ratio is high. On the other hand, a negative α_1 arises when signals are less informative. This suggests that the reduced form coefficient α_1 may be limited in what it reveals about genuine information frictions.

2.6.2 Implications for State Dependence

What does this mean for state dependence? If β_1 and α_1 were to rise in recessions, then the model would imply that the signal to noise ratio is countercyclical and information rigidities fall during economic downturns. If, on the other hand, these coefficients fall, then the signal-to-noise ratio is procyclical and information rigidities actually rise in recessions. CG document evidence indicating that α_1 falls in recessions. They interpret this as a reduction in information rigidities, however, my model would suggest that this implies a *rise* in information rigidities since it implies that the system experiences elevated amounts of noise. This is an important distinction between my model and the extant literature as it delivers an opposite answer to the question of whether individuals trust their signal more or less in recessions.

However, after performing a similar exercise to that in Table 2.4 by interacting a quarterly recession indicator with forecast revisions, I find no evidence that β_1 or α_1 changes with the business cycle. I also run this exercise by replacing the recession indicator with revised real GDP growth. It is possible, however, that the signal-to-noise ratio is insensitive to business cycle fluctuations because both the state and the signal experience stochastic volatility. I abstract away from volatility in signal precision, so there is a limit as to what one could glean from this model as it pertains to state dependence of information rigidities.

Nonetheless, one could distinguish between two types of uncertainty: fundamental uncertainty and information uncertainty. The first maps to time-varying volatility in the state while the latter arises when signal noise experiences stochastic volatility. There is a literature that stresses the importance of uncertainty shocks. These are often modeled as fundamental uncertainty shocks. Shocks to information precision are also studied in the literature (Dun Jia, 2016). According to my model, the state dependence of β_1 and α_1 depend on the signal-to-noise ratio which in turn depends on how these two types of uncertainty evolve relative to one another over the business cycle. If both rise in recessions, then is possible that the signal-to-noise ratio is acyclical thereby rendering β_1 and α_1 roughly constant over the cycle as well.

2.7 Conclusion

This paper documents that individual forecasters appear to simultaneous over- and underreact to new information. Existing models of belief formation are unable to flexibly accommodate these empirical patterns. This paper shows that a noisy information model incorporating unobserved time-varying volatility can make sense of these facts. Forecasters optimally select different models based on the complexity of the state dynamics. Heterogeneity in predictor functions can jointly deliver coincident over- and underreactions among forecasters. In particular, forecasters overreact to variables that exhibit more noise whereas they underreact to variables that are characterized by less noise. I uncover evidence in favor of this mechanism, demonstrating that fluctuations in volatility matter for belief dynamics.

Chapter 3

A New Fact to Discipline Models of Beliefs

3.1 Introduction

Expectations are ubiquitous in economics. Nonetheless, the manner in which expectations are formed remains an open question. The benchmark theory of full information rational expectations (FIRE) posits that agents form beliefs in an optimal manner such that ex-post forecast errors are unpredictable. Error orthogonality, however, has been found to be inconsistent with survey data on individual expectations. In particular, ex-ante updates made to forecasts are systematically correlated with ex-post errors. Such error predictability has motivated departures from FIRE in the literature, particularly in favor of non-rational theories (Bordalo et al., 2020; Broer & Kohlhas, 2019; D'Acunto, Hoang, Paloviita, & Weber, 2019; Kohlhas & Walther, 2020).

This paper shows that error predictability, while sufficient to reject FIRE, is not particularly informative about individual rationality. I show this by way of example. Specifically, I narrow my focus to models that explain overreaction, a salient feature of survey data. I consider three prominent non-FIRE theories, embedded in a noisy information environment: overconfidence (Daniel et al., 1998), diagnostic expectations (Bordalo et al., 2020), and strategic interaction (Woodford, 2001). I then prove that all three of these models can deliver identical patterns of error predictability.

Whereas the first two models feature non-rational expectations, the model of strategic interaction is rational. Intuitively, each of these theories features a key parameter that governs the extent of overreactive behavior. I show that the key parameters can be mapped to each other, thereby delivering identical relationships between ex-post expectation errors and ex-ante revisions.

The three candidate models, however, are not observationally equivalent in general. I offer a simple testable implication that is able to distinguish across all three of these models, and that can provide a way forward more broadly as the literature assesses which non-FIRE theories are consistent with the data. The relevant testable implication is the serial correlation of revisions. Models of overconfidence imply a negative autocorrelation coefficient for revisions, distinct from the aforementioned errors-on-revisions coefficient. Whereas diagnostic expectations also deliver a negative autocorrelation coefficient for revisions, this theory implies that the autocorrelation of revisions is *identical* to the errors-on-revisions coefficient. On the other hand, the model of strategic interactions requires that revisions made to a fixed event forecast be unpredictable.

When taking this testable implication to data from the Survey of Professional Forecasters (SPF), I document evidence in against models of strategic interaction. For most variables in the SPF, I find that estimates of error predictability and revision persistence are not statistically distinct from one another. As a result, of the three models considered, the data favor diagnostic expectations. Nonetheless, there are some variables for which these estimates diverge significantly, indicating that a theory of overconfidence provides a better fit for those variables.

My findings support non-rational theories of overreaction under standard linear dynamics.[1] Whereas models featuring Bayesian updating can deliver error predictability conditional on modifications made to the forecaster's objective, these mod-

[1] Models featuring nonlinear dynamics have also been found to rationalize the data (Ortiz, 2020).

els are generally unable to deliver revision predictability.[2] Bayesian updating requires that information be used efficiently, thereby preventing past revisions from holding predictive power over current revisions. Hence, analyzing the time series properties of revisions serves as a way to narrow the set of non-FIRE models that are consistent with the data.

Several studies have used survey data to test for FIRE.[3] In essence, FIRE implies that expectation errors, e_t, are orthogonal to the individual forecaster's information set, $\mathcal{I}_t$:

$$\mathbb{E}(e_t|\mathcal{I}_t) = 0.$$

This implication, however, is at odds with the data. In particular, ex-ante forecast revisions predict ex-post errors. Since a rejection of FIRE can be due to either a rejection of full information or a rejection of rationality, existing non-FIRE theories can be classified as fundamentally rational or non-rational. Rational non-FIRE theories typically incorporate a particular motive for deviating from the minimum mean square forecast. On the other hand, non-rational theories introduce a behavioral bias. In either case, expectation errors can be characterized as:

$$\widehat{\mathbb{E}}(e_t|\mathcal{I}_t) \neq 0.$$

Nordhaus (1987) first introduced the testable implication proposed in this paper which states that revisions are not serially correlated. This is an implication of the fact that under FIRE, revisions r_t, must be unpredictable conditioning on the forecaster's information set:

$$\mathbb{E}(r_t|\mathcal{I}_t) = 0.$$

[2] There are two documented exceptions in the extant literature: models featuring smoothing motives and noisy memory. Section 3.5 discusses these theories in more detail.

[3] Examples include Coibion and Gorodnichenko (2015), Bordalo et al. (2020), Fuhrer (2018), Dovern et al. (2015), Crowe (2010), Paloviita and Viren (2013), Bürgi (2016), Andrade and Bihan (2013)

The key contribution of this paper is to study non-FIRE theories in which:

$$\widehat{\mathbb{E}}_t(r_t|\mathcal{I}_t) = 0 \quad \text{despite} \quad \widehat{\mathbb{E}}_t(e_t|\mathcal{I}_t) \neq 0,$$

and to show that such theories are inconsistent with the data. Because serially correlated revisions imply error predictability (but not the other way around), this paper offers a stronger fact to discipline models of beliefs.

The focus of this paper is on overreaction. It should be noted that a recent literature has documented the existence of both over- and underreaction in survey expectations (Broer & Kohlhas, 2019; Kohlhas & Walther, 2020; Ortiz, 2020). There is at present no unified definition of over- and underreaction. For instance Broer and Kohlhas (2019) document overreaction in the form of a negative relation between errors and revisions, and underreaction to some endogenous public signals. On the other hand, Kohlhas and Walther (2020) define overreactions as a negative relation between errors and current output growth, while underreactions are defined as a the relation between errors and revisions at the consensus-level. Finally, Ortiz (2020) documents that the same forecaster simultaneously exhibits positive and negative coefficients of error predictability for distinct macroeconomic variables. Though I abstract from these empirical regularities here, my findings are relevant for this strand of the literature as well. In particular, theories of simultaneous over- and underreaction that are to be consistent with professional forecasts must be able to deliver serially correlated revisions.

The rest of the paper is organized as follows. In Section 3.2, I document empirical facts related to overreaction in survey expectations, based on both error and revision predictability. In Section 3.3, I introduce the three candidate models of overreaction. Section 3.4 maps the models' key parameters to each other and shows that all models can deliver the same relation between errors and revisions. I then characterize revision predictability from the lens of each model in Section 3.5. Section 3.6 concludes.

3.2 Evidence of Overreaction in Survey Expectations

Many non-FIRE theories have been devised to explain overreactive behavior (Afrouzi, Kwon, Landier, Ma, & Thesmar, 2020; Benigno & Karantounias, 2019; Bordalo et al., 2020; Fuster, Hebert, & Laibson, 2012b; Rozsypal & Schlafmann, 2019). The literature often utilizes survey data from the Survey of Professional Forecasters in order to assess the empirical performance of different theories. The SPF is a quarterly survey provided by the Federal Reserve Bank of Philadelphia. The survey provides forecasts from several forecasters across a number of macroeconomic variables over many horizons.

A popular test often implemented for empirical motivation projects expectation errors on revisions. Suppose that x_t is the target variable and $x^i_{t+h|t}$ is forecaster i's forecast devised at time t for horizon h. Then the empirical test is defined as:

$$\underbrace{x_{t+h} - x^i_{t+h|t}}_{\text{Error}} = \beta_0 + \beta_1[\underbrace{x^i_{t+h|t} - x^i_{t+h|t-1}}_{\text{Revision}}] + \epsilon^i_{t+h}. \tag{3.1}$$

At the forecaster-level, the coefficient in front of the revision is found to be negative for a range of macroeconomic variables, β_1.[4] Table 3.1 confirms this finding. Pooling across horizons, the results indicate that errors covary negatively with revisions. For instance, a one percentage point upward revision for real GDP is associated with a 0.23 percentage point more negative forecast error. Following positive news, forecasters tend to revise upward excessively such that the overly-optimistic update results in a systematically more negative subsequent error. Hence, professional forecasters tend

[4] Though there is also evidence of underreaction for some variables, I focus on overreaction due to its pervasiveness at the forecaster-level. Moreover, the model comparison exercise that I later undertake is without much loss of generality. In particular, a theory of underreaction that features strategic complementarity would imply a positive covariance between errors and revisions, but it would nonetheless require that revisions be serially uncorrelated.

Table 3.1: Forecast Error Predictability

Variable	Estimate	Standard Error	Observations
CPI	-0.103	0.122	7,052
Industrial production	-0.177	0.055	11,491
Nominal GDP	-0.291	0.044	11,994
GDP deflator	-0.297	0.057	11,951
Real consumption	-0.373	0.072	8,865
Real federal government spending	-0.499	0.061	8,273
Real GDP	-0.229	0.058	12,110
Real nonresidential investment	-0.228	0.088	8,604
Real residential investment	-0.192	0.059	8,587
Real state and local government spending	-0.542	0.027	8,282
Ten-year government bond	-0.079	0.059	7,246

Notes: The table reports the estimated β_1 coefficient of error predictability from (3.1). Samples for reach regression are pooled across horizons. Standard errors are clustered by forecaster and quarter.

to over-revise their forecasts.

Overreaction can similarly be observed by analyzing the persistence of fixed-event updates. Empirically, revision predictability estimates the following relationship:

$$\underbrace{x^i_{t+h|t} - x^i_{t+h|t+h-1}}_{\text{Revision}} = \gamma_0 + \gamma_1[\underbrace{x^i_{t+h|t+h-1} - x^i_{t+h|t+h-2}}_{\text{Previous revision}}] + \nu^i_{t+h}, \tag{3.2}$$

If $\gamma_1 \neq 0$, then the fixed event revisions are serially correlated (predictable). Whereas testing for error predictability is more common in the literature, revision predictability was previously analyzed in Dovern et al. (2015).

Table 3.2 reports estimates of (3.2). Across the same set of variables for which $\beta_1 < 0$, there is also ample evidence of $\gamma_1 < 0$. Focusing on real GDP, a one percentage point upward revision in the forecast for $t + h$ devised today is associated with a 0.23 percentage point downward revision in the same forecast made tomorrow. Jointly, the results could be interpreted as follows: amid the realization of positive news today, forecasters appear to over-revise their predictions upward only to observe a

Table 3.2: Forecast Revision Predictability

Variable	Estimate	Standard Error	Observations
CPI	-0.250	0.064	7,052
Industrial production	-0.202	0.030	11,491
Nominal GDP	-0.223	0.032	11,994
GDP deflator	-0.255	0.031	11,951
Real consumption	-0.240	0.042	8,865
Real federal government spending	-0.283	0.029	8,273
Real GDP	-0.211	0.033	12,110
Real nonresidential investment	-0.165	0.042	8,604
Real residential investment	-0.165	0.039	8,587
Real state and local government spending	-0.317	0.025	8,282
Ten-year government bond	0.003	0.052	7,246

Notes: The table reports the estimated β_1 coefficient of error predictability from (3.2). Samples for reach regression are pooled across horizons. Standard errors are clusted by forecaster and quarter.

systematically more negative forecast error tomorrow, which subsequently contributes to a more pessimistic revision tomorrow.

Overreaction is pervasive among professional forecasters. In the subsequent section, I review three prominent models of overreaction that have been proposed in the literature.

3.3 Three Models of Overreactive Behavior

All three models of overreaction are grounded in a noisy information setting with heterogeneous beliefs. Optimal forecasts in this context are consistent with the mathematical expectations operator, $\mathbb{E}$. This will be the benchmark from which all other expectations will deviate. In keeping with Coibion and Gorodnichenko (2015) and

Bordalo et al. (2020) consider the following set up:

$$\begin{aligned} \text{Exogenous State:} \quad & x_t = \rho x_{t-1} + w_t, \quad && w_t \sim \mathcal{N}(0, \sigma_w^2) \\ \text{Private Signal:} \quad & y_t^i = x_t + v_t^i, \quad && v_t^i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_v^2). \end{aligned}$$

The exogenous fundamental follows a simple AR(1) process.[5] Moreover, agents have access only to private information in the form of a noisy private signal y_t^i observed each period. I abstract away from more complex signal structures for simplicity, however, the assumptions on how agents receive information are unimportant for the results presented in subsequent sections.

From the Kalman filter (Kalman, 1960), the optimal current-period prediction of x_t is:

$$\mathbb{E}(x_t|\mathcal{I}_t^i) \equiv x_{t|t}^i = (1-\kappa)x_{t|t-1}^i + \kappa y_t^i,$$

where $\mathcal{I}_t^i$ denotes individual i's information set at time t, and κ refers to the steady-state Kalman gain, $\kappa = \frac{\text{Var}(x_t - x_{t|t-1}^i)}{\text{Var}(x_t - x_{t|t-1}^i) + \sigma_v^2}$ which is the optimal (in the mean-square sense) weight placed on new information.[6]

3.3.1 Overconfidence

First, I consider a non-rational model of overconfidence. Daniel et al. (1998) presents a theory of overconfidence in which individuals perceive their private signals to be more precise than they truly are. Intuitively, because forecasters believe their information to be more precise, they will tend to trust their signals more than is optimal. As a result, overconfident forecasters over-weight new information, thereby generating a over-revisions and delivering a negative covariance between updates and errors.

[5] The dynamics of the state are innocuous for the results of the paper.

[6] Note that the imperfect information enviornment is a more general formulation than a full-information environment. The model collapses to full-information rational expectations when $\sigma_v = 0$ so that $\kappa = 1$.

Specifically, forecasters perceive:

$$y_t^i = x_t + v_t^i \qquad v_t^i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathring{\sigma}_v^2)$$

where $\mathring{\sigma}_v = \alpha\sigma_v$ such that $\alpha \in [0, 1]$.

As a matter of notation, let the forecaster's current period forecast be denoted by $\mathring{x}_{t|t}^i$, and his one step ahead forecast by $\mathring{x}_{t|t-1}^i$. Forecasters invoke the Kalman filter in order to formulate their expectations. As a result, expectations are determined according to the following predict-update procedure:

$$\begin{aligned} \mathring{x}_{t|t-1}^i &= \rho \mathring{x}_{t-1|t-1}^i && \text{(Predict)} \\ \mathring{x}_{t|t}^i &= \mathring{x}_{t|t-1}^i + \mathring{\kappa}(y_t^i - \mathring{x}_{t|t-1}^i) && \text{(Update).} \end{aligned}$$

Having defined the expectations formation process under overconfidence, we obtain the following result.

Proposition 1. *The OLS coefficient arising from an errors-on-revisions regression in the overconfidence model is:*

$$\beta_1^{OC} = \frac{\mathbb{C}(x_t - \mathring{x}_{t|t}^i, \mathring{x}_{t|t}^i - \mathring{x}_{t|t-1}^i)}{\mathbb{V}(\mathring{x}_{t|t}^i - \mathring{x}_{t|t-1}^i)} = \frac{(\alpha^2 - 1)\sigma_v^2}{\mathbb{V}(x_t - \mathring{x}_{t|t-1}^i) + \sigma_v^2} < 0.$$

Proof. See Appendix A.3. □

The coefficient β_1^{OC} is always negative since $\alpha < 1$. In the absence of overconfidence, $\alpha = 1$ and $\beta_1 = 0$. Hence, underlying overconfidence on the part of forecasters generates observed overreaction.

3.3.2 Diagnostic Expectations

Second, I consider a non-rational theory of diagnostic expectations (Bordalo et al., 2020). Under diagnostic expectations, overreactions arise due to the representativeness heuristic of Tversky and Kahneman (1974). Intuitively, more recent information is more easily recalled and therefore overweighted when formulating beliefs. The predict-update procedure for diagnostic forecasters can be described as follows:

$$x^{i,\theta}_{t|t-1} = \rho x^{i,\theta}_{t-1|t-1} \qquad \text{(Predict)}$$
$$x^{i,\theta}_{t|t} = x^{i}_{t|t} + \theta(x^{i}_{t|t} - x^{i}_{t|t-1}) \qquad \text{(Update)},$$

where θ is the belief distortion parameter that governs the extent of overreaction. Under $\theta = 0$, there is no belief distortion, and expectations collapse to the standard rational expectations benchmark.

Bordalo et al. (2020) provide additional details including a derivation of the coefficient of error predictability:

$$\beta_1^{DE} = \frac{\mathbb{C}(x_t - x^{i,\theta}_{t|t}, x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1})}{\mathbb{V}(x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1})} = \frac{-\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2} < 0.$$

3.3.3 Strategic Interaction

Overreactions can also arise rationally, that is, as an optimal outcome. I next consider a model of strategic interaction. The model presented in this section draws from Morris and Shin (2002), Woodford (2001), and Coibion and Gorodnichenko (2015). To obtain overreactions, I assume that forecasters are characterized by strategic substitutability. Intuitively, forecasters have the dual objective to minimize their squared errors and also to distinguish themselves from the average forecast. More specifically, each forecaster wishes the minimize the following loss function:

$$\min_{\{\widetilde{x}^i_{t|t}\}} \mathbb{E}\left[(x_t - \widetilde{x}^i_{t|t})^2 + R(\widetilde{x}^i_{t|t} - F_t)^2 | \mathcal{I}^i_t\right]. \tag{3.3}$$

where x_t is the realized fundamental, $\widetilde{x}^i_{t|t}$ is forecaster i's *reported* current-period forecast, $\mathcal{I}^i_t$ is forecaster i's information set at time t, F_t is the consensus forecast at time t, and $R < 0$ is the degree of strategic substitutability.[7] There are a number of possible microfoundations for strategic substitutability. Most prominently, see Ottaviani and Sørensen (2006). When $R = 0$, the loss function collapses to the familiar mean-squared loss.

The presence of strategic incentives makes higher order beliefs crucial to this model. In particular, the consensus nowcast at time t is denoted by F_t and it is defined as

$$F_t = \frac{1}{1+R}\sum_{k=0}^{\infty}\left(\frac{R}{1+R}\right)^k \mathbb{E}^{(k)}(x_t) = \frac{1}{1+R}x_{t|t} + \frac{R}{1+R}F_{t|t}.$$

where $\mathbb{E}^{(k)}$ is the k^{th}-order expectation of x_t.

Taking the first order conditions of (3.3), it follows that the optimal reported prediction is:

$$\widetilde{x}^i_{t|t} = \frac{1}{1+R}x^i_{t|t} + \frac{R}{1+R}F^i_{t|t}.$$

where $x^i_{t|t}$ is the optimal current-period forecast for the state and $F^i_{t|t}$ is forecaster i's prediction about what the consensus nowcast is at time t.

From this, it follows that the forecast error in this model is:[8]

$$x_t - \widetilde{x}^i_{t|t} = (1-\lambda)\left[x_t - \frac{1}{1+R}x^i_{t|t-1} - \frac{R}{1+R}\rho F^i_{t-1|t-1}\right] - \lambda v^i_t.$$

and the forecast revision is defined as:

$$\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} = \lambda(x_t - x^i_{t|t-1} + v^i_t).$$

where $\lambda = \frac{\kappa_1 + R\kappa_2}{1+R}$ and κ_1 and κ_2 are the elements of the $2\times$ 1 Kalman gain vector, κ.[9]

[7] For $R > 0$, forecasters exhibit strategic complementarities. That is, forecasters have an incentive to stay close to the consensus forecast.

[8] See Appendix A.3 for details.

[9] The Kalman gain vector is two-dimensional because forecasters generate predictions for the

Having defined the errors and revisions according to this model, we can derive the following result.

Proposition 2. *The errors-on-revisions regression coefficient in the strategic interaction model is:*

$$\beta_1^{SI} = \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} < 0.$$

Proof. See Appendix A.3. □

As expected, when $R = 0$, the coefficient $\beta_1^{SI} = 0$, consistent with rational expectations under standard mean squared loss. Given the assumption placed on R, this model can generate overreactions.[10]

All three of these models are capable of explaining the over-revisions observed in the survey data and reported in Table 3.1. Collectively, these three models encompass both rational and non-rational theories of expectation formation. I next demonstrate how these three models can deliver the same patterns of error predicability, implying that the forecast error predictability, while evidence against FIRE, is not particularly informative about rationality. Following this, I show how revision predictability is a more desirable test because it can speak to deviations from FIRE and can deliver sharper implications about rationality.

3.4 Matching Error Predictability

All three models discussed in the previous section are able to generate forecaster-level overreactions. Table 3.3 summarizes the updating rules for each of the three models. Note that these models can be expressed as a (positive) deviation from the conditional

unobserved state, x_t and the unobserved consensus forecast, F_t.

[10] One might wonder whether forecasters can exhibit overreactions under strategic complementarities ($R > 0$). This can only occur if the weight placed private information when predicting the consensus forecast exceeds the weight placed on private information when predicting the state ($\kappa_2 > \kappa_1$). Given that the signal is more informative about x_t than F_t, it is never optimal for the forecaster to set $\kappa_2 > \kappa_1$.

Table 3.3: Update Rules Across Models

Model	Update rule
Overconfidence	$\widehat{x}^i_{t\|t} = x^i_{t\|t} + (1-\widehat{\kappa})\widehat{x}^i_{t\|t-1} - (1-\kappa)x^i_{t\|t-1} + (\widehat{\kappa}-\kappa)y^i_t$
Diagnostic Expectations	$x^{i,\theta}_{t\|t} = x^i_{t\|t} + \theta(x^i_{t\|t} - x^i_{t\|t-1})$
Strategic Interaction	$\widetilde{x}^i_{t\|t} = x^i_{t\|t} - \frac{R}{1+R}(x^i_{t\|t} - F^i_{t\|t})$

Notes: The table reports the updating rules for models of overconfidence, diagnostic expectations, and strategic interaction.

expectation, $\mathbb{E}(x_t|\mathcal{I}^i_t) \equiv x^i_{t|t}$. Hence, forecast updates exceed what is called for by the optimal minimum mean square estimate. Importantly, all three theories can deliver an identical β_1 coefficient.

Proposition 3. *Fix the state and signal parameters, ρ, σ_w, and σ_v. Given one of the parameters in $\{\alpha, \theta, R\}$, there exist values for the other two such that:*

$$\beta_1^{OC} = \beta_1^{DE} = \beta_1^{SI}.$$

Proof. See Appendix A.3. □

This is fundamentally an exercise in matching moments. In light of the recent use of β_1 to motivate non-FIRE models, this result demonstrates that a key assumption made in non-FIRE models (rationality or non-rationality) are equally capable of delivering the same patterns of error predictability. I next summarize precisely how these coefficients are equalized across theories of overreaction with details provided in Appendix A.3.

Consider first the mapping to the strategic interaction model. Given $\beta_1^{DE} = \beta_1(\rho, \sigma_v, \sigma_w, \theta)$ or $\beta_1^{OC} = \beta_1(\rho, \sigma_v, \sigma_w, \alpha)$, we can match β_1 in the strategic interaction model by setting the degree of strategic substitution to be:

$$R = \frac{[\kappa_1(1+\theta) - \kappa_2](1+\theta) + (\kappa_1 - \kappa_2)\rho^2\theta^2}{(1+\theta)^2 + \rho^2\theta^2}, \text{ and}$$
$$R = \frac{(\alpha^2 - 1)\sigma_v^2\kappa_1}{(\kappa_1 - \kappa_2)[\mathbb{V}(x_t - \mathring{x}^i_{t|t-1}) + \sigma_v^2] - (\alpha^2-1)\sigma_v^2\kappa_2},$$

respectively.

Similarly, there is a mapping to diagnostic expectations. Given $\beta_1^{SI} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, R)$ or $\beta_1^{OC} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, \alpha)$, one could construct a model diagnostic expectations that delivers an equivalent β_1 by setting the degree of diagnosticity, θ, to be equal to the largest root of:

$$0 = \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} + \left(\frac{3R\kappa_1 - R\kappa_2}{\kappa_1 + R\kappa_2}\right)\theta + \left[\frac{(1+R)\kappa_1 + \rho^2 R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}\right]\theta^2, \text{ and}$$
$$0 = \frac{(\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \mathring{x}^i_{t|t-1}) + \sigma_v^2} + \left(\frac{\mathbb{V}(x_t - \mathring{x}^i_{t|t-1}) + (2\alpha^2 - 1)\sigma_v^2}{\mathbb{V}(x_t - \mathring{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta$$
$$+ \left(\frac{\mathbb{V}(x_t - \mathring{x}^i_{t|t-1}) + [\alpha^2 + \rho^2(\alpha^2-1)]\sigma_v^2}{\mathbb{V}(x_t - \mathring{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta^2,$$

respectively.

Finally, consider the mapping to overconfidence. Given $\beta_1^{DE} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, \theta)$ or $\beta_1^{SI} = \beta_1(\rho, \sigma_v^2, \sigma_w^2, R)$ one could construct a model of overconfidence by finding the α parameter such that:

$$\alpha = \sqrt{\frac{(1+\theta)[\sigma_v^2 - \theta\mathbb{V}(x_t - \mathring{x}^i_{t|t-1})]}{[(1+\theta)^2 + \rho^2\theta^2]\sigma_v^2}}, \text{ and}$$
$$\alpha = \sqrt{\frac{(1+R)\kappa_1\sigma_v^2 + R(\kappa_1 - \kappa_2)\mathbb{V}(x_t - \mathring{x}^i_{t|t-1})}{(\kappa_1 + R\kappa_2)\sigma_v^2}},$$

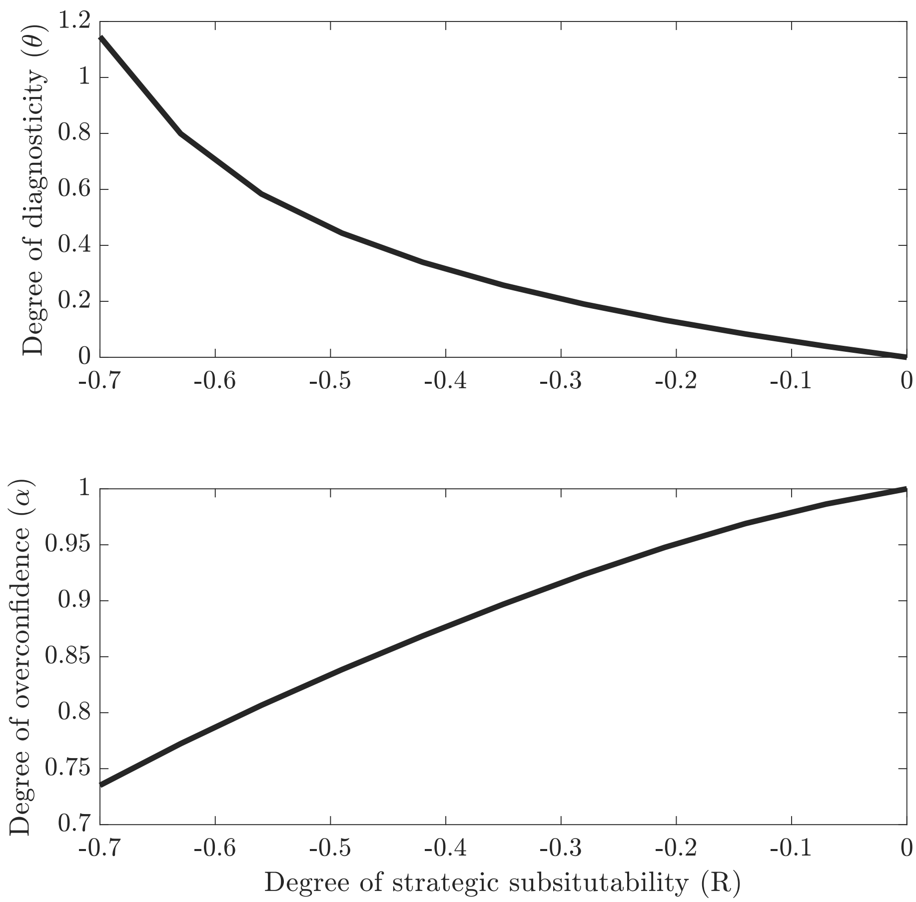
respectively.[11]

Hence, by simply assessing the regression coefficient in an errors-on-revisions re-

[11] Note that $\mathbb{V}(x_t - \mathring{x}^i_{t|t-1}) \equiv \mathring{\Psi}^i_{t|t-1}(\alpha)$ is the forecast error variance in the overconfidence model which is itself a function of α.

gression, one cannot necessarily distinguish across noisy information models of rational and non-rational expectations. The panels in Figure 3·1 plot the relationship between the θ, α, and R that are key in delivering identical β_1 coefficients.

Figure 3·1: Mapping β_1 Across Models of Overreactions



Notes: The figure plots the degree of strategic substitutability, R, that generates the same β_1 that is obtained by non-rational models of overreactions denoted by the degree of diagnosticity, θ (for diagnostic expectations) and the parameter governing perceived information precision, α (for overconfidence).

3.5 Revision Predictability

While both models can deliver identical β_1 coefficients, they are not observationally equivalent, in general. Hence, with enough data, one can discern between the two. More broadly, with enough data, we can distinguish between two subsets of non-FIRE models. I show that we can make progress on this front by merely focusing on an additional fact: the persistence of fixed-event revisions.

Beyond forecast error orthogonality, Nordhaus (1987) notes that revisions must be "informationally efficient." This requires the following condition to hold:

$$\mathbb{E}(x^i_{t|t} - x^i_{t|t-1} | \mathcal{I}^i_t) = 0.$$

In words, forecast revisions must be orthogonal to any variable residing in the forecasters information set,

$$\mathbb{E}[(x^i_{t|t} - x^i_{t|t-1})\mu] = 0 \quad \text{for } \mu \in \mathcal{I}^i_t.$$

This is akin to the error orthogonality condition which has been the focus of conventional efficiency tests. However, whereas error orthogonality can be violated for some linear noisy information rational expectations models, revision orthogonality cannot. This is an artifact of Bayesian updating in a linear setting. In such models, the forecast revision is equal to the innovation error observed when the signal is received, scaled by the optimal Kalman gain. These innovation errors are unpredictable by definition. Motivated by this insight, I present the next result.

Proposition 4. *The three models deliver distinct implications about the serial correlation of revisions. In particular:*

(i) The overconfidence model implies: $\gamma_1 = \frac{(\alpha^2-1)\sigma_v^2\hat{\kappa}}{\mathbb{V}(x_t - \hat{x}^i_{t|t-1}) + \sigma_v^2} \leq \beta_1$.

(ii) The diagnostic expectations model implies: $\gamma_1 = \frac{-\theta(1+\theta)}{(1+\theta)^2+\theta^2\rho^2} = \beta_1$.

(iii) The strategic interaction model implies: $\gamma_1 = 0$.

Proof. See Appendix A.3. □

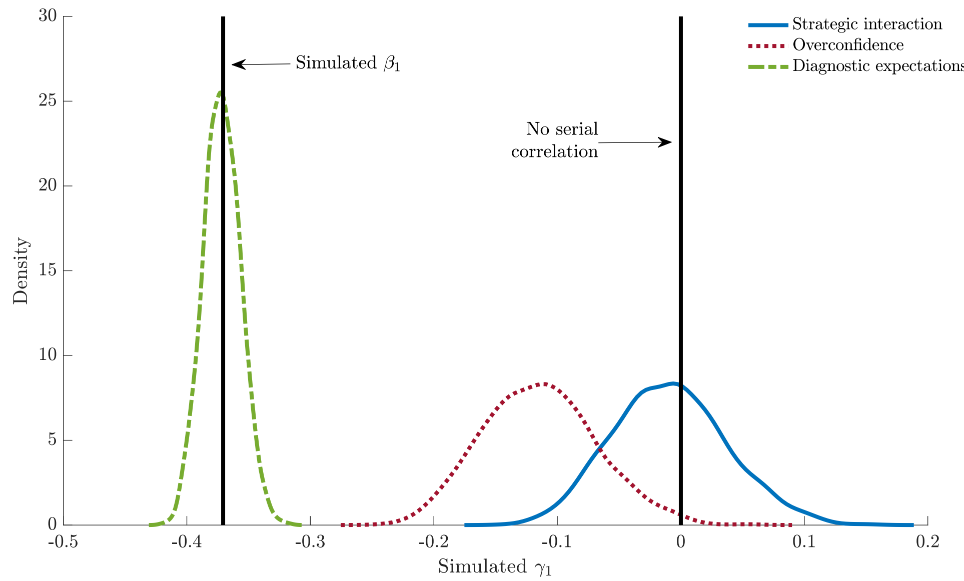
The autocorrelation of revisions is able to distinguish between all three models. In particular, the overconfidence model implies that revision persistence is negative and distinct from the errors-on-revisions coefficient. On the other hand, diagnostic expectations models imply that the persistence of revisions is identical to the errors-on-revisions coefficient. Finally, the strategic interaction model requires that revisions be serially uncorrelated.

Figure 3·2 plots a set of simulations results from all three models. I first fix the parameters of the strategic interaction model and I find the $\{\alpha, \theta\}$ that replicate β_1^{SI} in the other two models. I then compute the simulated revision persistence coefficient, γ_1 for the three models. The figure verifies that while all models can share identical β_1 coefficients, they have distinct implications for revision persistence. Consistent with Proposition 4, the overconfidence model implies that the first-order autocorrelation of revisions is smaller than the error predictability coefficient. Furthermore, diagnostic expectations requires that these two coefficients be equal. Finally, strategic interaction model requires lagged revisions to have no predictive power over current revisions. Observed overreaction on the basis of errors-on-revisions regressions is consistent with several theories, rational and non-rational alike. However, revision persistence is capable of discerning across non-FIRE models, and in the process, is more information about rationality. Based on negatively serially correlated revisions in the data, a rational model of strategic diversification can be rejected.

3.5.1 Empirical Evidence Against Strategic Interaction

Based on Table 3.2, it is clear that the data are inconsistent with serially uncorrelated revisions. Furthermore, Figure C1 visualizes the 95% confidence intervals of the error

Figure 3·2: Simulated First-Order Autocorrelation of Revisions

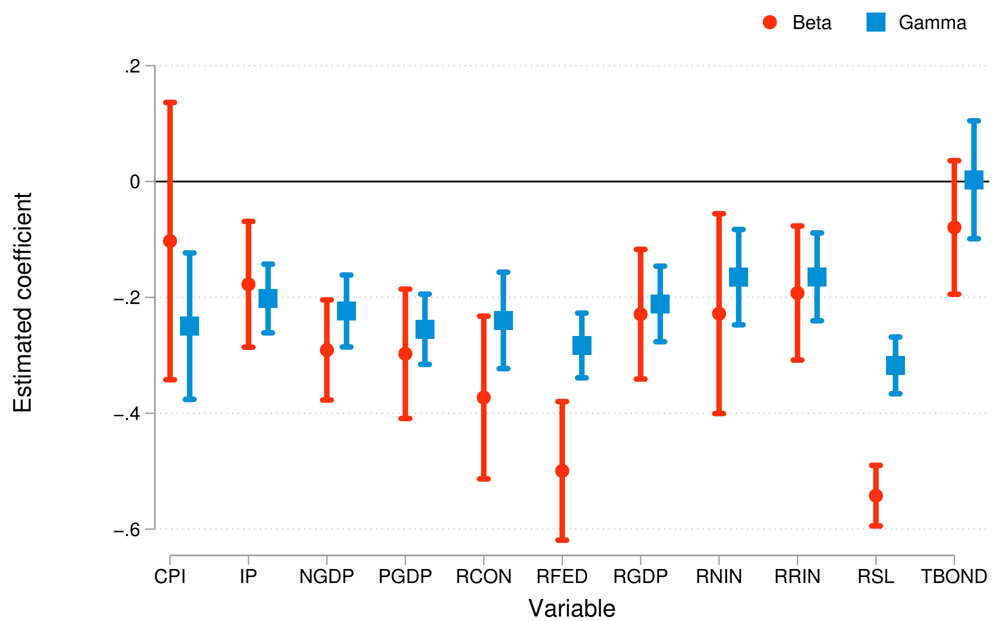


Notes: The figure plots the simulated γ_1 coefficients from the three models, each of which delivers the same simulated β_1 coefficient of -0.37. The mean first-order autocorrelations of revisions are: -0.009, -0.11, and -0.37 for the strategic interaction, overconfidence, and diagnostic expectations models, respectively.

predictability estimates alongside revision persistence.[12] The estimates of β_1 and γ_1 are mostly statistically indistinguishable from one another. This lends support to the of overreaction based on diagnostic expectations. Nonetheless, for some variables there is evidence that the error predictability coefficient is different from the autocorrelation of revisions, namely, for macroeconomic variables related to government spending. In these cases, the data support a model of overconfidence.

[12]Estimates are pooled across horizon. Appendix A.3 details horizon-by-horizon results.

Figure 3·3: Error Predictability and Revision Persistence



Notes: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. EMP- Employment, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, and RSL - Real State and Local Government Spending.

3.5.2 Broader Implications of Revision Predictability

As demonstrated in Table 3.3, non-FIRE forecasts can be generally expressed as some deviation from the conditional expectation:

$$x^i_{t|t} = \mathbb{E}(x_t|\mathcal{I}_t) + \tau^i_{t|t}$$

$$x^i_{t|t-1} = \mathbb{E}(x_t|\mathcal{I}_{t-1}) + \tau^i_{t|t-1},$$

where $\tau^i_{t|t}$ and $\tau^i_{t|t-1}$ describe the nature of the deviation from the minimum mean square forecast.[13] These deviations, which are either due to a strategic motive or a behavioral bias, can generate error predictability. However, models featuring strategic motives (i.e. alterations to the standard mean square error objective) tend to imply zero autocovariance of revisions, implying:

$$\text{Cov}(\tau^i_{t|t} - \tau^i_{t|t-1}, \tau^i_{t|t-1} - \tau^i_{t|t-2}) = 0.$$

Examples of such models include those incorporating strategic interaction, heterogeneous priors, and asymmetric attention.[14]

Notably, models featuring smoothing motives (Scotese, 1994) and noisy memory (Afrouzi et al., 2020; Azeredo da Silvera & Woodford, 2019) are an exception to this condition. In the case of smoothing motives, the forecast revision is explicitly incorporated in the forecaster's objective. As a result, the optimal reported forecast is a linear combination of the conditional expectation and the forecast revision. The canonical model featuring smoothing motives will imply a counterfactually positive autocorrelation of revisions. Nonetheless, as shown in Appendix A.3, a model of multi-horizon forecasts and a smoothing motive with respect to the long term forecast, can generate negatively serially correlated revisions.

Moreover, models featuring noisy memory can successfully deliver serially correlated revisions because forecasters optimize over the set of past signal realizations.[15] In particular, forecasters face a cost to retrieve past information. Intuitively, the presence of such a cost allows for information sets across adjacent periods to differ, thereby admitting serially correlated revisions. These two theories serve as exceptions

[13]The benchmark noisy information model is a special case where $\tau^i_{t|t} = \tau^i_{t|t-1} = 0$.

[14]See Appendix A.3 for a discussion of alternative models.

[15]Such models imply that the scope for serially correlated revisions is decreasing in the underlying persistence of the variable in question. In particular, when $\rho = 1$, these models imply zero autocorrelation of revisions.

since they are both rational models that are able to generate non-zero revision persistence. Finding ways to further discern between these models and those featuring non-rational expectations can be a promising area of future research (see, for instance D'Haultfoeuielle, Gaillac, and Maruel (2020)).

In general, theories that invoke Bayesian updating require informational efficiency despite the modified incentives to optimally deviate from the conditional expectation. Besides models in which forecasters strategically track their path of revisions or optimally choose their information sets each period, these models preclude revisions from being serially correlated. The class of models featuring behavioral biases, on the other hand, are generally better able to produce persistence in fixed-event revisions. Besides the overconfidence and diagnostic expectations models considered here, other theories that can make sense of this empirical fact include, for instance, models of relative overconfidence (Broer & Kohlhas, 2019) and natural expectations (Fuster, Laibson, & Mendel, 2010).

3.6 Conclusion

In this paper, I argue that the popular errors-on-revisions coefficient used in the expectations formation literature is insufficient to motivate departures from rationality. By way of example, I demonstrate that two prominent models of non-rational expectations can deliver the same errors-on-revisions coefficient as in a rational strategic interaction model. Motivated by this finding, I offer a new fact for further discerning across non-FIRE models which requires projecting revisions on their past values. Using data from the Survey of Professional Forecasters, I find evidence favoring diagnostic expectations and overconfidence over strategic interaction. To make progress on understanding how different economic actors formulate their expectations, it is necessary to have a rich set of theories. Conditional on a set of existing theories, however, it is important to assess the ways in which the models the fit the data. Studying the time series properties of updates made to fixed-event forecasts serves as a powerful way to discerning across existing non-FIRE models.

Appendix A

Appendix

A.1 Appendix for Chapter 1

A.1.1 Empirics

This section provides summary statistics of the data used in Section 2 of the main text. The section also includes further details on the JIT adopters data obtained, the weather regression results, and industry-level results.

Sample Construction

Table A1: Compustat Summary Statistics

	Mean	Median	Standard Deviation	25%	75%
Employment growth	0.005	0.005	0.210	-0.075	0.094
Inventory-to-sales	0.190	0.157	0.244	0.103	0.231
Inventory investment rate	0.035	0.035	0.333	-0.104	0.180
Log sales	4.881	4.769	2.092	3.369	6.292
Sales growth	0.065	0.057	0.261	-0.049	0.167
Log cash-to-assets	-2.533	-2.254	1.546	-3.524	-1.338
Log inventories	2.982	2.885	2.024	1.576	4.348
Log sales per worker	5.093	5.063	0.784	4.545	5.596
Cash-to-assets growth	0.025	0.116	0.868	-0.333	0.360
Log employment	-0.213	-0.330	1.899	-1.635	1.082
Inventory-to-sales growth	-0.018	-0.013	0.305	-0.148	0.120

Notes: The table reports summary statistics for the relevant variables in estimation in the main text. The sample is constructed from Compustat Fundamentals Annual files for 1980-2018. Sample consists of 5,099 unique firms.

I make use of Compustat Fundamentals Annual data from 1980-2018. Upon downloading this data, I keep only manufacturing firms (NAICS 31-33). In addition, I drop firm years in which acquisitions exceed 5% of total assets (to avoid influence of large mergers). To mitigate for any measurement error, I keep only those firms with non-missing and positive book value of assets, number of employees, total inventories, and sales. In addition, I keep only firms that exist in the data for at least two years. My final sample consists of 5,099 unique firms. Table A1 reports summary statistics for the variables used.

Adopters Dataset

The data for adopters was kindly provided to me by William Wempe from his joint work with Michael Kinney. Xiaodan Gao also provided me with updated data. These data include the years in which a Compustat manufacturer was identified to be a JIT adopter (via Form 10-K filings, press releases, among other communications. See Kinney and Wempe (2002) and Gao (2018) for further details). After verifying these data, I conduct an search of my own and identify an additional 18 firms (reported in Table A2). After linking these identified firm-years to those in my Compustat dataset, I identify a total of 185 adopters in the manufacturing sector. Figure A1 plots the empirical CDF of the adopters in my sample over time.

Local Weather Events

I consider a number of weather events reported by NOAA from 1980-2018. These events are reported at the county level. I keep only weather events that caused at least $1 million in property damage in a given county and link these local disasters to firm headquarter zip codes.

Using the links provided in Barrot and Sauvagnat (2016), I first map the county-level weather events to firms' headquarter zip codes in Compustat. Following this,

Table A2: Additional JIT Adopters

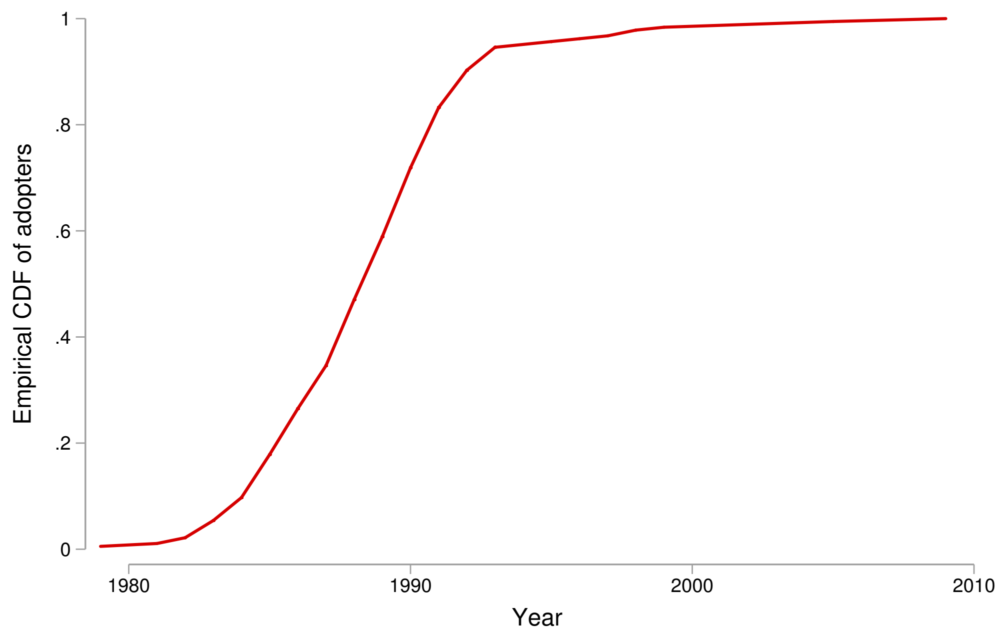
Firm	Compustat `gvkey`
Ford Motors	4839
General Motors	5073
Dell	14489
Motorola	7585
NCR	7648
Sunrise Medical	10185
Tellelabs	10420
Van Dorn Co	11101
Donnelly Corp	14462
Tuscarora	14578
Selectron	17110
Honeywell Inc	5693
ADC Telecommunications	1013
Sunbeam	1278
Boeing	2285
Campbell	2663
Cascade Corporation	2802
Caterpillar	2817

Notes: The table reports the additional JIT adopters that were added to the original set of adopters.

I link firms based on their customer-supplier relationships.[1] In the end, I have a dataset of Compustat firm i in industry j with supplier k in year t. I consider weather disasters that hit supplier k's headquarters. The idea is that customer firm i, if it is an adopter, should see a sharper decline in sales growth when its supplier k's headquarters experiences an unexpected weather event. The regression I run is

[1]This is based on a regulation requiring firms to disclose customers representing more than 10% of total reported sales (Financial Accounting Standard Board regulation No. 131).

Figure A1: Adopters, by Year



Notes: The figure plots the empirical cumulative density function for JIT adoption in the sample.

as follows

$$\text{salegrowth}_{ijt} = \vartheta_1 \text{adopter}_{ijt} + \vartheta_2 \log\left(\sum_k \text{disaster}_{ijkt}\right) + \vartheta_3 \left[\text{adopter}_{ijt} \times \log\left(\sum_k \text{disaster}_{ijkt}\right)\right] + \delta_i + \delta_{jt} + v_{ijt}$$

The coefficient of interest is ϑ_3 which describes the interaction between a JIT adoption indicator and the total number of weather events hitting suppliers of a given firm i. The results are reported in Table 1.4 in the main text.

Industry Results

The facts presented in Section 2 are robust to aggregation. Below, I provide evidence that these patterns hold at the four-digit NAICS level.

Table A3: Industry-Level Growth Regressions (Five-Year)

Change	Inventory-to-sales	Sales
Change in adopt share	-0.081**	0.207***
	(0.035)	(0.069)
Fixed effects	Industry, Year	Industry, Year
Observations	2,159	2,159

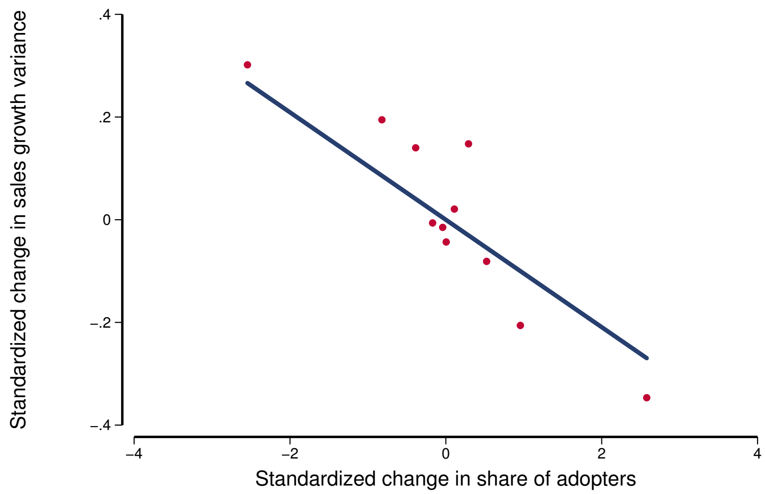
Notes: The table reports panel regression results from Compustat Annual Fundamentals of manufacturing firms (NAICS 31-33). The dependent variables are five-year changes in logs of inventory-to-sales and sales. The regressor of interest is the five-year change in share of adopters within a given industry. All variables are standardized, industry and year fixed effects are specified, and standard errors are clustered at the industry level. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

I begin by estimating

$$\Delta y_{jt} = \gamma \Delta \text{adoptshare}_{jt} + \delta_j + \delta_t + \nu_{jt} \tag{A.1}$$

where Δy_{jt} refers to the five-year difference in a given outcome for industry j in year t. I consider the log the inventory-to-sales ratio and log sales as the outcomes of interest. The regressor, $\Delta\text{adoptshare}_{jt}$, denotes the five-year difference in the share of JIT firms residing in a given industry in year t. Table A3 reports the results. For inventory-to-sales growth and sales growth, a one standard deviation increase in the difference in the share of adopters is associated with a change of -8% and 21% of one standard deviation in the outcomes, respectively.

Figure A2: Smoother Outcomes in Industries with More Adoption



Notes: The figure displays a binned scatter plot of the normalized five-year difference in sales growth variance against the normalized five-year difference in the share of adopters within an industry.

Furthermore, JIT adoption is associated with less variability in outcomes. Figure A2 plots the change in a measure of sales growth variance within an industry against the change in the share of adopters within that industry. The slope of this line is estimated from the following regression:

$$\Delta y_{jt} = \gamma \Delta \text{adoptshare}_{jt} + \delta_j + \delta_t + \nu_{jt} \tag{A.2}$$

The outcome variables specified are five-year differences in the variance of inventory-to-sales growth, employment growth, and sales growth. Table A4 reports the results. A one standard deviation increase in the share of JIT firms within an industry is associated with a 5%-11% standard deviation reduction in variability of firm outcomes within that industry.

Table A4: Industry Level Variance Regressions (Five-Year)

Change	Inventory-to-sales	Employment	Sales
Change in adopt share	-0.053**	-0.080***	-0.112***
	(0.022)	(0.028)	(0.018)
Fixed effects	Industry, Year	Industry, Year	Industry, Year
Observations	2,159	2,159	2,159

Notes: The table reports panel regression results from Compustat Annual Fundamentals of manufacturing firms (NAICS 31-33). The dependent variables are five-year changes in interquartile range of (1) inventory-to-sales growth, (2) employment growth, and (3) sales growth. The regressor of interest is the five-year change in share of adopters within a given industry over the same horizon. All variables are standardized. Industry and year fixed effects are also specified. Standard errors are clustered at the industry level. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

At the same time, industries with more adopters exhibit heightened cyclicality. I show this by running a regression similar to (1.2) at the industry level. Table A5 reports the results. In an industry comprised of approximately 10% of JIT adopters, an empirically relevant share, industry sales growth tends to rise by 0.7% above the reported baseline.

Table A5: JIT Adoption and Cyclicality (Industry-Level)

	Sales growth	Sales growth
GDP growth	1.854***	
	(0.626)	
Adopter share × GDP growth	6.928**	7.568*
	(3.384)	(3.851)
Fixed Effect	Industry	Industry, Year
Observations	3,044	3,044

Notes: The table reports regression results from Compustat Annual Fundamentals of manufacturing firms (NAICS 31-33). The dependent variable is sales growth and the independent variable of interest is the interaction between the share of JIT firms in an industry and GDP growth. Control variables include logs of sales per worker, firm size, cash-to-assets, and inventory stock, as well as the share of adopters. Standard errors are clustered at the industry level. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% signficance.

Regarding sensitivity to weather events, I run a regression similar to (1.3), however, I achieve broader coverage by instead looking at disasters originating at the customer firm's zip code. To this end, I do away with the customer-supplier links and focus on industry-level evidence by running the following regression:

$$\text{salegrowth}_{jt} = \psi_1 \text{adoptshare}_{jt} + \psi_2 \sum_i \sum_k \text{disaster}_{ijkt} + \psi_3 \big[\text{adoptshare}_{jt} \times \sum_i \sum_k \text{disaster}_{ijkt}\big] + \delta_j + \delta_t + v_{jt}$$

Rather than focusing on customer-supplier links, this regression provides evidence that industries with a larger share of JIT adopters appear to be more exposed to local weather disasters. Table A6 reports the results.

Table A6: JIT Adoption and Local Disasters (Industry-Level)

	Sales growth
Adopt share	0.011
	(0.017)
Total disasters within industry	-0.0004*
	(0.0002)
Adopt share × Total disasters within industry	-0.007**
	(0.003)
Fixed Effects	Industry, Year
Observations	3,045

Notes: The table reports weather event regressions from a sample of Compustat manufacturing firms (NAICS 31-33). The dependent variable is sales growth and the independent variable of interest is the interaction between the share of JIT firms in an industry and total disasters within the industry. Standard errors are clustered at the firm and industry levels respecitvely. *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

A.1.2 Model

Order Threshold for Final Goods Firm

The firm's problem delivers a threshold rule for placing an order. In particular, a firm places an order if and only if the order cost draw is lower than a threshold order cost: $\xi < \xi^*(z,s,a)$ where

$$\widetilde{\xi}(z,s,a) = \frac{pqs + V^*(z,s,a) - V^P(z,s,a)}{\phi} \tag{A.3}$$

and

$$\xi^*(z,s,a) = \min\Big(\max\big(0, \widetilde{\xi}(z,s,a)\big), \overline{\xi}\Big) \tag{A.4}$$

Intermediate Goods Firm

The intermediate good firm's problem is

$$\max_{K,L}\left(qK^\alpha L^{1-\alpha} - RK - wL\right)$$

One can solve for K L, and q analytically. In particular, the relative price of the intermediate good is

$$q = \left(\frac{1}{\beta\alpha}\right)^\alpha \left(\frac{w}{1-\alpha}\right)^{1-\alpha}$$

Equilibrium

An equilibrium is a set of functions,

$$\left\{V^A, V^O, V^*, V^P, s^*, s', \xi^*, a', K, L, p, w, q, \Gamma_\mu\right\},$$

such that:

1. The household's first order conditions hold:

$$p = \frac{1}{C}, \quad w = \frac{\phi}{p}.$$

2. The intermediate goods firm first order conditions hold:

$$R = \alpha q \left(\frac{L}{K}\right)^{1-\alpha} \quad w = (1-\alpha) q \left(\frac{K}{L}\right)^{\alpha}.$$

3. V^A, V^O, V^*, V^P solve the final good firm's problem.

4. Market for final goods clears:

$$\begin{aligned} C = & \int\int y(z, s^*, s', a, \xi) dH(\xi^*) d\mu(z, s, a) \\ & + \int\int y(z, s, s', a, \xi)[1 - dH(\xi^*)] d\mu(z, s, a) - K. \end{aligned}$$

5. Market for orders clears:

$$O = \int\int [s^*(z, s, a, \xi) - s] dH(\xi^*) d\mu(z, s, a).$$

6. Market for labor clears:

$$\begin{aligned} N^h = & \int\int n(z, s^*, s', \xi) dH(\xi^*) d\mu(z, s, a) \\ & + \int\int n(z, s, s', a, \xi)[1 - dH(\xi^*)] d\mu(z, s, a) + \int \left[\int_0^{\xi^*(z,s,a)} \xi dH(\xi) \right] d\mu(z, s, a) \\ & + \int a'(z, s, a)[(1-a)c_s + ac_f] d\mu(z, s, a) + \frac{q(1-\alpha)}{w}^{\frac{1}{\alpha}} K. \end{aligned}$$

7. The evolution of the distribution of firms is consistent with individual decisions:

$$\begin{aligned} \Gamma_\mu(z, s, a) &= \int\int\int 1_{\mathbb{A}} d\mu(z, s, a) dH(\xi) d\Phi(\varepsilon_z) \\ \mathbb{A}(z', s', a', \xi, \varepsilon_z; \mu) &= \{(z, s, a) | s'(z, s, a, \xi; \mu) = s', \\ & \qquad z' = \rho_z z + \sigma_z \varepsilon_z, a'(z, s, a, \xi; \mu) = a'\} \\ \Phi(x) &= \mathbb{P}(\varepsilon_z \leq x). \end{aligned}$$

Numerical Solution

The model is solved using methods that are standard in the heterogeneous firms literature. The exogenous productivity process is discretized following Tauchen (1986) which allows me to express the AR(1) process for log firm productivity as a Markov process. I select $N_z = 11$ grid points for idiosyncratic productivity. Furthermore, I select $N_s = 200$ grid points for the endogenous inventory holdings state. Finally, the binary adoption state implies that the discretized model has 4,400 grid points.

I solve for the policy functions via value function iteration which is accelerated by the use of the MacQueen-Porteus error bounds (MacQueen, 1966; Porteus, 1971). This acceleration method makes use of the contraction mapping theorem to obtain bounds for the true (infinite horizon) value function. These bounds are used in order to produce a better update of the value function. The ergodic distribution of firms is obtained via nonstochastic simulation as in Young (2010). This histogram-based method overcomes sampling error issues associated with simulating individual firms in order to obtain the stationary cross-sectional distribution.

Operationally, I solve the model by initiating a guess of the final goods price, p_0. I then compute q_0 and w_0 given the guess p_0. From here, I solve the firm's problem via value function iteration and then obtain the ergodic distribution. From the policies and ergodic distribution, I compute aggregates and the associated market clearing error using the household's optimality condition. I update the price based on this error using bisection.

A.1.3 Estimation

Simulated Method of Moments

The parameter vector to be estimated is $\theta = \big(\rho_z \;\; \sigma_z \;\; \overline{\xi}_{NA} \;\; \overline{\xi}_A \;\; c_s \;\; c_f \;\; c_m \;\; \tau\big)'$. Operationally, this requires solving my plant-level model, given θ, and simulating a panel of firms from which I compute the different moments. I define a firm to be composed of ten plants and simulate a panel of firms roughly eight times the size of the panel in Compustat. A firm is defined to be an adopter if at least one of the ten plants adopt JIT, consistent with the classification of JIT firms in my sample. I discard the first 25 years of simulated data so as to minimize the impact of initial values. I then collect the targeted empirical moments in a stacked vector $m(X)$ which comes from my Compustat sample. I next stack the model-based moments, which depend on θ, in the vector $m(\theta)$. Finally I search the parameter space to find the $\widehat{\theta}$ that minimizes the following objective

$$\min_{\theta} \big(m(\theta) - m(X)\big)' W \big(m(\theta) - m(X)\big)$$

where W is the optimal weighting matrix, defined to be the inverse of the covariance matrix of the moments. I obtain the covariance matrix via a clustered bootstrap, allowing for correlation within firms. I estimate the parameter vector via particle swarm, a standard stochastic global optimization solver.[2] [3]

The limiting distribution of the estimated parameter vector $\widehat{\theta}$ is

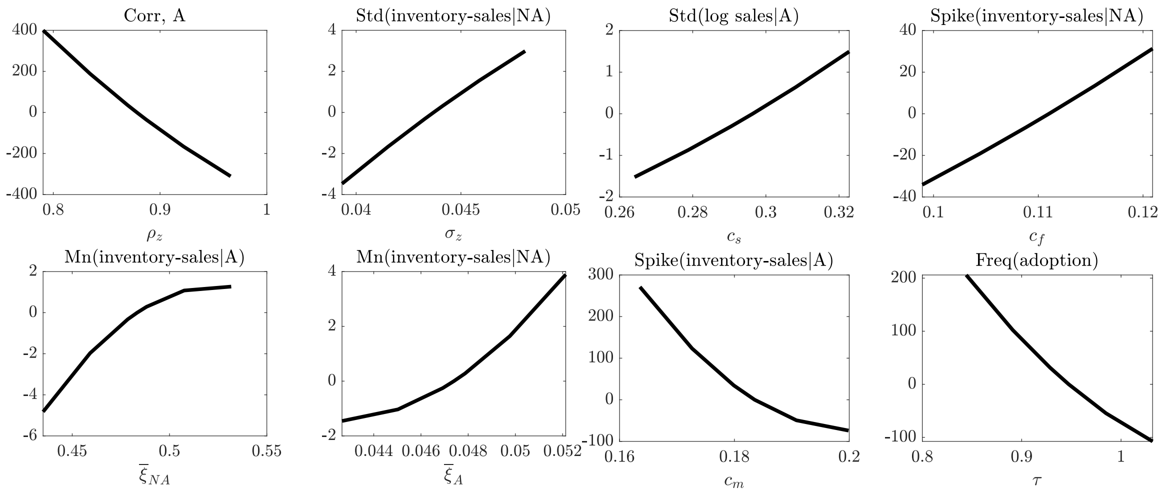
$$\sqrt{N}(\widehat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, \Sigma)$$

[2] I specify 100 particles.

[3] Considering all of the moments used in the overidentified SMM estimation, the J-test of overidentifying restrictions rejects the null hypothesis.

Figure A1: Monotonic Relationships

(a) Sales Volatility **(b)** Firm Profits



Notes: The figure plots the changes in select moments to changes in the parameters, in percentage points relative to moment at estimated parameter values.

where

$$\Sigma = \left(1 + \frac{1}{S}\right)\left[\left(\frac{\partial m(\theta)}{\partial \theta}\right)' W \left(\frac{\partial m(\theta)}{\partial \theta}\right)\right]^{-1}$$

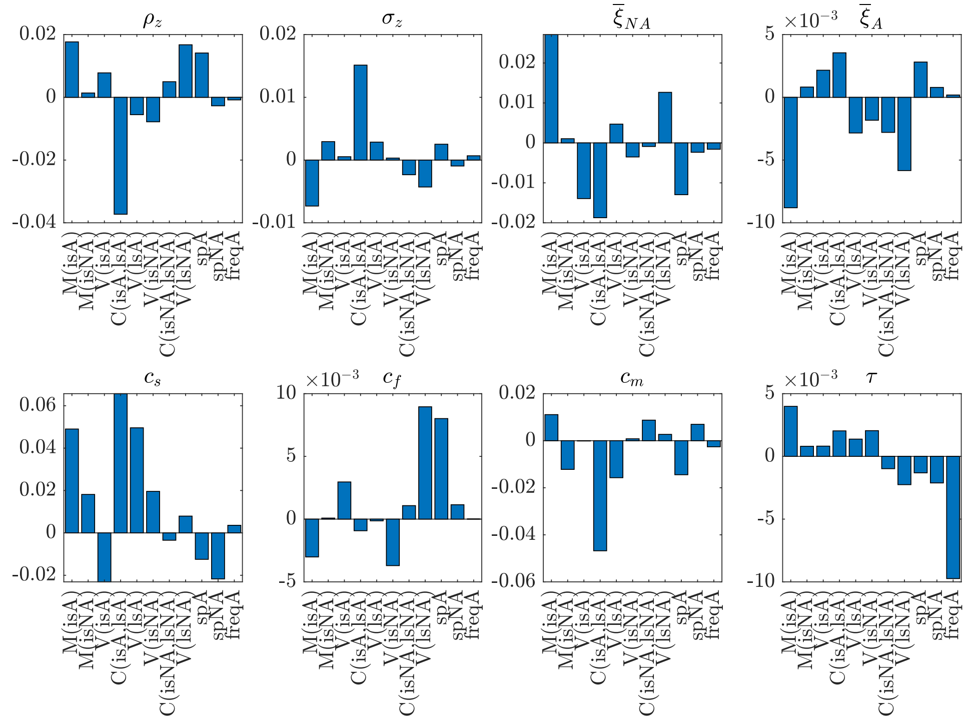
and S is the ratio of the number of observations in the simulated data to the number of observations in the sample.[4] I obtain the standard errors by computing the secant approximation to the partial derivative of the simulated moment vector with respect to the parameter vector. Given the discontinuities induced by the discretized state space, I select ϵ to be a step size of 1.0%.

Identification

The 11 moments jointly determine the eight parameters that reside in vector θ. To supplement the discussion on monotone relationships from the main text, Figure A1 reports the monotone relationships between selected moments and parameters. Figure A2 reports the sensitivity of each of the seven parameters to changes in each of the moments. These results come from an implementation of Andrews et al. (2017).

[4] I simulate 40,000 firms, thereby setting S to be approximately 7.8.

Figure A2: Sensitivity



Notes: The figure plots sensitivity estimates as in Andrews et al. (2017). These estimates describe the changes in each of the eight parameters to a one standard deviation increase in each moment.

In particular, the sensitivity of $\widehat{\theta}$ to $m(\theta)$ is

$$\Lambda = -\left[\left(\frac{\partial m(\theta)}{\partial\theta}\right)' W \left(\frac{\partial m(\theta)}{\partial\theta}\right)\right]^{-1}\left(\frac{\partial m(\theta)}{\partial\theta}\right)' W$$

I then transform this matrix so as that the interpretation of the coefficients is the effect on each parameter of a one standard deviation change in the respective moments.

Estimating Subperiod Model

The counterfactual is estimated using data from 1980-1989. More specifically, I fix all parameters to be as in the estimated full sample adoption model, and re-estimate only the adoption costs, $\{c_s, c_f\}$. Table A1 reports the estimated parameters.

Table A1: Estimated Parameters (1980-1989 Subperiod)

Description	Parameter	Estimate
Sunk cost of adoption	c_s	0.237 (0.024)
Continuation cost of adoption	c_f	0.123 (0.0004)

Notes: The table reports the estimated parameters for the 1980-1989 subperiod (standard errors in parentheses). Parameters were estimated by targeting 11 moments. J-test of overidentifying restrictions rejects null hypothesis. Firms in the model are defined to consist of ten plants.

The parameters are estimated precisely. The continuation cost of adoption is higher at around 52% of the initial sunk cost relative to the estimated adoption model in the main text. As a result, there is less adoption persistence in the counterfactual model. The steady state of the counterfactual model has a mass of 0.09 adopters, 60% of the mass of adopters in the estimated model. Table A2 reports the model fit. Table A3 reports the full parameterization for the counterfactual model.

Alternate Counterfactual: Re-estimated Order Costs

The incentives to adopt JIT in the model are governed by adoption costs as well as order costs. As a robustness check to my counterfactual economy, I consider an alternate counterfactual in which I also re-estimate the parameters governing the order cost distributions, $\overline{\xi}_{NA}$ and $\overline{\xi}_{A}$ for the 1980s. I find that the estimated order and adoption costs are little changed from the parameterization in the original counterfactual. Table A4 details the estimated parameters and long-run aggregates of the estimated full sample JIT economy vs. this alternate counterfactual. Comparing these two models, I find that consumption-equivalent welfare rises 0.8% in the JIT model relative to this counterfactual, slightly above the 0.6% increase relative to the counterfactual detailed in the main text. Moreover, I find a comparable contraction amid the disaster: this counterfactual economy contracts by 8.4% amid the same dis-

Table A2: Model vs. Empirical Moments (1980-1989 Subperiod)

Moment	Model	Data
Mean(inventory-sales ratio|adopter)	0.176	0.150
		(0.007)
Mean(inventory-sales ratio|non-adopter)	0.208	0.213
		(0.003)
Std(inventory-sales ratio|adopter)	0.056	0.042
		(0.0004)
Corr(inventory-sales ratio, log sales|adopter)	-0.106	-0.309
		(0.001)
Std(log sales|adopter)	0.218	0.169
		(0.008)
Std(inventory-sales ratio|non-adopter)	0.066	0.070
		(0.0002)
Corr(inventory-sales ratio, log sales|non-adopter)	-0.287	-0.346
		(0.0004)
Std(log sales|non-adopter)	0.277	0.228
		(0.002)
Spike(inventory-sales|adopter)	0.095	0.070
		(0.022)
Spike(inventory-sales|non-adopter)	0.285	0.284
		(0.008)
Frequency of adoption	0.038	0.015
		(0.002)

Notes: The table reports the model-based moments and the empirical moments for the estimated 1980-1989 model. Standard errors in parentheses.

Table A3: Counterfactual Parameterization

ρ_z	σ_z	$\overline{\xi}_{NA}$	$\overline{\xi}_A$	c_s	c_f	c_m	τ
0.878	0.044	0.483	0.047	0.237	0.123	0.182	0.938

Notes: The table reports the parameterization used to define the counterfactual model.

aster as described in the main text. This amounts to a 1.1 percentage point sharper contraction in the JIT economy, consistent with the headline results.

Table A4: Alternate Counterfactual Estimation

Description	Parameter	Estimate
Order cost distribution (non-adopters)	$\overline{\xi}_{NA}$	0.479 (0.046)
Order cost distribution (adopters)	$\overline{\xi}_{A}$	0.042 (0.001)
Sunk cost of adoption	c_s	0.229 (0.043)
Continuation cost of adoption	c_f	0.130 (0.008)

Notes: The table reports the estimated parameters for the alternate counterfactual detailed above (standard errors in parentheses). Parameters were estimated by targeting 11 moments. J-test of overidentifying restrictions rejects null hypothesis. Firms in the model are defined to consist of ten plants.

A.1.4 Robustness

In this section I provide different robustness checks related to the JIT tradeoff presented in the main text. I begin by examining the sensitivity of the tradeoff to different parameter values. I then consider alternate disaster persistence specifications.

Alternate Parameterizations

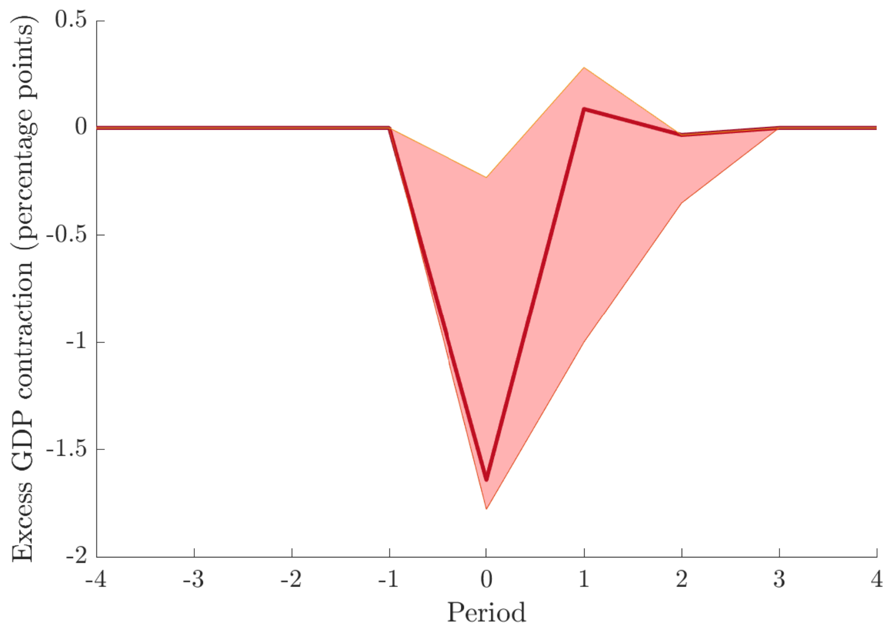
Table A1: Robustness Parameterization

Description	Parameter	Value	Value
Idiosyncratic productivity persistence	ρ_z	0.900	0.450
Idiosyncratic productivity volatility	σ_z	0.250	0.025
Order cost upper bound (non-adopters)	$\overline{\xi}_{NA}$	0.600	0.200
Order cost upper bound (adopters)	$\overline{\xi}_{A}$	0.100	0.025
Carrying cost	c_m	0.200	0.100

Notes: The table reports the alternate parameterizations chosen to compute the excess GDP contraction in the JIT economy relative to the counterfactual economy.

Table A1 reports a number of different parameter specifications. I vary all parameters in different directions with the exception of the adoption costs which trace out the frontier displayed in Figures 1·8. Figure A1 plots the gap in GDP growth amid a disaster between the estimated and counterfactual economies. The solid line depicts the figure in the main text while the shaded area captures the different tradeoffs reflected in the alternate parameterizations. Across all specifications, there is a robust negative gap indicating a sharper contraction in the estimated economy relative to the counterfactual.

Figure A1: Robustness Checks to Disaster



Notes: The figure plots the exceess GDP contraction in the estimated model relative to the counterfactual. The thick solid line refers to the estimated model parameterization used in the main text. The shaded area is obtained by considering the maximal and minimal gap across the two models in each period across the parameterizations detailed in Table A1.

Disaster Size

Table A2: GDP Contractions by Disaster Severity

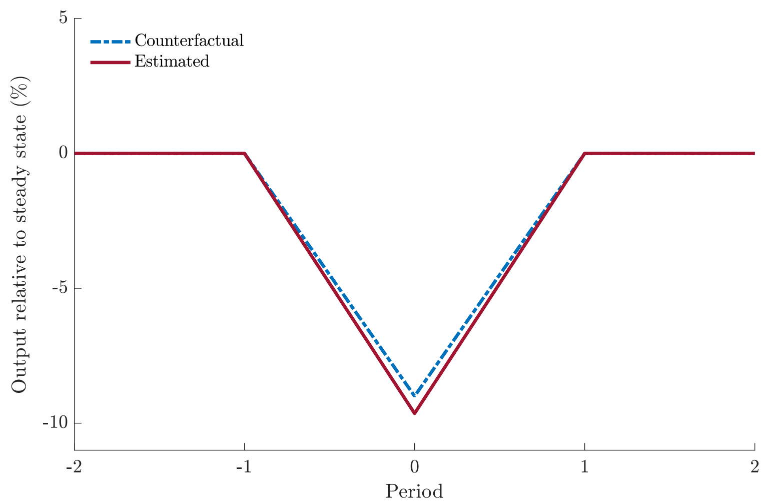
Relative shock size	Excess contraction (percentage points)
0.10	-0.189
0.25	-0.204
0.50	-0.624
0.75	-0.788
0.90	-1.328
1.00	-1.642

Notes: The table reports GDP contractions by disaster size. Column (1) reports the size of the unanticipated shock relative to the baseline shock size reported in the main text (baseline= 1.00). The second column reports the excess GDP contraction in the estimated vs. the counterfactual (in percentage points).

Table A2 reports the excess output contraction in the estimated economy relative to the counterfactual for lesser disasters to the one calibrated in the main text. Across all magnitudes, the JIT economy experiences a deeper output contraction relative to the counterfactual as indicated by the negative numbers in column 2. The deadline result is reported in the final row.

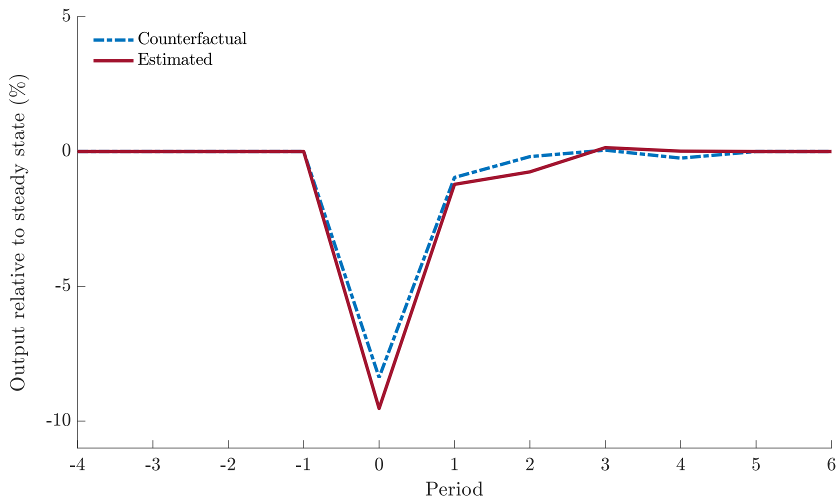
Alternate Disaster Persistence

Figure A2: One-Year Disaster



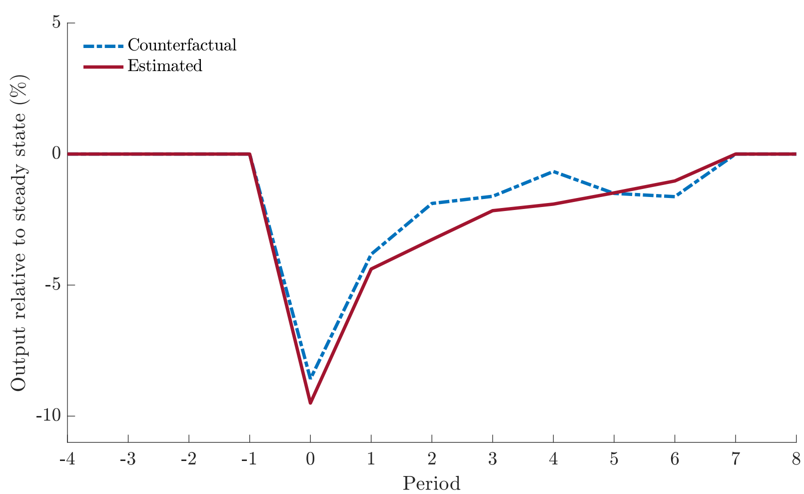
Notes: The figure plots the evolution of GDP amid a one year unanticipated disaster episode. The estimated JIT economy contracts 0.6 percentage points further than the counterfactual economy with less JIT.

Figure A3: Five-Year Disaster



Notes: The figure plots the evolution of GDP amid a five year unanticipated disaster episode. The estimated JIT economy contracts 1.5 percentage points further than the counterfactual economy with less JIT.

Figure A4: Seven-Year Disaster



Notes: The figure plots the evolution of GDP amid a seven year unanticipated disaster episode. The estimated JIT economy contracts 1 percentage point further than the counterfactual economy with less JIT.

Incorporating Anticipation

In this sub-section, I allow for there to be uncertainty as to whether the disaster occurs in period t. This uncertainty is fully resolved following t regardless of whether or not the disaster comes to pass.

Let λ denote the probability that the large aggregate shock to A is realized at time t. Recall that firms face the following problem in the production stage:

$$V^P(z, \widetilde{s}, a) = \max_{s' \in [0, \widetilde{s}]} \pi(z, \widetilde{s}, s', a) + \beta \mathbb{E}\big[V^A(z', s', a')\big]$$

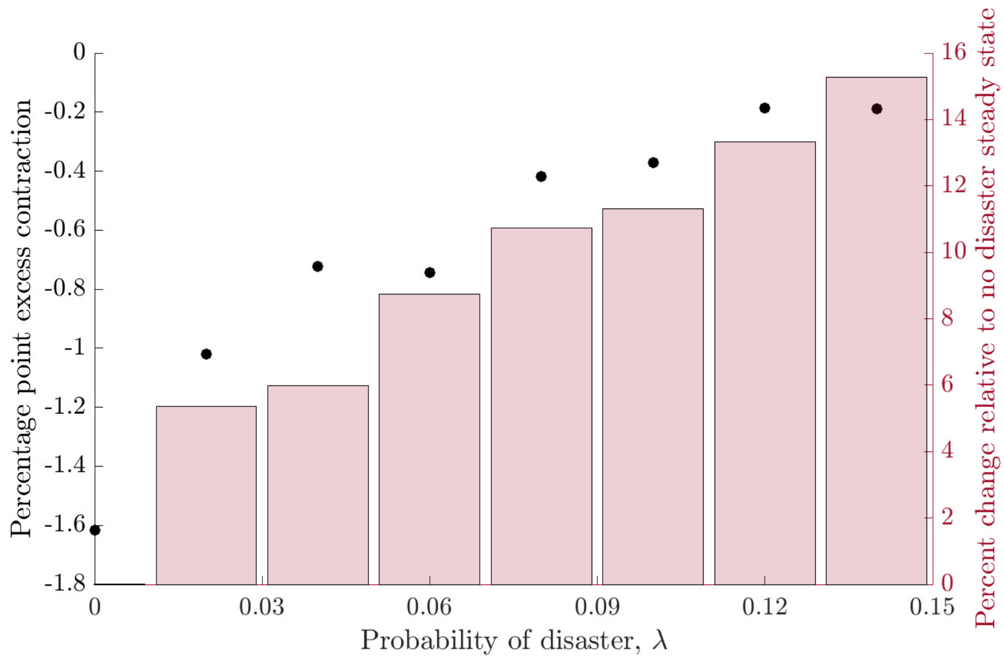
In period $t-1$, however, the expectation is not only taken across idiosyncratic productivity realizations but across the realization of the disaster as well:

$$V^A(z', s', a') = \lambda V^A_{\text{Disaster}}(z', s', a') + (1-\lambda) V^A_{SS}(z', s', a')$$

I evaluate the dynamics amid the disaster shock by numerically implementing an algorithm similar to the unanticipated case. Assuming the disaster lasts from $t = 2, \ldots, T$, I first guess a price p_0 and work backwards from T to obtain a sequence of time-indexed value and policy functions. With these in hand, I proceed to a forward step in which I push the distribution of firms forward across time utilizing the optimal policies from the backward step. From here, I compute aggregates, check for market clearing, and update the price until convergence.

Figure A5 plots two relevant quantities. On the left axis, I plot the percent increase in economy-wide inventory stocks accumulated by firms in anticipation of the disaster. Intuitively, with the prospect of a widespread disaster on the horizon, firms will optimally hold an added precautionary stock of inventories. For each probability, λ, considered I plot the percent increase in inventory holdings prior to the shock relative to the estimated baseline steady state economy. On the right axis, I plot the excess drop in output in the JIT economy relative to the counterfactual economy.

Figure A5: JIT Tradeoff Robust to Anticipation



Notes: The dots, which correspond to the left axis, display the excess output contraction (relative to the less-JIT counterfactual) for different disaster probabilties. On the right axis, the bar plot reports the percent increase in inventory holdings due to the precautionary motive that arises with anticipation of the disaster.

Importantly, despite the added precautionary inventory holdings among firms, there is still a sizable excess drop in output across the two economies, indicating that the JIT tradeoff documented in the main text is robust to the anticipation modeled here.

A.2 Appendix for Chapter 2

A.2.1 Empirics

SPF: Variable Descriptions

While the paper focuses on inflation forecasts based on the GDP deflator, in this subsection I report additional results that make use of several other variables. Before presenting these results, I provide the variable descriptions below:

- NGDP–Quarterly nominal GDP growth forecast (seasonally adjusted, annual rate). Prior to 1992, these are forecasts for nominal GNP.

- RGDP–Quarterly real GDP growth forecast (seasonally adjusted, annual rate).

- PGDP–Quarterly GDP price index growth forecast (seasonally adjusted, annual rate). From 1992 - 1995, GDP implicit deflator is used, and prior to 1992, GNP implicit deflator.

- UNEMP–Forecasts for the quarterly average unemployment rate (seasonally adjusted, average of underlying monthly levels).

- EMP–Quarterly average growth of nonfarm payroll employment (seasonally adjusted, average of underlying monthly levels).

- RNRESIN–Quarterly growth forecast of real nonresidential fixed investment. Also known as business fixed investment (seasonally adjusted, annual rate).

- RRESINV–Quarterly growth forecast of real residential fixed investment (seasonally adjusted, annual rate).

- TBILL–Quarterly forecast of average three-month Treasury bill rate (percentage points, average of underlying daily levels).

- HOUSING– Quarterly growth forecast of average housing starts (seasonally adjusted, annual rate, average of underlying monthly levels).

- CPI–Quarterly forecasts of the headline CPI inflation rate (percentage points, seasonally adjusted, annual rate). Quarterly forecasts are annualized q/q percent changes of quarterly average price index level (average of underlying monthly levels).

- RCONSUM – Quarterly growth forecast of real personal consumption expenditures (seasonally adjusted, annual rate).

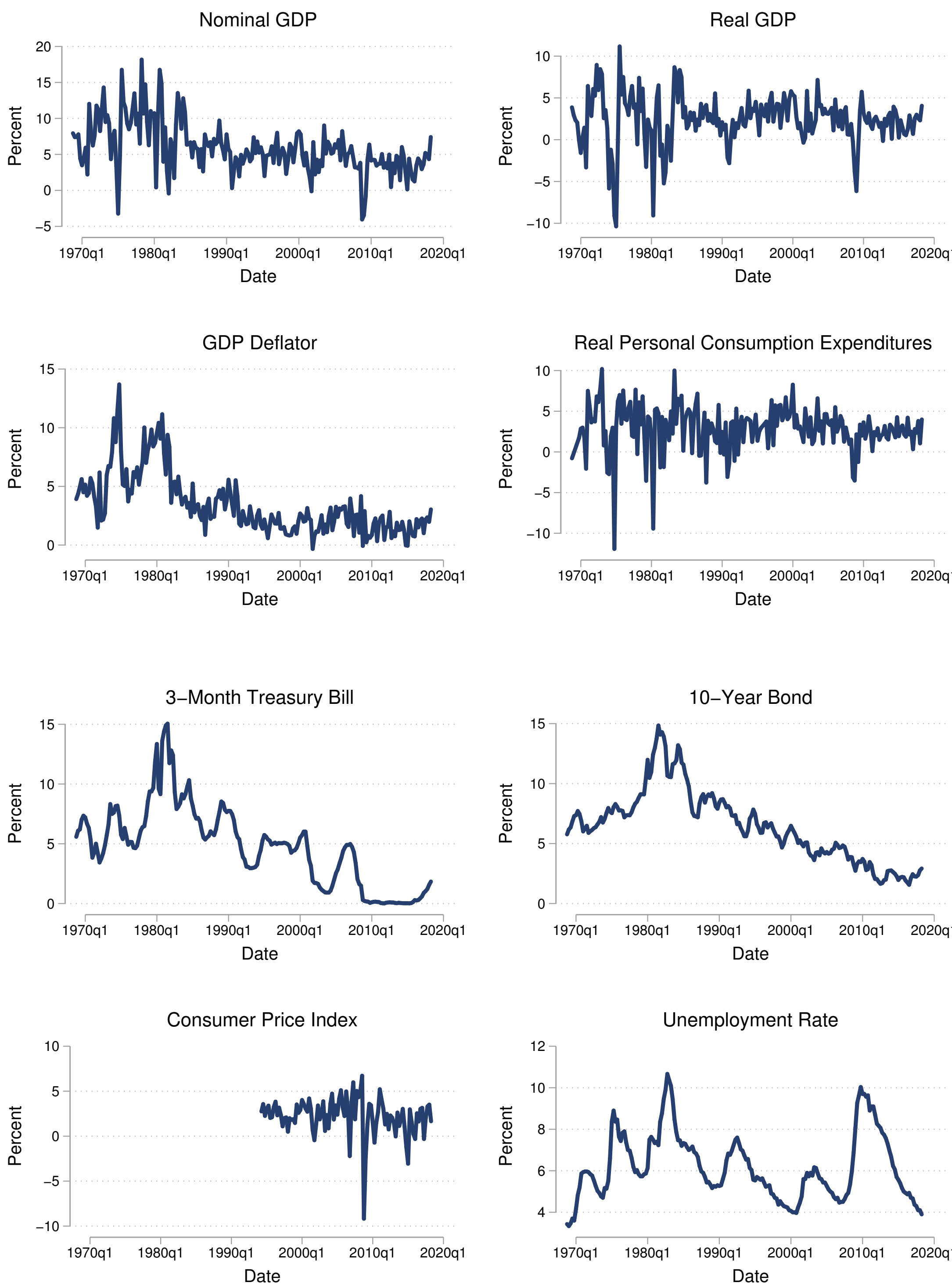
- RFEDGOV –Quarterly growth forecast of real federal government consumption and gross investment (seasonally adjusted, annual rate).

- INDPROD – Quarterly forecasters of level of the index of industrial production, seasonally adjusted (quarterly forecasts are for quarterly average of underlying monthly levels).

- TBOND–Quarterly average 10-year Treasury bond rate (percentage points, average of the underlying daily levels). the underlying daily levels

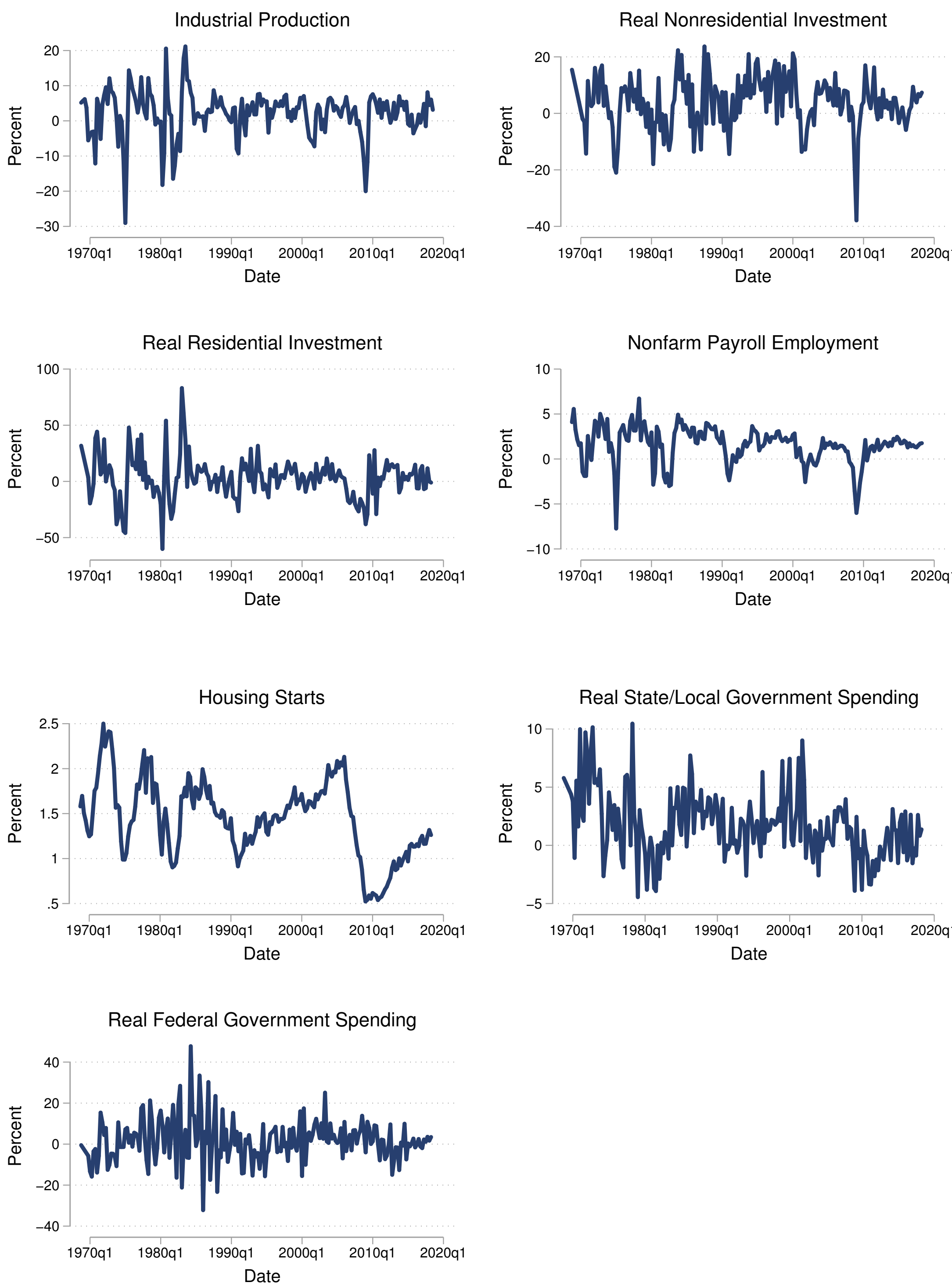
- RSLGOV–Quarterly growth forecast of real state and local government consumption and gross investment (seasonally adjusted, annual rate).

Figure B1: Real-Time Macroeconomic Time Series



Source: Survey of Professional Forecasters

Figure B2: Real-Time Macroeconomic Time Series



Source: Survey of Professional Forecasters

Modified Error Predictability Regressions (Robustness)

This subsection reports the robustness results for the modified regressions in Section 5 of the main text. Table B1 reports the results by defining 7-year year windows for the rolling standard deviations. Table B2 reports the results from a 10-year rolling window specification.

Table B1: Modified Forecast Error Predictability Regressions (7-Year Window)

	Forecast Error	
	(1)	(2)
Revision	-0.308***	-0.136*
	(0.052)	(0.071)
Revision × Quarterly		-0.194**
		(0.090)
Revision × Fundamental Volatility		0.093**
		(0.040)
Observations	58,739	58,739

Notes: The table reports estimated coefficients of forecast error predictability across two specifications. The variable *Quarterly* is equal to 0 if the SPF variable is released at a non-quarterly frequency and 1 otherwise. The *Fundamental Volatility* variable is a time-varying index of fundamental volatility constructed as described in the text. Column (1) reports a simple regression of errors-on-revisions while columns (2) includes the two proxies described. In addition to the variables reported in the table, column (2) includes the proxies individually as well as all of their interactions . Standard errors for forecaster-level regressions are as in Driscoll and Kraay (1998). Data used for estimation come from SPF (1964Q4-2018Q3). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

Table B2: Modified Forecast Error Predictability Regressions (10-Year Window)

	Forecast Error	
	(1)	(2)
Revision	-0.278***	-0.143**
	(0.053)	(0.071)
Revision × Quarterly		-0.209**
		(0.091)
Revision × Fundamental Volatility		0.118**
		(0.036)
Observations	48,827	48,827

Notes: The table reports estimated coefficients of forecast error predictability across two specifications. The variable *Quarterly* is equal to 0 if the SPF variable is released at a non-quarterly frequency and 1 otherwise. The *Fundamental Volatility* variable is a time-varying index of fundamental volatility constructed as described in the text. Column (1) reports a simple regression of errors-on-revisions while columns (2) includes the two proxies described. In addition to the variables reported in the table, column (2) includes the proxies individually as well as all of their interactions . Standard errors for forecaster-level regressions are as in Driscoll and Kraay (1998). Data used for estimation come from SPF (1964Q4-2018Q3). *** denotes 1% significance, ** denotes 5% significance, and * denotes 10% significance.

A.2.2 Model

General Linear Noisy Information RE Model

Theories of linear rational expectations are unable to account for over- and underreactions. Full information rational expectations counterfactually imply that errors are unpredictable. In addition, models of sticky information imply that forecast errors and forecast revisions are unrelated at the forecaster level.[5] In this sub-section, I focus on a general linear rational expectations model and provide analytical results about error predictability.

Consider a linear Gaussian state space model. Suppose there are n latent state variables and m exogenous signals.

$$\begin{aligned}\mathbf{s_t} &= \mathbf{A}\mathbf{s_{t-1}} + \mathbf{B}\mathbf{w_t} \\ \mathbf{z^i_t} &= \mathbf{C}\mathbf{s_t} + \mathbf{D}\mathbf{v^i_t}\end{aligned} \tag{A.5}$$

Note that $\mathbf{s_t}$ is an $n\times1$ vector, $\mathbf{A}$ is $n\times n$, $\mathbf{B}$ is $n\times n$ and $\mathbf{w_t}$ is $n\times1$. Furthermore, $\mathbf{z_t}$ is $m\times1$, $\mathbf{C}$ is $m\times n$, $\mathbf{D}$ is $m\times m$ and $\mathbf{v^i_t}$ is $m\times1$. There are no other restrictions placed on the model. In particular, $\mathbf{s_t}$ can be a vector of many different state variables, or lags of itself. $\mathbf{B}$ need not be a diagonal matrix. Furthermore, $\mathbf{z^i_t}$ can include an arbitrary finite number of observed signals. The noise vector $\mathbf{v^i_t}$ can include private or public noise.[6] From the Kalman filter, the optimal state estimate is defined as

$$\mathbf{s^i_{t|t}} = \mathbf{s^i_{t|t-1}} + \kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{A.6}$$

where κ is the (constant) Kalman gain. Since the state is unobservable, forecasters can only formulate predictions of the signals and assess the mistakes made with regard

[5] If a forecaster updates, he does so with full information rational expectations so that the subsequent error is unrelated to the revision. On the other hand, if the forecaster does not update, then there is no revision.

[6] I index this vector by i in general to allow for forecaster-specific signals.

to these observables. The optimal forecast of the signal vector $\mathbf{z^i_t}$ is

$$\mathbf{z^i_{t+1|t}} = \mathbf{z^i_{t+1|t-1}} + \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{A.7}$$

Forecast errors for the generalized linear model can be expressed as follows

$$\mathbf{z^i_{t+1}} - \mathbf{z^i_{t+1|t}} = (\mathbf{z^i_{t+1}} - \mathbf{z^i_{t+1|t-1}}) - \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{A.8}$$

Furthermore, the forecast revision is

$$\mathbf{z^i_{t+1|t}} - \mathbf{z^i_{t+1|t-1}} = \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}}) \tag{A.9}$$

Using these expressions, one can derive the two testable implications presented in the main text.

Proposition 5. *The generalized linear model implies the following:*

(i) $\beta_1 = 0$

(ii) $\alpha_1 = \mathbf{CA(I - C}\kappa)(\kappa\mathbf{C})^{-1}(\mathbf{CA})^{-1} > 0$

Proof. We have the following expressions for forecast errors and revisions, respectively:

$$\mathbf{FE^i} = \mathbf{CA(I} - \kappa\mathbf{C})(\mathbf{s_t} - \mathbf{s^i_{t|t-1}}) + \mathbf{CBw_{t+1}} + \mathbf{Dv^i_{t+1}} - \mathbf{CA}\kappa\mathbf{Dv^i_t}$$

$$\mathbf{FR^i} = \mathbf{CA}\kappa\mathbf{Dv^i_t} + \mathbf{CA}\kappa\mathbf{C}(\mathbf{s_t} - \mathbf{s^i_{t|t-1}})$$

Then,

(a) $\beta_1 \propto \mathrm{Cov}(FE^i, FR^i) = \mathbf{CA(I} - \kappa\mathbf{C})\boldsymbol{\Psi}(\mathbf{CAKC})' - \mathbf{CAK}(\mathbf{Dv^i_t v^i_t D})(\mathbf{CA}\kappa)'$ where $\boldsymbol{\Psi}$ denotes the state estimation error variance. This becomes

$$\begin{aligned} \beta_1 \quad &\propto \quad CA(I - \kappa C)\Psi(CA\kappa C)' - CAK(Dv^i_t v^i_t D)(CA\kappa)' \\ &= CA\Big\{(I - \kappa C)\Psi C - \kappa D v^i_t v^i_t D\Big\}(CA\kappa)' \\ \beta_1 \quad &= 0 \end{aligned}$$

because the term in brackets is zero by the definition of the Kalman gain.

(b) Denoting $\overline{FE}$ and $\overline{FR}$ as the cross-sectional mean of the forecast error and revision, respectively, we have

$$\alpha_1 \propto \text{Cov}(\overline{\mathbf{FE}}, \overline{\mathbf{FR}}) = CA(I - \kappa C)\overline{\Psi}(CA\kappa C)'$$

The variance of the average revision is $\text{Var}(\overline{FR}) = CA\kappa C\Psi(CA\kappa C)'$ Thus, we have

$$\alpha_1 = CA(I - \kappa C)(CA\kappa C)^{-1}$$

□

The proofs are straightforward: (a) holds given the orthogonality condition that must be satisfied at the individual-level under rational expectations. Forecast error orthogonality implies that $\mathbb{E}\big[(\mathbf{z^i_t} - \mathbf{z^i_{t|t}})\mu\big] = \mathbf{0}$ for any μ residing in the forecaster's information set.[7] Put another way, rationality implies the optimal use of information so that no variable residing in one's information set may predict the forecast error. This very general model precludes the predictability of forecast errors at the individual-level. As a result, any such linear Gaussian model with mean square loss cannot generate error predictability, regardless of the signal structure.

Moreover, (b) is a generalization of the CG result. The extent to which the mean revision predicts mean errors is determined by the Kalman gain matrix and the matrix $\mathbf{C}$ which maps the underlying state to the observed signal vector. The generalized linear model nests the CG result. Letting $\mathbf{C} = 1$, $\mathbf{D} = \sigma_v$, $\mathbf{A} = \rho$ and $\mathbf{B} = \sigma_w$, it follows that $\alpha_1 = \frac{1-\kappa}{\kappa}$. In this limiting case, one can recover an estimate of information rigidity by projecting consensus errors on consensus revisions. Importantly, the signal structure must be such that $\mathbf{C} = 1$. If, instead, the elements of $\mathbf{C}$ include additional parameters, or there is common noise in the signal vector, then it is no longer possible

[7] Similarly, there is a revision orthogonality condition implied by rationality which states that $\mathbb{E}(\mathbf{z^i_{t|t}} - \mathbf{z^i_{t|t-1}})|\mathcal{I}^i_t) = 0$. See Pesaran and Weale (2006).

to cleanly extract an the Kalman gain from a standard OLS regression.[8]

As a result, a highly generalized linear rational expectations model is unable to explain the patterns in the data.

Error Predictability Under Time-Varying Volatility

From the general nonlinear model described in the main text, the covariance of errors and revision can be signed as follows:

$$\beta_1 \propto \mathbb{C}\Bigg(\mathbf{\bar{s}_t}, \int \mathbf{\bar{s}_t}[\mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})]\mathbf{d\bar{s}_t}\Bigg) \\ - \mathbb{C}\Bigg(\int \mathbf{\bar{s}_t}\mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}})d\mathbf{\bar{s}_t}, \int \mathbf{\bar{s}_t}[\mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})]\mathbf{d\bar{s}_t}\Bigg)$$

When there are no approximation errors, error orthogonality holds and $\beta_1 = 0$. In the case of non-zero approximation errors, however, the first term is the source of observed underreaction while the second term governs the extent of overreaction. When forecast revisions are more closely related to the underlying state, then underreactions arise as the first term dominates the second. If instead, forecast revisions covary more with the current prediction than the underlying state, then overreactions result. In essence, when the approximate revision incorporates more noise than is optimally called for, then forecasters will overreact.

Similar to the approximate prediction defined above, the consensus forecast arising from approximate predictions is defined as follows

$$\alpha_1 \propto \mathbb{C}\Bigg(\mathbf{\bar{s}_t}, \int\int \mathbf{\bar{s}_t}\Big[\mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})\Big]d\mathbf{\bar{s}_t}di\Bigg) \\ - \mathbb{C}\Bigg(\int\int \mathbf{\bar{s}_t}\mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}})d\mathbf{\bar{s}_t}di, \int\int \mathbf{\bar{s}_t}\Big[\mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \mathbf{\widehat{p}}(\mathbf{\bar{s}_t}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})\Big]d\mathbf{\bar{s}_t}di\Bigg)$$

[8] See CG for a discussion of the bias in estimated information rigidities induced by public noise.

More volatile revisions increase the scope for overreaction. Upon aggregating (symmetrically) across several individual forecasts, the consensus revision will exhibit more persistence than the individual revisions. This motivates the following result

Proposition 6. *In the nonlinear noisy information model, $\alpha_1 \geq \beta_1$.*

Proof. Recall that

$$\alpha_1 = [(\widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})'(\widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})]^{-1}(\widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})'(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})$$

and

$$\beta_1 = [(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})'(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})]^{-1}(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})'(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})$$

To prove the proposition, I will show that the covariance between consensus errors and revisions is weakly greater than that for pooled errors and revisions. I will then show that the variance of the consensus revision is weakly smaller than the variance of the pooled variance.

We can express the covariance between errors and revisions as

$$\begin{aligned}\mathbb{C}(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t}, \widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}) = &\int\int(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt \\ &- \int\int(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})didt \\ &- \int\int(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})didt\end{aligned}$$

and at the consensus level

$$\begin{aligned}\mathbb{C}(\mathbf{z}_{t+h} - \widehat{\mathbf{z}}_{t+h|t}, \widehat{\mathbf{z}}_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1}) = &\int\left(\mathbf{z}_{t+h} - \int\widehat{\mathbf{z}}^i_{t+h|t}di\right) \\ &\times\left(\int[\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}]di\right)dt \\ &- \int\left(\mathbf{z}_{t+h} - \int\widehat{\mathbf{z}}^i_{t+h|t}di\right)dt \\ &- \int\int(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}_{t+h|t-1})didt\end{aligned}$$

We wish to show that the second equation is weakly greater than the first. One can note immediately that the second and third terms of both equations are equal (given the linearity of the expectations operator), and so they cancel out. The resulting inequality that we wish to verify is

$$\int \left(\mathbf{z}_{t+h} - \int \widehat{\mathbf{z}}^i_{t+h|t} di \right) \left(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}] di \right) dt$$

$$\geq \int \int (\mathbf{z}_{t+h} - \widehat{\mathbf{z}}^i_{t+h|t})(\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}) didt$$

By distributing the revision into the error on either side of the inequality, we can express each side as the sum of two terms. The first of these will drop out as we will have

$$\int \mathbf{z}_{t+h} \left(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}] di \right) dt$$

on the LHS and

$$\int \int \mathbf{z}_{t+h} (\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}) didt$$

on the RHS. Again, due to the linearity of the expectations operator, these terms cancel out. The remaining inequality is therefore:

$$
\begin{aligned}
-\int\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)\left(\int [\widehat{\mathbf{z}}^i_{t+h|t}-\widehat{\mathbf{z}}^i_{t+h|t-1}]di\right)dt &\geq -\int\int \widehat{\mathbf{z}}^i_{t+h|t}(\widehat{\mathbf{z}}^i_{t+h|t} \\
&\quad -\widehat{\mathbf{z}}^i_{t+h|t-1})didt \\
\int\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)\left(\int [\widehat{\mathbf{z}}^i_{t+h|t}-\widehat{\mathbf{z}}^i_{t+h|t-1}]di\right)dt &\leq \int\int \widehat{\mathbf{z}}^i_{t+h|t}(\widehat{\mathbf{z}}^i_{t+h|t} \\
&\quad -\widehat{\mathbf{z}}^i_{t+h|t-1})didt \\
\int\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)^2 dt-\int\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)\left(\int \widehat{\mathbf{z}}^i_{t+h|t-1}di\right)dt &\leq \int\int \widehat{\mathbf{z}}^{i2}_{t+h|t}didt \\
&\quad -\int\int \widehat{\mathbf{z}}^i_{t+h|t}\widehat{\mathbf{z}}^i_{t+h|t-1}didt \\
\int\int \widehat{\mathbf{z}}^i_{t+h|t}\widehat{\mathbf{z}}^i_{t+h|t-1}didt-\int\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)\left(\int \widehat{\mathbf{z}}^i_{t+h|t-1}di\right)dt &\leq \int\int \widehat{\mathbf{z}}^{i2}_{t+h|t}didt \\
&\quad -\int\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)^2 dt \\
\int\left[\int \widehat{\mathbf{z}}^i_{t+h|t}\widehat{\mathbf{z}}^i_{t+h|t-1}di-\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)\left(\int \widehat{\mathbf{z}}^i_{t+h|t-1}di\right)\right]dt &\leq \int\left[\int \widehat{\mathbf{z}}^{i2}_{t+h|t}di\right. \\
&\quad \left.-\left(\int \widehat{\mathbf{z}}^i_{t+h|t}di\right)^2\right]dt
\end{aligned}
$$

which is true since the terms in hard brackets on the RHS is the cross-sectional variance of the forecast whereas the term in hard brackets on the LHS is a cross-sectional covariance. Hence, the covariance of the consensus errors with consensus revisions is weakly greater than the covariance of individual-level pooled errors and revisions.

Finally, I show that the variance of the consensus revision is weakly smaller than the variance of the pooled revision. This is simpler to verify. Note that

$$
\mathbb{V}(\widehat{\mathbf{z}}^i_{t+h|t}-\widehat{\mathbf{z}}^i_{t+h|t-1}) = \int\int (\widehat{\mathbf{z}}^i_{t+h|t}-\widehat{\mathbf{z}}^i_{t+h|t-1})^2 didt - \left(\int\int [\widehat{\mathbf{z}}^i_{t+h|t}-\widehat{\mathbf{z}}^i_{t+h|t-1}]didt\right)^2
$$

and

$$
\begin{aligned}
\mathbb{V}(\widehat{\mathbf{z}}_{t+h|t}-\widehat{\mathbf{z}}_{t+h|t-1}) = &\int\left(\int [\widehat{\mathbf{z}}^i_{t+h|t}\right. \\
&\left. -\widehat{\mathbf{z}}^i_{t+h|t-1}]di\right)^2 dt - \left(\int\left[\int [\widehat{\mathbf{z}}^i_{t+h|t}-\widehat{\mathbf{z}}^i_{t+h|t-1}]di\right]dt\right)^2
\end{aligned}
$$

Once again, the second term in each of the above revision variance equations will cancel out. The resulting condition that we wish to verify is

$$\int \left(\int [\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1}] di \right)^2 dt \leq \int \int (\widehat{\mathbf{z}}^i_{t+h|t} - \widehat{\mathbf{z}}^i_{t+h|t-1})^2 di dt$$

which holds by Jensen's inequality. □

The model implies that the OLS coefficient estimated from an errors on revisions regression will be weakly greater than the analogous coefficient obtained from a pooled regression of individual forecasters. This result does depend on the presence of nonlinear dynamics. In fact, this holds in the linear setting as well (see Appendix B).[9]

[9] This need not be the case in other economic settings in which agents take actions, and their decisions are aggregated in a manner other than by taking a simple mean. Depending on the context, it is possible for the aggregate decision to also exhibit overreaction or excess volatility. Bordalo et al. (2020) provide a discussion of this.

A.2.3 Calibration

I internally calibrate four parameters: $\{\sigma_{v,1}, \sigma_{v,2}, \bar{c}_{PF,1}, \bar{c}_{PF,2}\}$ where the subscript one denotes the first variable (real GDP) and two denotes the second variable (unemployment). These parameters are calibrated to match four moments: $\{\beta_{1,1}, \beta_{1,2}, \alpha_{1,1}, \alpha_{1,2}\}$.

Based on the calibration, the two state variables must only be simulated once. Following this, I simulate a panel of forecasters who select KF or PF depending on the mean square errors and their model adoption cost draw. Following the endogenous model selection decision, I simulate a panel of errors and revisions from which I then compute model-implied errors-on-revisions coefficients. The simulated panel of forecasters is roughly 7 times the size of the size of the panel of SPF forecasters.[10].

I then collect the targeted empirical moments in a stacked vector $m(X)$ which comes from the SPF sample. I next stack the model-based moments, which depend on $\theta = (\beta_{1,1} \;\; \alpha_{1,1} \;\; \beta_{1,2} \;\; \alpha_{1,2})'$, in the vector $m(\theta)$. Finally I search the parameter space to find the $\widehat{\theta}$ that minimizes the following objective

$$\min_{\theta} \big(m(\theta) - m(X)\big)' W \big(m(\theta) - m(X)\big)$$

where the weighting matrix is set to be the identity matrix, $W = I$.

[10] I also discard the first 1,000 observations of the simulated state variables

A.2.4 Details on Particle Filtering

In this section I briefly summarize the particle filter which is a popular nonlinear filter that have been devised to handle state dynamics such as unobserved stochastic volatility.

In their seminal paper, Gordon, Salmond, and Smith (n.d.) propose the bootstrap filter which is a popular variant to the particle filter. In principle, this approach makes use to mass points (particles) to approximate the underlying filtering density, $p(s_t|Z_t^i)$. This is done by defining the set of particles and associated weights: $\chi = \{s^{(n)}, \omega^{(n)}\}_{n=1}^N$.

Importantly, the filter still follows a general predict-update algorithm. For each particle n, the forecaster propagates the estimate through the nonlinear system

$$s_t^{i,(n)} = F\big(s_{t-1}^{i(n)}, w_t\big)$$

and then updates the weight,[11]

$$\widetilde{\omega}_t^{i,(n)} = \omega_{t-1}^{i,(n)} \cdot p\big(z_t^i|s_t^{i,(n)}\big)$$

The forecaster then normalizes the weights

$$\omega_t^{i,(n)} = \frac{\widetilde{w}_t^{i,(n)}}{\sum_{n=1}^N \widetilde{\omega}_t^{i,(n)}}$$

so that they sum to one. Lastly, the nowcast of the state is computed as a weighted average of the particles

$$\widehat{s}_{t|t}^i = \sum_{n=1}^N s_t^{i,(n)} \cdot \omega_t^{i,(n)}.$$

One common issue with sequential importance sampling is that the sample of particles tends to degenerate as few particles are given most of the weight. As a result, I make

[11] The precise manner in which the weights are updated depends on choices for the importance distribution.

use of the common sequential importance resampling scheme in which I resample the particles, each with a probability equal to its weight.

Forecast errors and revisions are analogous to the formulation with the Kalman filter generalizations. The only difference is that the particle filtered estimates are not formulated by making use of the Kalman filtering equations. Nonetheless, these estimates approximate the optimal forecast.

A.3 Appendix for Chapter 3

A.3.1 Derivations

Proof of Proposition 1

Recall that $\widehat{x}^i_{t|t} = \widehat{x}^i_{t|t-1} + \widehat{\kappa}(y^i_t - \widehat{x}^i_{t|t-1})$. So the forecast error is $x_t - \widehat{x}^i_{t|t} = (1-\widehat{\kappa})(x_t - \widehat{x}^i_{t|t-1}) - \widehat{\kappa}v^i_t$, and the revision is $\widehat{x}^i_{t|t} - \widehat{x}^i_{t|t-1} = \widehat{\kappa}[(x_t - \widehat{x}^i_{t|t-1}) + v^i_t]$. Then, evaluating the population coefficient:

$$\begin{aligned}
\beta_1^{OC} &= \frac{\mathbb{C}(x_t - \widehat{x}^i_{t|t}, \widehat{x}^i_{t|t} - \widehat{x}^i_{t|t-1})}{\mathbb{V}(\widehat{x}^i_{t|t} - \widehat{x}^i_{t|t-1})} \\
&= \frac{(1-\widehat{\kappa})\widehat{\kappa}\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) - \widehat{\kappa}^2\sigma_v^2}{\widehat{\kappa}^2[\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2]} \\
&= \frac{(1-\widehat{\kappa})\mathbb{V}(x_t - x^i_{t|t-1}) - \widehat{\kappa}\sigma_v^2}{\widehat{\kappa}[\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2]}
\end{aligned}$$

Noting that the biased Kalman gain is, $\widehat{\kappa} = \frac{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1})}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \widehat{\sigma}_v^2}$, we can use this definition in the numerator to obtain:

$$\begin{aligned}
\beta_1^{OC} &= \frac{\frac{\widehat{\sigma}_v^2\mathbb{V}(x_t - \widehat{x}^i_{t|t-1})}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \widehat{\sigma}_v^2} - \frac{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1})\sigma_v^2}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \widehat{\sigma}_v^2}}{\widehat{\kappa}[\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2]} \\
&= \frac{(\widehat{\sigma}_v^2 - \sigma_v^2)\widehat{\kappa}}{\widehat{\kappa}[\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2]}
\end{aligned}$$

Finally, since $\widehat{\sigma}_v = \alpha\sigma_v$, and $\alpha \in (0,1)$, we have:

$$\beta_1^{OC} = \frac{(\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2} < 0.$$

Deriving Errors and Revisions for SI Model

In state space form, we have

$$\begin{bmatrix} x_t \\ F_t \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ \lambda\rho & (1-\lambda)\rho \end{bmatrix} \begin{bmatrix} x_{t-1} \\ F_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ \lambda \end{bmatrix} w_t$$
$$y_t^i = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ F_t \end{bmatrix} + v_t^i$$

Or more compactly,

$$\mathbf{s_t} = \mathbf{A}\mathbf{s_{t-1}} + \mathbf{B}w_t$$
$$y_t^i = \mathbf{C}\mathbf{s_t} + v_t^i$$

Invoking the Kalman filter

$$\mathbf{s^i_{t|t}} = (\mathbf{I} - \kappa\mathbf{C})\mathbf{A}\mathbf{s^i_{t-1|t-1}} + \kappa\mathbf{C}\mathbf{A}\mathbf{s_{t-1}} + \kappa\mathbf{C}\mathbf{B}w_t + \kappa v_t^i$$
$$\mathbf{s^i_{t|t-1}} = \mathbf{A}\mathbf{s^i_{t-1|t-1}}$$

which implies that

$$x^i_{t|t} = (1-\kappa_1)\rho x^i_{t-1|t-1} + \kappa_1 y_t^i$$
$$F^i_{t|t} = (\lambda-\kappa_2)\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1} + \kappa_2 y_t^i$$

and

$$x^i_{t|t-1} = \rho x^i_{t-1|t-1}$$
$$F^i_{t|t-1} = \lambda\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1}$$

where $\lambda = \frac{\kappa_1 + R\kappa_2}{1+R}$ and κ_1, κ_2 are the Kalman gains belonging to the two-dimensional column vector κ. Letting $\xi = \left[\frac{1}{1+R} \quad \frac{R}{1+R}\right]$, we have

$$\begin{aligned}
\widetilde{x}^i_{t|t} &= \xi \mathbf{s^i_{t|t}} \\
&= \frac{1}{1+R} x^i_{t|t} + \frac{R}{1+R} F^i_{t|t} \\
&= \frac{1}{1+R}\left[(1-\kappa_1)\rho x^i_{t-1|t-1} + \kappa_1 y_{it}\right] \\
&+ \frac{R}{1+R}\left[(\lambda-\kappa_2)\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1} + \kappa_2 y_{it}\right] \\
&= \frac{1-\kappa_1 + R(\lambda-\kappa_2)}{1+R} x^i_{t|t-1} + \lambda y_{it} + \frac{(1-\lambda)R\rho}{1+R} F^i_{t-1|t-1} \\
&= \frac{1-\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)R\rho}{1+R} F^i_{t-1|t-1} + \lambda y_{it}
\end{aligned}$$

and

$$\begin{aligned}
\widetilde{x}^i_{t|t-1} &= \xi \mathbf{s^i_{t|t-1}} \\
&= \frac{1}{1+R}\rho x^i_{t-1|t-1} + \frac{R}{1+R}\left[\lambda\rho x^i_{t-1|t-1} + (1-\lambda)\rho F^i_{t-1|t-1}\right] \\
&= \frac{1+R\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)\rho R}{1+R} F^i_{t-1|t-1}
\end{aligned}$$

Hence,

$$\begin{aligned}
\widetilde{x}^i_{t|t} &= \frac{1-\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)R\rho}{1+R} F^i_{t-1|t-1} + \lambda y^i_t \\
\widetilde{x}^i_{t|t-1} &= \frac{1+R\lambda}{1+R} x^i_{t|t-1} + \frac{(1-\lambda)\rho R}{1+R} F^i_{t-1|t-1}
\end{aligned}$$

Furthermore,

$$F_t = \widetilde{x}_{t|t} = (1-\lambda)\rho F_{t-1} + \lambda x_t$$

$$\widetilde{x}_{t|t-1} = \lambda\rho x_{t-1|t-1} + (1-\lambda)\rho F_{t-1}$$

From here, it follows that

$$\begin{aligned} x_t - \widetilde{x}^i_{t|t} &= x_t - \frac{1-\kappa_1+(\lambda-\kappa_2)R}{1+R}x^i_{t|t-1} - \lambda y^i_t - \frac{(1-\lambda)R\rho}{1+R}F^i_{t-1|t-1} \\ &= (1-\lambda)x_t - \frac{1-\kappa_1+(\lambda-\kappa_2)R}{1+R}x^i_{t|t-1} - \lambda v^i_t - \frac{(1-\lambda)R\rho}{1+R}F^i_{t-1|t-1} \\ &= (1-\lambda)x_t - \frac{1-\lambda}{1+R}x^i_{t|t-1} - \frac{(1-\lambda)R\rho}{1+R}F^i_{t-1|t-1} - \lambda v^i_t \\ &= (1-\lambda)\left[x_t - \frac{1}{1+R}x^i_{t|t-1} - \frac{R}{1+R}\rho F^i_{t-1|t-1}\right] - \lambda v^i_t \end{aligned}$$

and the forecast revision is

$$\begin{aligned} \widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} &= \frac{(1-\kappa_1)+(\lambda-\kappa_2)R-(1+R\lambda)}{1+R}x^i_{t|t-1} + \lambda x_t + \lambda v^i_t \\ &= -\frac{\kappa_1-\kappa_2 R}{1+R}x^i_{t|t-1} + \lambda x_t + \lambda v^i_t \\ &= \lambda(x_t - x^i_{t|t-1} + v^i_t) \end{aligned}$$

Proof of Proposition 2

$$\begin{aligned}
\beta_1^{SI} &= \frac{\text{Cov}(x_t - \widetilde{x}^i_{t|t}, \widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1})}{\text{Var}(\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1})} \\
&= \frac{(1-\lambda)\lambda\text{Cov}(x_t, x_t - x^i_{t|t-1}) - \lambda^2\sigma_v^2}{\lambda^2\left[\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2\right]} \\
&= \frac{1-\lambda}{\lambda}\frac{\text{Var}(x_t) - \text{Cov}(x_t, x^i_{t|t-1})}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - \frac{\sigma_v^2}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} \\
&= \frac{1-\lambda}{\lambda}\frac{\frac{\sigma_w^2}{1-\rho^2} - \frac{\rho^2\kappa_1 \cdot \frac{\sigma_w^2}{1-\rho^2}}{1-(1-\kappa_1)\rho^2}}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - (1-\kappa_1) \\
&= \frac{1-\lambda}{\lambda}\frac{\frac{\sigma_w^2}{1-\rho^2}\frac{1-\rho^2}{1-(1-G)\rho^2}}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - (1-\kappa_1) \\
&= \frac{1-\lambda}{\lambda}\frac{\text{Var}(x_t - x^i_{t|t-1})}{\text{Var}(x_t - x^i_{t|t-1}) + \sigma_v^2} - (1-\kappa_1) \\
&= \frac{1-\lambda}{\lambda}\kappa_1 - (1-\kappa_1) \\
&= \frac{1-\kappa_1 + R(1-\kappa_2)}{\kappa_1 + R\kappa_2}\kappa_1 - (1-\kappa_1) \\
&= \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}
\end{aligned}$$

Proof of Proposition 3

Each model delivers a coefficient of error predictability that is a function of parameters θ, R, or α.

$$\beta_1^{DE} = \frac{-\theta(1+\theta)}{(1+\theta)^2 + \rho^2\theta^2}, \quad \beta_1^{SI} = \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}, \quad \beta_1^{OC} = \frac{(\alpha^2 - 1)\sigma_v^2}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2}.$$

To show how the three models can match the identical error predictability coefficient, β_1, I consider the three pairs of models and draw the implications out in either direction to establish equality.

DE → SI

Given $\{\rho, \sigma_w, \sigma_v, \theta\}$ we set $\beta_1^{DE} = \beta_1^{SI}$ and solve for R:

$$\begin{aligned}
\frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} &= \beta_1^{DE} \\
R(\kappa_1 - \kappa_2) &= \beta_1^{DE}(\kappa_1 + R\kappa_2) \\
R(\kappa_1 - \kappa_2) &= \beta_1^{DE}\kappa_1 + \beta_1^{DE}R\kappa_2 \\
R(\kappa_1 - \kappa_2) - \beta_1^{DE}R\kappa_2 &= \beta_1^{DE}\kappa_1 \\
R &= \frac{\beta_1^{DE}\kappa_1}{\kappa_1 - (1 + \beta_1^{DE})\kappa_2}
\end{aligned}$$

which, from the definition of β_1^{DE}, is equal to

$$R = \frac{[\kappa_1(1+\theta) - \kappa_2](1+\theta) + (\kappa_1 - \kappa_2)\rho^2\theta^2}{(1+\theta)^2 + \rho^2\theta^2}.$$

SI→DE

Given $\{\rho, \sigma_w, \sigma_v, R\}$ we solve for θ by setting $\beta_1^{DE} = \beta_1^{SI}$

$$\begin{aligned}
\frac{-\theta(1+\theta)}{(1+\theta)^2 + \rho^2\theta^2} &= \beta_1^{SI} \\
-\theta(1+\theta) &= \beta_1^{SI}(1+\theta)^2 + \beta_1^{SI}\rho^2\theta^2 \\
-\theta - \theta^2 &= \beta_1^{SI}(1 + 2\theta + \theta^2) + \beta_1^{SI}\rho^2\theta^2 \\
0 &= \beta_1^{SI} + 2\beta_1^{SI}\theta + \theta + \beta_1^{SI}\theta^2 + \theta^2 + \beta_1^{SI}\rho^2\theta^2 \\
0 &= \beta_1^{SI} + (2\beta_1^{SI} + 1)\theta + [1 + (1+\rho^2)\beta_1^{SI}]\theta^2.
\end{aligned}$$

From the definition of β_1^{SI}, we can express the above quadratic as:

$$\left[\frac{(1+R)\kappa_1 + \rho^2 R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2}\right]\theta^2 + \left(\frac{3R\kappa_1 - R\kappa_2}{\kappa_1 + R\kappa_2}\right)\theta + \frac{R(\kappa_1 - \kappa_2)}{\kappa_1 + R\kappa_2} = 0$$

DE→OC

Given $\{\rho, \sigma_w, \sigma_v, \theta\}$, we solve for α by setting $\beta_1^{DE} = \beta_1^{OC}$

$$
\begin{aligned}
\frac{(\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2} &= \beta_1^{DE} \\
\alpha^2 &= \frac{\beta_1^{DE}}{\sigma_v^2}[\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2] + 1 \\
\alpha^2 - \frac{\beta_1^{DE}}{\sigma_v^2}\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) &= \beta_1^{DE} + 1.
\end{aligned}
$$

from the definition of β_1^{DE}, the above equality can be expressed as:

$$
\alpha = \sqrt{\frac{(1+\theta)[\sigma_v^2 - \theta\mathbb{V}(x_t - \widehat{x}^i_{t|t-1})]}{[(1+\theta)^2 + \rho^2\theta^2]\sigma_v^2}}
$$

Note that due to the recursive nature of the overconfidence model, $\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) \equiv \widehat{\Psi}^i_{t|t-1}(\alpha)$ is itself a function of α. As a result, there is no closed form expression that characterizes the mapping from θ to α.

OC→DE

Similar to the result relating SI→DE, we begin with:

$$
0 = \beta_1^{OC} + (2\beta_1^{OC} + 1)\theta + [1 + (1+\rho^2)\beta_1^{OC}]\theta^2.
$$

from the definition of β_1^{OC}, we obtain

$$
\begin{aligned}
0 = &\frac{(\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2} + \left(\frac{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + (2\alpha^2-1)\sigma_v^2}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta \\
&+ \left(\frac{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + [\alpha^2 + \rho^2(\alpha^2-1)]\sigma_v^2}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2}\right)\theta^2
\end{aligned}
$$

SI→OC

Similar to the result relating DE→OC, we have

$$\alpha^2 - \frac{\beta_1^{SI}}{\sigma_v^2}\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) = \beta_1^{SI} + 1,$$

which implies:

$$\alpha = \sqrt{\frac{(1+R)\kappa_1\sigma_v^2 + R(\kappa_1 - \kappa_2)\mathbb{V}(x_t - \widehat{x}^i_{t|t-1})}{(\kappa_1 + R\kappa_2)\sigma_v^2}}.$$

Again, there is no closed form expression that maps α to R since the state estimation error, $\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) \equiv \widehat{\Psi}(\alpha)$, is a nonlinear function of α.

OC→SI

Finally, given α in the overconfidence model, we can equate β_1^{OC} to β_1^{SI} by setting R equal to

$$R = \frac{(\alpha^2 - 1)\sigma_v^2\kappa_1}{(\kappa_1 - \kappa_2)[\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}) + \sigma_v^2] - (\alpha^2 - 1)\sigma_v^2\kappa_2}.$$

Proof of Proposition 4

Recall that $\gamma_1 = \frac{\mathbb{C}(x^i_{t|t} - x^i_{t|t-1})}{\mathbb{V}(x^i_{t|t-1} - x^i_{t|t-2})}$.

(i) For the overconfidence model, the constant Kalman gain is:

$$\widehat{\kappa} = \frac{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1}}{\mathbb{V}(x_t - \widehat{x}^i_{t|t-1} + \sigma_v^2} \in [0,1].$$

Given this, we can evaluate γ_1:

$$\begin{aligned}
\gamma_1 &= \frac{\mathbb{C}\Big(\widehat{\kappa}(x_t - \widehat{x}^i_{t|t-1}) + \widehat{\kappa} v^i_t, \rho\widehat{\kappa}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \rho\widehat{\kappa} v^i_{t-1}\Big)}{\mathbb{V}\Big(\rho\widehat{\kappa}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \rho\widehat{\kappa} v^i_{t-1}\Big)} \\
&= \frac{\mathbb{C}[\widehat{\kappa}(x_t - \widehat{x}^i_{t|t-1}), \rho\widehat{\kappa}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \rho\widehat{\kappa} v^i_{t-1}]}{\rho^2\widehat{\kappa}^2[\mathbb{V}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \sigma_v^2]} \\
&= \frac{\rho^2\widehat{\kappa}^2(1-\widehat{\kappa})\mathbb{V}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) - \widehat{\kappa}^3\rho^2\sigma_v^2}{\rho^2\widehat{\kappa}^2[\mathbb{V}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \sigma_v^2]} \\
&= \frac{(1-\widehat{\kappa})\mathbb{V}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) - \widehat{\kappa}\sigma_v^2}{\mathbb{V}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \sigma_v^2} \\
&= \frac{\widehat{\sigma}_v^2\widehat{\kappa} - \sigma_v^2\widehat{\kappa}}{\mathbb{V}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \sigma_v^2} \\
&= \frac{(\alpha^2 - 1)\sigma^2\widehat{\kappa}}{\mathbb{V}(x_{t-1} - \widehat{x}^i_{t-1|t-2}) + \sigma_v^2} = \widehat{\kappa}\beta_1^{OC} \leq \beta_1^{OC}.
\end{aligned}$$

(ii) In the diagnostic expectations model, the current-period prediction is:

$$x^{i,\theta}_{t|t} = \mathbb{E}(x_t|\mathcal{I}_t) + \theta\big[\mathbb{E}(x_t|\mathcal{I}^i_t - \mathbb{E}(x_t|\mathcal{I}^i_{t-1})\big],$$

where $\mathbb{E}(x_t|\mathcal{I}^i_t) \equiv x^i_{t|t}$ and $\mathbb{E}(x_t|\mathcal{I}^i_{t-1}) \equiv x^i_{t|t-1}$. The forecast revision is therefore:

$$x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1} = (1+\theta)(x^i_{t|t} - x^i_{t|t-1}) - \theta\rho(x^i_{t-1|t-1} - x^i_{t-1|t-2})$$

As a result, the first-order autocovariance of revisions is:

$$\mathbb{C}(x^{i,\theta}_{t|t} - x^{i,\theta}_{t|t-1}, x^{i,\theta}_{t|t-1} - x^{i,\theta}_{t-1|t-2}) = -(1+\theta)\theta\rho^2\mathbb{V}(x^i_{t-1|t-1} - x^i_{t-1|t-2})$$

and the variance of the lagged revision is:

$$\mathbb{V}(x^{i,\theta}_{t|t-1} - x^{i,\theta}_{t-1|t-2}) = \rho^2(1+\theta)^2\mathbb{V}(x^i_{t-1|t-1} - x^i_{t-1|t-2}) + \theta^2\rho^4\mathbb{V}(x^i_{t-2|t-2} - x^i_{t-2|t-3})$$

Noting that the error variance is time invariant, and evaluating the covariance divided by the variance, we obtain:

$$\gamma_1 = \frac{-\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2} = \beta_1^{DE}$$

(iii) It suffices to show that $\mathrm{Cov}(\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1}, \widetilde{x}^i_{t|t-1} - \widetilde{x}^i_{t|t-2}) = 0$. The current period revision is:

$$\begin{aligned}\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} &= \frac{1-\lambda-1-R\lambda}{1+R}x^i_{t|t-1} + \lambda y_{it} \\ &= \lambda(y_{it} - x^i_{t|t-1})\end{aligned}$$

and the previous period revision can be expressed as:

$$
\begin{aligned}
\widetilde{x}^i_{t|t-1} - \widetilde{x}^i_{t|t-2} &= \frac{1}{1+R}(x^i_{t|t-1} - x^i_{t|t-2}) + \frac{R}{1+R}(F^i_{t|t-1} - F^i_{t|t-2}) \\
&= \frac{\rho}{1+R}(x^i_{t-1|t-1} - x^i_{t-1|t-2}) + \frac{R}{1+R}\big[\lambda\rho x^i_{t-1|t-1} \\
&+ (1-\lambda)\rho F^i_{t-1|t-1} - \lambda\rho x^i_{t-1|t-2} - (1-\lambda)\rho F^i_{t-1|t-2}\big] \\
&= \frac{\rho}{1+R}\big[\kappa_1(y_{it-1} - x^i_{t-1|t-2})\big] + \frac{R}{1+R}\big[\lambda\rho(x^i_{t-1|t-1} - x^i_{t-1|t-2}) \\
&+ (1-\lambda)\rho(F^i_{t-1|t-1} - F^i_{t-1|t-2})\big] \\
&= \frac{\rho\kappa_1}{1+R}\big[x_{t-1} - x^i_{t-1|t-2} + v_{it-1}\big] \\
&+ \frac{R}{1+R}\big[\lambda\rho\kappa_1(x_{t-1} - x^i_{t-1|t-2} + v_{it-1}) \\
&+ (1-\lambda)\rho\kappa_2(x_{t-1} - x^i_{t-1|t-2} + v_{it-1})\big] \\
&= \left[\frac{\rho\kappa_1 + R\lambda\rho\kappa_1 + R(1-\lambda)\rho\kappa_2}{1+R}\right](x_{t-1} - x^i_{t-1|t-2} + v_{it-1}) \\
&= \frac{\rho(\kappa_1 + R\kappa_2) + \rho R\lambda(\kappa_1 - \kappa_2)}{1+R}(x_{t-1} - x^i_{t-1|t-2} + v_{it-1}) \\
&= \rho\lambda\left(1 + \frac{R(\kappa_1 - \kappa_2)}{1+R}\right)(x_{t-1} - x^i_{t-1|t-2} + v_{it-1})
\end{aligned}
$$

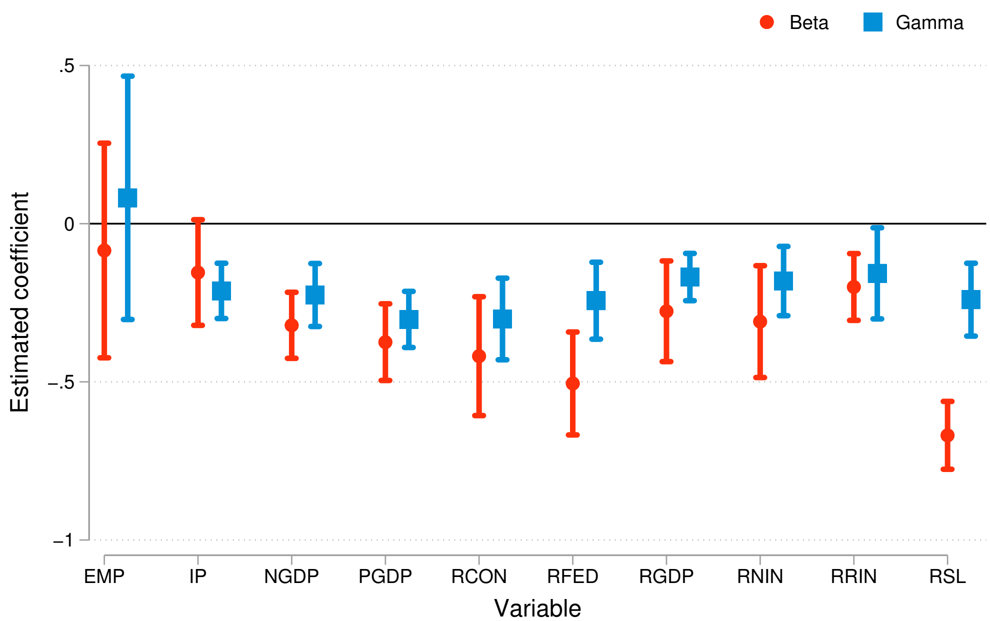
Then,

$$
\mathrm{Cov}(\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1}, \widetilde{x}^i_{t|t-1} - \widetilde{x}^i_{t|t-2}) = 0
$$

By definition, $\kappa_1 = \frac{\Psi}{\Psi+\sigma_v^2}$ where Ψ is the steady state forecast error variance (i.e. the variance that solves the Ricatti equation: $\Psi = (1-\kappa_1)\rho^2\Psi + \sigma_w^2$). From this it follows that the last term in hard brackets is zero.

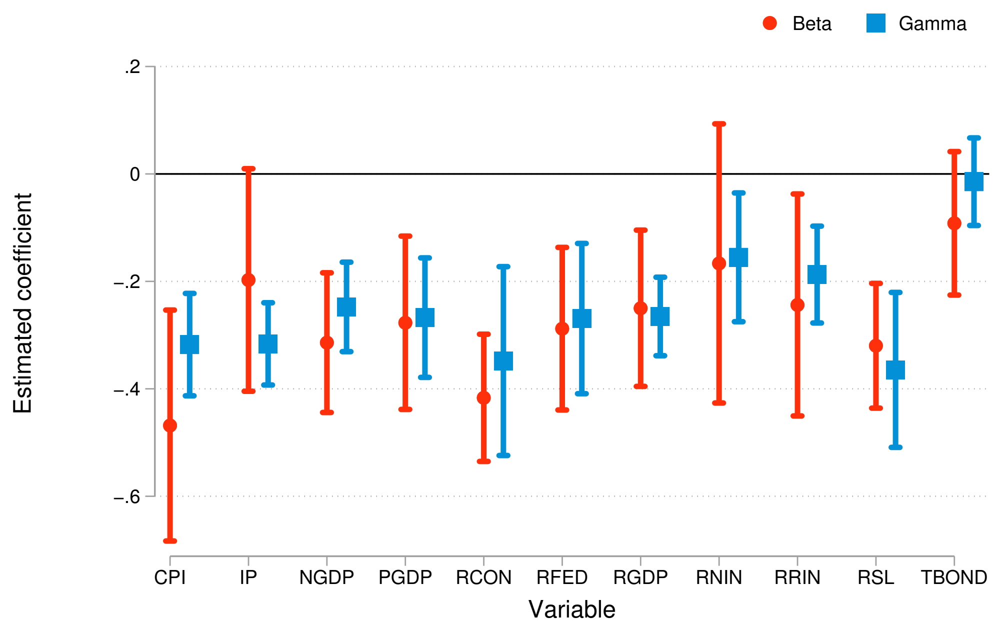
A.3.2 Empirics

Figure C1: Error Predictability and Revision Persistence, $H = 0$



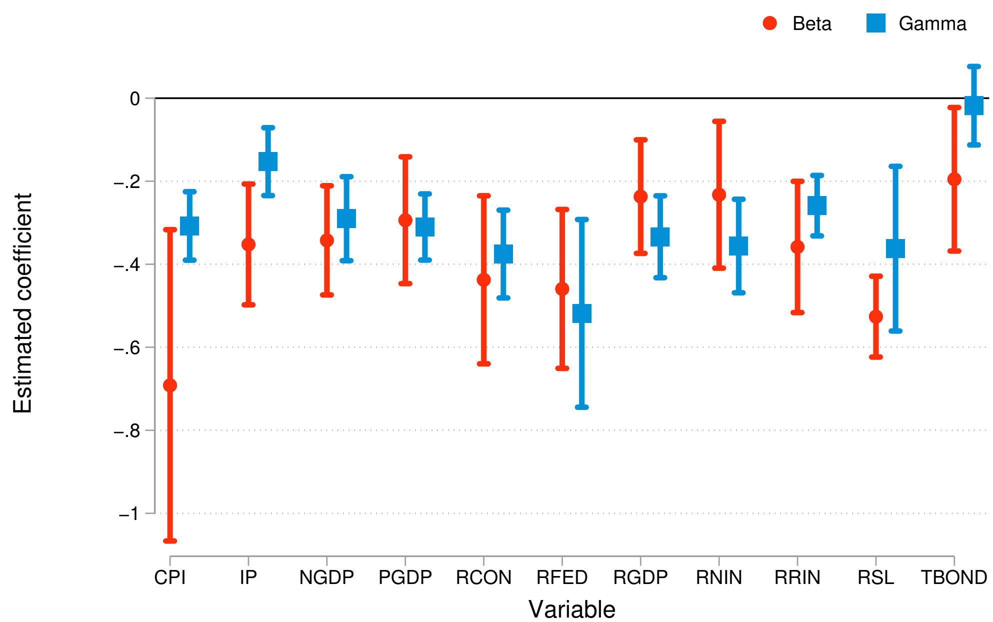
Notes: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. EMP- Employment, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, and RSL - Real State and Local Government Spending.

Figure C2: Error Predictability and Revision Persistence, $H = 1$



Notes: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. CPI - Consumer Price Index, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, RSL - Real State and Local Government Spending, and TBOND - 10 Year Government Bond.

Figure C3: Error Predictability and Revision Persistence, $H = 2$



Notes: Figure displays 95% confidence intervals for estimated coefficients of error predictability and revision persistence, (β_1 and γ_1 in the main text, respectively). Double clustered standard errors are specified. CPI - Consumer Price Index, IP - Industrial Production, NGDP - Nominal GDP, PGDP - GDP Deflator, RCON - Real Consumption, RFED - Real Federal Government Spending, RGDP - Real GDP, RNIN - Real Nonresidential Investment, RRIN - Real Residential Investment, RSL - Real State and Local Government Spending, and TBOND - 10 Year Government Bond.

A.3.3 Alternate Models

Heterogeneous Priors

In a model with heterogeneous priors (Patton & Timmermann, 2010), forecasters wish to minimize their mean squared forecast errors while tracking their long-run beliefs. Let the prior be μ^i. Then the forecaster's objective is:

$$\min_{\widetilde{x}^i_{t|t}} \left((x_t - \widetilde{x}^i_{t|t})^2 + \omega(\widetilde{x}^i_{t|t} - \mu^i)^2 \right)$$

The optimal reported forecast is:

$$\widetilde{x}^i_{t|t} = \frac{1}{1+\omega}\mathbb{E}(x_t|\mathcal{I}^i_t) + \frac{\omega}{1+\omega}\mu^i.$$

In this case, the deviation from the conditional expectation is $\tau^i_{t|t} = \frac{\omega}{1+\omega}(\mu^i - \mathbb{E}(x_t|\mathcal{I}^i_t)$. Furthermore, this model implies that revisions are unpredictable:

$$\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} = \frac{1}{1+\omega}\left[\mathbb{E}(x_t|\mathcal{I}^i_t) - \mathbb{E}(x_t|\mathcal{I}^i_{t-1})\right]$$

Measurement Error

In a model of measurement error, or trembling-hand noise as in Branch (2004), implies that the reported forecast is $\widetilde{x}^i_{t|t} = \mathbb{E}(x_t|\mathcal{I}^i_t) + \xi^i_t$. The measurement error, ξ^i_t is distributed $\xi^i_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2_\xi)$. While such models can deliver a negative coefficient of error predictability, the forecast revision is simply $\widetilde{x}^i_{t|t} - \widetilde{x}^i_{t|t-1} = \xi^i_t - \xi^i_{t-1}$, and hence, unpredictable.

Smoothing Motives

First, I discuss a standard model of smoothing motives. If forecasters wish to strategically temper their revisions, as in the smoothing motive models of Coibion and

Gorodnichenko (2015) and Kucinskas and Peters (2019), then the optimal forecast is

$$\widetilde{x}^i_{t|t} = (1-\phi)\mathbb{E}(x_t|\mathcal{I}^i_t) + \phi\widetilde{x}^i_{t|t-1}$$

where $\phi = \frac{\alpha(1-\delta)}{1+\alpha(1-\delta)}$. The parameter α governs the smoothing motive. As $\alpha \to \infty$, the smoothing motive (or quadratic revision cost) increases, and forecasters are less willing to update their forecasts. On the other hand, if $\alpha = 0$, then the smoothing motive disappears. Furthermore, δ is the discount factor. From this model, the autocorrelation of revisions is simply $\phi > 0$. In other words, revisions are positively autocorrelated.

I next, introduce a model with multi-horizon forecasting and smoothing with respect to the long term forecast. This model is rational, however, it is capable of generating negatively serially correlated revisions. Suppose that in each period, t, forecasters generate predictions for horizons $\{t+h\}_{h=0}^{H}$ with $H > 0$. At the same time, there is a long-run forecast, or prior, that forecasters wish to track for reputational reasons. As a result, a forecaster's objective is:

$$\max_{\widehat{x}^i_{t+h|t}} \sum_{h=0}^{H}(x_{t+h} - \widehat{x}^i_{t+h|t})^2 + \lambda\left[\left(\frac{1}{H+1}\sum_{h=0}^{H}\widehat{x}^i_{t+h|t}\right) - \overline{x}\right]^2$$

Asymmetric Attention

Kohlhas and Walther (2020) present an elegant theory of asymmetric attention which is able to generate overreaction to current output realizations. This theory, however, is unable to replicate overreaction on the basis of forecast revisions. Furthermore, I demonstrate here that it is also unable to generate serially correlated revisions.[12]

I consider a stylized version of the baseline asymmetric attention model in Kohlhas and Walther (2020). In particular, suppose the target forecast variable is a function

[12]The extended version of this model that incorporates irrational overconfidence, however, would be able to deliver these facts.

of two structural components:

$$y_t = x_{1t} + x_{2t},$$

where

$$x_{1t} = a_1\theta_t + b_1 u_{1t}$$
$$x_{2t} = a_2\theta_t + b_2 u_{2t}$$
$$\theta_t = \rho\theta_{t-1} + \eta_t,$$

where $u_{jt} \sim \mathcal{N}(0,1)$ and $\eta_t \sim \mathcal{N}(0,\sigma_\eta^2)$. Forecasters do not observe the structural components or the latent state, θ. Instead, they observe two private signals each period,

$$y^i_{1t} = x_{1t} + q_1\epsilon^i_{jt}$$
$$y^i_{2t} = x_{2t} + q_2\epsilon^i_{2t},$$

with $\epsilon^i_{jt} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. The precision of signals is component-specific so that weight placed on new information (i.e. attention) is asymmetric across components. This weight can be defined as:

$$m_j = \frac{b_j^2}{b_j^2 + q_j^2}.$$

As a result, the conditional expectation for a given component is:

$$\mathbb{E}(x_{jt}|\mathcal{I}^i_t) = m_j z^i_{jt} + (1-m_j)\mathbb{E}(x_{jt}|\mathcal{I}^i_{t-1}),$$

and the conditional expectation of the latent state is:

$$\mathbb{E}(\theta_t|\mathcal{I}^i_t) = \mathbb{E}(\theta_t|\mathcal{I}^i_{t-1}) + g_1[z^i_{1t} - \mathbb{E}(z^i_{1t}|\mathcal{I}^i_{t-1})] + g_2[z^i_{2t} - \mathbb{E}(z^i_{2t}|\mathcal{I}^i_{t-1})],$$

where g_j refers to the component-specific Kalman gain (a function of m_j).

From the above set up, one can define the forecast revision for y_{t+1} to be equal to the sum of the revisions made to the unobserved components x_{1t+1} and x_{2t+1}:

$$y^i_{t+1|t} - y^i_{t+1|t-1} = (x^i_{1,t+1|t} - x^i_{1,t+1|t-1}) + (x^i_{2,t+1|t} - x^i_{2,t+1|t-1}).$$

The righthand side of this equation, however, is simply equal to the sum of the innovation errors scaled by the attention coefficients,

$$y^i_{t+1|t} - y^i_{t+1|t-1} = m_1[z^i_{1,t} - \mathbb{E}(x_{1,t}|\mathcal{I}^i_{t-1})] + m_2[z^i_{2,t} - \mathbb{E}(x_{2,t}|\mathcal{I}^i_{t-1})],$$

which are themselves serially uncorrelated as a result of Bayesian updating. Hence, this theory is unable to accommodate autocorrelated revisions.

References

Afrouzi, H., Kwon, S., Landier, A., Ma, Y., & Thesmar, D. (2020). *Overreaction and Working Memory.* (Working Paper)

Ahmed, S., Levin, A., & Wilson, B. A. (2002). Recent U.S. Macroeconomic Stability: Good Policies, Good Practices or Good Luck?

Alessandria, G., & Choi, H. (2007). Do Sunk Costs of Exporting Matter for Net Export Dynamics? *The Quarterly Journal of Economics*, *122*, 289–336.

An, Z., Liu, D., & Wu, Y. (2021). Expectation Formation Following Pandemic Events. *Economics Letters*, *200*.

Andrade, P., & Bihan, H. L. (2013). Inattentive Professional Forecasters. *Journal of Monetary Economics*, *60*, 967–982.

Andrews, I., Gentzkow, M., & Shapiro, J. M. (2017). Measuring the Sensitivity of Parameter Estimates to Estimation Moments. *The Quarterly Journal of Economics*, *132*, 1553–1592.

Arellano, C., Bai, Y., & Mihalache, G. (2020). *Deadly Debt Crises: COVID-19 in Emerging Markets.* (Working Paper)

Azeredo da Silvera, R., & Woodford, M. (2019). *Noisy Memory and Overreaction to News.* (NBER Working Paper 25478)

Bachmann, R., Caballero, R. J., & Engel, E. M. (2013). Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model. *American Economic Journal: Macroeconomics*, *5*, 29–67.

Bachmann, R., & Ma, L. (2016). Lumpy Investment, Lumpy Inventories. *Journal of Money, Credit and Banking*, *48*, 821–855.

Barrero, J. M. (2020). *The Micro and Macro of Managerial Beliefs.* (Working Paper)

Barro, R., & Ursua, J. (2008). Macroeconomic Crises Since 1870. *Brookings Papers* *on Economic Activity*, 255–335.

Barrot, J.-N., & Sauvagnat, J. (2016). Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks. *The Quarterly Journal of Economics*, *131*, 1543–1592.

Bazdresch, S. R., Kahn, J., & Whited, T. M. (2018). Estimating and Testing Dynamic Corporate Finance Models. *Review of Financial Studies*, *31*, 322–361.

Benigno, P., & Karantounias, A. G. (2019). Overconfidence, Subjective Perception and Pricing Behavior. *Journal of Economic Behavior and Organization*, *164*, 107–132.

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020). Overreaction in Macroeconomic Expectations. *American Economic Review*, *110*, 2748–2782.

Branch, W. (2004). The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations. *The Economic Journal*, *114*, 592–621.

Brock, W. A., & Hommes, C. H. (1997). A Rational Route to Randomness. *Econometrica*, *65*, 1059–1095.

Broer, T., & Kohlhas, A. (2019). *Forecaster (Mis-)Behavior.* (Working Paper)

Bürgi, C. (2016). *What Do We Lose When We Average Expectations?* (Working Paper)

Cachon, G., Randall, T., & Schmidt, G. (2007). In Search of the Bullwhip Effect. *Manufacturing and Service Operations Management*, *4*, 457–479.

Caplin, A. S. (1985). The Variability of Aggregate Demand with (S,s) Inventory Policies. *Econometrica*, *53*, 1395–1409.

Carvalho, V., Nirei, M., Saito, Y., & Tahbaz-Salehi, A. (2016). *Supply Chain Disruptions: Evidence from the Great East Japan Earthquake.* (Working Paper)

Clementi, G., Castro, R., & Lee, Y. (2015). Cross-Sectoral Variation in the Volatility of Plant-Level Idiosyncratic Shocks. *Journal of Industrial Economics*, *63*, 1–29.

Coibion, O., & Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic*

Review, *109*, 465–490.

Crisan, D., & Doucet, A. (2002). A Survey of Convergence Results on Particle Filtering Methods for Practitioners. *IEEE Transactions on Signal Processing*, *50*, 736–746.

Crowe, C. (2010). *Consensus Forecasts and Inefficient Information Aggregation.* (Working Paper)

Cua, K. O., McKone, K. E., & Schroeder, R. G. (2001). Relationships Between Implementation of TQM, JIT, and TPM, and Manufacturing Performance. *Journal of Operations Management*, *19*, 675–694.

D'Acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2019). *IQ, Expectations, and Choice.* (Becker Friedman Institute Working Paper)

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. *Journal of Finance*, *53*, 1839–1885.

Davis, S., & Kahn, J. (2008). Interpreting the Great Moderation: Changes in the Volatility of Economic Activity at the Macro and Micro Levels. *Journal of Economic Perspectives*, *22*, 155–180.

D'Haultfoeuielle, X., Gaillac, C., & Maruel, A. (2020). *Rationalizing Rational Expectations: Characterizations and Tests.* (Working Paper)

Doucet, A., & Johansen, A. M. (2009). A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later. *Handbook of Nonlinear Filtering*.

Dovern, J., Fritsche, U., Loungani, P., & Tamirisa, N. (2015). Information Rigidities: Comparing Average and Individual Forecasts for a Large International Panel. *International Journal of Forecasting*, *31*, 144–154.

Driscoll, J. C., & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statistics*, *80*, 549–560.

Duffie, D., & Singleton, K. J. (1993). Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica*, *61*, 929–952.

Dun Jia, C. (2016). *Disagreement vs. Uncertainty: Investment Dynamics and Business Cycles.* (Working Paper)

Eichenbaum, M. S. (1984). Rational Expectations and the Smoothing Properties of Inventories of Finished Goods. *Journal of Monetary Economics*, *14*, 71–96.

Espino, E., Kozlowski, J., Martin, F. M., & Sanchez, J. M. (2020). *Seigniorage and Sovereign Default: The Response of Emerging Markets to COVID-19.* (Working Paper)

Evans, G. W., & Ramey, G. (1992). Expectation Calculation and Macroeconomic Dynamics. *American Economic Review*, *82*, 207–224.

Fuhrer, J. (2018). *Intrinsic Expectations Persistence: Evidence from Professional and Household Survey Expectations.* (Working Paper)

Fullerton, R. R., & McWatters, C. S. (2001). The Production Performance Benefits from JIT Implementation. *Journal of Operations Management*, *19*, 81–96.

Fuster, A., Hebert, B., & Laibson, D. (2012a). Natural Expectations, Macroeocnomic Dynamics, and Asset Pricing. *NBER Macroeconomics Annual*, *26*, 1–48.

Fuster, A., Hebert, B., & Laibson, D. (2012b). Natural Expectations, Macroeocnomic Dynamics, and Asset Pricing. *NBER Macroeconomics Annual*, *26*, 1–48.

Fuster, A., Laibson, D., & Mendel, B. (2010). Natural Expectations and Macroeconomic Fluctuations. *Journal of Economic Perspectives*, *24*, 67–84.

Gabaix, X. (n.d.). Variable Rare Disasters: An Exactly S]olved Framework for Ten Puzzles in Macro-Finance, journal=The Quarterly Journal of Economics, volume=127, year=2012, pages=645–700.

Gao, X. (2018). Corporate Cash Hoarding: The Role of Just-in-Time Adoption. *Management Science*, *64*, 4471–4965.

Gilchrist, S., Sim, J. W., & Zakrajsek, E. (2014). *Uncertainty, Financial Frictions, and Investment Dynamics.* (NBER Working Paper 20038)

Gordon, N., Salmond, D., & Smith, A. F. M. (n.d.). Novel Approach to Nonlinear and Non-Gaussian Bayesian State Estimation. *IEE*, *140*.

Hennessy, C. A., & Whited, T. M. (2007). How Costly is External Financing? Evidence from a Structural Estimation. *Journal of Finance*, *62*, 1705–1745.

Hu, X.-L., Schon, T. B., & Ljung, L. (2011). A General Convergence Result for

Particle Filtering. *IEEE Transactions on Signal Processing*, *59*, 3424–3429.

Iacoviello, M., Schiantarelli, F., & Schuh, S. (2011). Input and Output Inventories in General Equilibrium. *International Economic Review*, *52*, 1179–1213.

Irvine, O. F., & Schuh, S. (2005). Inventory Investment and Output Volatility. *International Journal of Production Economics*, *93-94*, 75–86.

Julier, S. J., & Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *IEEE Transactions on Signal Processing*, *92*, 401–422.

Justiniano, A., & Primiceri, G. E. (2008). The Time Varying Volatility of Macroeconomic Fluctuations. *American Economic Review*, *98*, 604–641.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME–Journal of Basic Engineering*, *82*, 35–45.

Kastner, G., & Frühwirth-Schnatter, S. (2014). Ancillarity-Sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Estimation of Stochastic Volatility Models of the Business Cycle. *Computational Statistics and Data Analysis*, *76*, 408–423.

Khan, A., & Thomas, J. (2007). Inventories and the Business Cycle: An Equilibrium Analysis of (S,s) Policies. *American Economic Review*, *97*, 1165–1188.

Khan, A., & Thomas, J. (2008). Idiosyncratic Shocks and the Role of Nonconvexitiies in Plant and Aggregate Investment Dynamics. *Econometrica*, *76*, 395–436.

Kim, C.-J., & Nelson, C. (1999). Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle. *Review of Economics and Statistics*, *81*, 608–616.

Kinney, M. R., & Wempe, W. F. (2002). Evidence on the Extent and Origins of JIT's Profitability Effects. *The Accounting Review*, *77*, 203–225.

Kohlhas, A. N., & Walther, A. (2020). *Asymmetric Attention.* (Working Paper)

Krusell, P., Mukoyama, T., Sahin, A., & Smith, J. A. A. (2009). Revisiting the Welfare Effects of Eliminating Business Cycles. *Review of Economic Dynamics*, *12*, 393–402.

Kucinskas, S., & Peters, F. (2019). *Measuring Biases in Expectation Formation.*

(Working Paper)

Lansing, K. J., LeRoy, S. F., & Ma, J. (2020). *Examining the Sources of Excess Return Predictability: Stochastic Volatility or Market Inefficiency?* (Federal Reserve Bank of San Francisco Working Paper 2018-14)

Larsen, V. H., Thorsrud, L. A., & Zhulanova, J. (2020). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics (Forthcoming).*

Lewis, K. (1989). Changing beliefs and systematic rational forecast errors with evidence from foreign exchange. *American Economic Review*, *99*, 769–803.

MacQueen, J. B. (1966). A Modified Dynamic Programming Method for Markovian Decision Problems. *Journal of Mathematical Analysis and Applications*, *14*, 38–43.

McCarthy, J., & Zakrajsek, E. (2007). Inventory Dynamics and Business Cycles: What Has Changed? *Journal of Money, Credit and Banking*, *39*, 591–613.

McConnell, M. M., & Perez-Quiros, G. (2000). Output Fluctuations in the United States: What Has Changed Since the Early 1980's? *American Economic Review*, *90*, 1464–1476.

McMahon, M., & Wanengkirtyo, B. (2015). *Beyond Inventory Management: The Bullwhip Effect and the Great Moderation.* (Working Paper)

Meier, M. (2020). *Supply Chain Disruptions, Time to Build and the Business Cycle.* (Working Paper)

Morris, S., & Shin, H. S. (2002). Social Value of Public Information. *American Economic Review*, *92*, 1521–1534.

Nakamura, M., Sakakibara, S., & Schroeder, R. (1998). Adoption of Just in Time Manufacturing Methods at U.S. and Japanese-Owned Plants: Some Empirical Evidence. *IEEE Transactions on Engineering Management*, *45*, 230–240.

Nordhaus, W. D. (1987). Forecasting Efficiency: Concepts and Applications. *The Review of Economics and Statistics*, *69*, 667–674.

Ohno, T. (1988).
In *Toyota Production System: Beyond Large-Scale Production.* Productivity

Press.

Ortiz, J. (2020). *Time-Varying Volatility as a Source of Error Predictability.* (Working Paper)

Ottaviani, M., & Sørensen, P. (2006). The Strategy of Professional Forecasting. *Journal of Financial Economics*, *2*, 441–466.

Ottonello, P., & Winberry, T. (2020). *Financial Heterogeneity and the Investment Channel of Monetary Policy.* (Working Paper)

Paloviita, M., & Viren, M. (2013). *Are Individual Survey Expectations Internally Consistent?* (Working Paper)

Patton, A. J., & Timmermann, A. (2010). Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion. *Journal of Monetary Economics*, *57*, 803-820.

Pesaran, M. H., & Weale, M. (2006). Chapter 14 Survey Expectations. *Handbook of Economic Forecasting*, *1*, 715-776.

Porteus, E. L. (1971). Some Bounds for Discounted Sequential Decision Processes. *Management Science*, *18*, 7–11.

Ramey, V. A., & Vine, D. J. (2004). *Tracking the Source of the Decline in GDP Volatility: An Analysis of the Automobile Industry.* (NBER Working Paper 10384)

Rogerson, R. (1988). Indivisible Labor, Lotteries and Equilibrium. *Journal of Monetary Economics*, *21*, 3–16.

Roumiantsev, S., & Netessine, S. (2008). *Should Inventory Policy Be Lean or Responsive? Evidence for US Public Companies.* (Working Paper)

Rozsypal, F., & Schlafmann, K. (2019). *Overpersistence Bias in Individual Income Expectations and its Aggregate Implications.* (Working Paper)

Scarf, H. E. (1960). The Optimality of (S,s) Policies in the Dynamic Inventory Problem. In *Mathematical Methods in the Social Sciences* (pp. 196–202). Stanford: Stanford University Press.

Scotese, C. A. (1994). Forecast Smoothing and the Under-Utilization of Information at the Federal Reserve. *Journal of Macroeconomics*, *16*, 653–670.

Senga, T. (2018). *A New Look at Uncertainty Shocks: Imperfect Information and Misallocation.* (Working Paper)

Stock, J., & Watson, M. (2002). Has the Business Cycle Changed and Why? *NBER Macroeconomics Annual*, *17*, 159–224.

Stock, J., & Watson, M. (2007). Why has U.S. Inflation Become Harder to Forecast? *Journal of Money, Credit and Banking*, *39*, 4–33.

Tauchen, G. (1986). Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions. *Economics Letters*, *20*, 177–181.

Taylor, L. A. (2010). Why Are CEOs Rarely Fired? Evidence from Structural Estimation. *Journal of Finance*, *65*, 2051–2087.

Terry, S. (2017). *The Macro Impact of Short-Termism.* (Working Paper)

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*, 1124–1131.

Veronesi, P. (2015). Stock Market Overreactions to Bad News in Good Times: A Rational Expectations Equilibrium Model. *The Review of Financial Studies*, *12*, 975–1007.

Wen, Y. (2011). Input and Output Inventory Dynamics. *American Economic Journal: Macroeconomics*, *3*, 181–212.

Woodford, M. (2001). Imperfect Common Knowledge and the Effects of Monetary Policy. *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund Phelps*.

Young, E. R. (2010). Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell-Smith Algorithm and Non-Stochastic Simulations. *Journal of Economic Dynamics and Control*, *34*, 36–41.

CURRICULUM VITAE

[redacted]