What do you expect?
Revisiting Expectation Formation Models

By

Calvin McElvain

An honors thesis submitted in partial fulfillment
of the requirements for Honors in the Major
in Analytical Economics in the
Tippie College of Business
The University of Iowa

Spring 2024

Advisor: Dr. Anne P. Villamil

ACKNOWLEDGMENTS

This thesis is dedicated to the relentless support of my advisor, Dr. Anne P. Villamil, and the invaluable guidance of several faculty members and mentors: Dr. Alexandra Nica, Dr. David J. Cooper, Dr. Seongjoo Min, Dr. Sarah Frank, and Dr. Julia Garlick. As well as to my parents, Sheila and William, and my two brothers, Collin and Brandon.

What do you expect?
Revisiting Expectation Formation Models

By Calvin McElvain

This paper replicates prominent expectation formation models using contemporary inflation data across full and restricted samples. Employing the approach of Coibion and Gorodnichenko (2015a), estimates of mean ex-ante forecast revisions on mean ex-post forecasting errors consistently indicate systematic underpredictions, aligning with prior literature. Structural parameter estimates within models incorporating information costs reveal notable information rigidities. Conversely, individual-level analysis reveals patterns of overreaction among forecasters. Notably, parameter estimates from a diagnostic expectations model suggest that individual forecasters assign approximately 56 percent more weight to new information than those of rational forecasters. However, these findings are tempered when observed across the full sample.

In economic interactions, expectations play a fundamental role, serving as the cornerstone for decision-making, price setting, and policy formation. Over time, expectation modeling has evolved, initially characterized by the dominance of backward-looking models. Termed Adaptive Expectations, these models portrayed agents with ”lagged” or anchored beliefs, yet they proved inconsistent and failed to explain phenomena like stagflation during the 1970s. This deficiency was attributed to their inability to incorporate current information, such as policy regime shifts, as emphasized in the critique by Lucas (1976).

Subsequent to the Lucas critique, Sargent and Lucas (1979) advocated for forward-looking expectations, termed Full-information Rational Expectations (FIRE). This adoption of FIRE was somewhat belated, with its initial proposition by Muth (1961), who argued against the passive, backward-looking nature of adaptive expectations, favoring forward-looking, information-driven expectations instead. The conceptual strides of FIRE were underscored by their utilization in various other models. Notably, Calvo (1983) introduced a staggered prices model, incorporating FIRE, which gained traction for its widespread adoption in New Keynes Phillips Curves (NKPCs).

However, subsequent survey evidence revealed certain shortcomings and complexities within real-world settings that FIRE struggled to capture. One proposed shortcoming comes in the form of information rigidities, where agents do not promptly or accurately incorporate all available information into their expectations. Models such as the sticky-information model (Mankiw & Reis, 2002) and noisy-information models (Sims, 2003; Woodford, 2003) emerged as attempts to ride the wave of rational expectations with adjustments to information. Alternatively, models such as the diagnostic expectations model (Bordalo et al., 2019), proposed deeper departures from rational expectations, building off the behavioral model of Kahneman and Tversky (1973) representative heuristic. Other explanatory models includes an asymmetric attention model (Kohlhas & Walther, 2021), which incorporates both information rigidities and extrapolation but argues economic agents give asymmetric attention to structural components and thus build their expectations accordingly.

This thesis will strongly follow and add to existing literature with focus on expectation formation models including Angeletos et al. (2021), Bordalo et al. (2020), Coibion and Gorodnichenko (2015a), Coibion, Gorodnichenko, and Kumar (2018), and Kohlhas and Walther (2021). Specifically, this thesis evaluates models using contemporary inflation data (CPI) with a focus on a sticky-information model, a noisy-information model, and a diagnostic expectations model. While all macroeconomic indicators are valuable for understanding the general formation of expectations, this thesis focuses exclusively on inflation, justified by the availability of data and the vital role that inflation plays. Additionally, this thesis will contribute to an open-source Python code of these models[1].

The following sections of this thesis will begin with an overview of related literature on expectations, their biases, and the role of survey data. Proceeding this, the following models section will give the technical details and specifcations of each structural model evaluated and the approach to evaluating them. The next section, Data, will describe the characteristics and summary statistics of the data as well as provide preliminary details of statistical frameworks. Following this, the Results section will present the estimates and interpretations of each model as well as a discussion on the robustness of estimates. Finally, the Conclusion will synthesize the thesis, offering insights, implications, and directions for future research.

[1] See https://github.com/calvinmcelvain/expectation_replication

I. Related Literature

The foundational premise of expectations under the FIRE framework is that forecasting errors should lack predictability. This assumption was quickly called into question with the use of survey data as a proxy for expectations. Pesando (1975) and Carlson (1977), revealed departures from FIRE in inflation expectations among respondents in the Livingston Survey. This evidence was further pointed out in Zarnowitz and Braun (1992) who identified rejections of FIRE in inflation expectations through a joint macroeconomic survey conducted by the NBER and ASA. Given the significance of the rational revolution, this evidence was faced with considerable resistance. Some critiques focused on econometric techniques, advocating for methodologies accommodating weak-form rationality (Grant & Thomas, 1999; Paquet, 1992), while others proposed alternative explanations, such as limited survey sample sizes (Andolfatto et al., 2008) and age-dependent expectations (Malmendier & Nagel, 2015).

Arguments in favor of traditional FIRE models failed to gain momentum, giving rise to a substantial body of literature further unveiling survey biases and proposing alternative explanatory models. Mankiw et al. (2003) illistrate departures, in one form or another, from rational expectations in median inflation expectations across the MSC, Livingston survey, and the SPF. Furthermore, they observed systematic under-predictions in median inflation forecasts by professional forecasters (SPF) and propose a model, termed the sticky-information model, in which firms fractionally aquire new information about the state of the economy in each period. They show that the sticky-information model offered a more robust explanation for deviations not explained by the conventional FIRE model.

Coibion and Gorodnichenko (2012) adds to this evidence, finding under-reactions in average inflation forecasts by professional forecasters, consistent with models of information rigidities. However, Bordalo et al. (2020) and Capistrán and Timmermann (2009) report estimates indicating overreactions in forecasting data when evaluating professional forecasters at the individual-level. Angeletos et al. (2021) attempted to reconcile these disparities, finding both under-reactions in average inflation forecasts and overreactions in individual-level forecasts. They suggested the existence of an initial under-reaction followed by subsequent overreactions in expectations when responding to shocks, advocating for expectation models capable of incorporating both information rigidities and over-extrapolation. Chen et al. (2022) evaluates survey data in the Euro-area and identifies similar under- and over-reactions among professional forecasters at the aggregate and individual levels, respectively, compared to the United States.

More recently, as survey data has become increasingly available, literature has shifted focus to firm expectations and their motives in firm-level decision making. Coibion, Gorodnichenko, and Kamdar (2018) study New Zealand firms, finding large dispersion in inflation expectations with overestimates of inflation greater than that of professional forecasters. They concluded that models of information rigidities could predict these expectations due to infrequent information updating, as many firms perceive inflation as unimportant in firm decision-making. Coibion et al. (2019) use an information treatment approach among Italian firms to reveal the causal effect of information-driven expectations on firm decision-making. Similar to Coibion, Gorodnichenko, and Kumar (2018), they identified systematic over-estimations of inflation among firms, yet observed that firms in the control group (i.e., not subjected to the information treatment) exhibited significantly higher prices, reduced employment, and increased demand for capital. Furthermore, they found that firms in the information treatment group reported moderately higher profits than those in the control.

Collectively, evidence consistently refutes the assumptions of FIRE, yet the next expectation revolution remains untold. While many advocate for explanations such as information rigidities, over-extrapolation, learning, or hybrid approaches, substantial disagreement persists. This thesis aims to contribute to ongoing discussions by providing estimates of prevalent expectation formation models using contemporary inflation data from an ever-evolving economic landscape.

II. Expectation Formation Models

This section presents the 3 model frameworks evaluated in this thesis. Beginning with models of information rigidities, namely the sticky-information model (Mankiw & Reis, 2002,

2006) and a noisy information model (Woodford, 2003). And finally, shifting our focus to the diagnostic expectations model (Bordalo et al., 2019).

i. Sticky-Information Model

Mankiw and Reis (2002, 2006) introduce a model wherein agents' expectations stem from a rational framework incorporating costs of acquiring new information. In this model, agents update their information set at a rate λ. Given an agent updates their information in period $t-1$, they will incorporate the new information optimally (i.e. rationally) for period t. Otherwise, the agent will use information from the previous period (and periods thereafter). Following Mankiw and Reis (2002), let $f_t\pi_{t+j}$ denote aggregate (mean) $t+j$ forecasts *across agents* in period t. The sticky-information model is then represented by

$$f_t\pi_{t+j} = \lambda \sum_{k=0}^{\infty}(1-\lambda)^k \mathbb{E}_{t-k}\left[\pi_{t+j}\right]^j \tag{1}$$

In this context, given an agent does not update their information set in period t (i.e. $\lambda = 0$), equation (1) collapses into equation (2), a FIRE model. Here, we assume normality of the respective error term $\epsilon_{t+j,t}$ with mean zero and variance σ_ϵ^2. Additionally, we assume that π_t follows an AR(1) process with corresponding regressor ϕ ($\phi \in [0,1)$) and shock term w_t:

$$\mathbb{E}_t\left[\pi_{t+j}\right] = \pi_{t+j} - \epsilon_{t+j,t} \tag{2}$$

Utilizing a unique and effective approach by Coibion and Gorodnichenko (2015b), we can show the relationship between the mean ex-post forecast errors and mean ex-ante forecast revisions. Acknowledging $f_t\pi_{t+j}$ as the weighted average between $f_t\pi_{t+j}$ and $f_{t-1}\pi_{t+j}$, we can rewrite equation (1) as:

$$f_t\pi_{t+j} = (1-\lambda)\mathbb{E}_t\left[\pi_{t+j}\right] + \lambda f_{t-1}\pi_{t+j}$$

Hence, when the sticky-information model is expressed in terms of the forecast error, $\pi_{t+j} - f_t\pi_{t+j}$ (denote as $FE_{t+j|t}$), it can be written as:

$$\begin{aligned} FE_{t+j|t} &= f_t\pi_{t+j} + \epsilon_{t+j,t} - (1-\lambda)\mathbb{E}_t\left[\pi_{t+j}\right] - \lambda f_{t-1}\pi_{t+j} \\ &= f_t\pi_{t+j} + \epsilon_{t+j,t} - (1-\lambda)\left(\frac{f_t\pi_{t+j} - \lambda f_{t-1}\pi_{t+j}}{1-\lambda}\right) - \lambda f_{t-1}\pi_{t+j} \\ &= \frac{(1-\lambda)\mathbb{E}_t\left[\pi_{t+j}\right] + \lambda f_{t-1}\pi_{t+j} - f_t\pi_{t+j}}{1-\lambda} + \epsilon_{t+j,t} \\ &= \frac{\lambda}{1-\lambda}\left(f_t\pi_{t+j} - f_{t-1}\pi_{t+j}\right) + \epsilon_{t+j,t} \end{aligned} \tag{3}$$

This approach has several advantages. First, it offers a straightforward and practical means to examine FIRE. As per FIRE, this association ought to be unpredictable, where $\lambda = 0$, and equation (3) simplifies to equation (2). Additionally, it provides a straightforward and elegant method to recover the intrinsic structural parameter λ. In cases where the coefficient of revisions is greater than zero, this association maps directly to the degree of information rigidities ($\lambda > 0$). Lastly, because the associated error term, $\epsilon_{t+j,t}$, is the FIRE error term which is orthogonal to information in time t, we can use OLS[2].

ii. Noisy-Information Model

Sims (2003) and Woodford (2003) introduce models wherein agents gather and incorporate information rationally with the exception being that the signal or state is noisy, implying that agents never fully observe the true state of inflation. Following Woodford (2003), suppose inflation follows an AR(1) process with agent i receiving a signal equal to $z_t^i = \pi_t + w_t^i$ with w_t^i representing an agent-specifc shock. Then the noisy-information model is written as:

$$f_t^i\pi_t = f_{t-1}^i\pi_t + G\left[z_t^i - f_{t-1}^i\pi_t\right] \tag{4}$$

[2]While in theory this is true, we explore other theories that may introduce potential endogeneity in the robustness checks of the results section.

Here, G is a Kalman gains paramenter, where we assume $G \geq 0$. Similar to the sticky-information model, we can apply and follow Coibion and Gorodnichenko (2015b) to get the relationship between ex-post forecast errors and ex-ante forecast revisions. First, we describe the mean $t + j$ forecast across agents in periods t and $t - 1$. Since we assume $\epsilon_{t+j,t}$ is mean zero in period t, we can denote the mean forecast across agents as $f_t \pi_{t+j} = \mathbb{E}_t \left[\pi_{t+j} \right]$. Using equation (4) we obtain:

$$
\begin{aligned}
f_t \pi_{t+j} &= \rho^j f_t \pi_t \\
&= \left(1 - G\rho^j\right) f_{t-1}\pi_t + G\left(\pi_t + \epsilon_t\right) \qquad (5) \\
f_{t-1}\pi_{t+j} &= \rho^j f_{t-1}\pi_t \qquad (6)
\end{aligned}
$$

Finally, we express the noisy-information model in terms of the forecast error, again denoted by $FE_{t+j|t}$, as:

$$
\begin{aligned}
FE_{t+j|t} &= \pi_{t+j} - \left(1 - G\rho^j\right) f_{t-1}\pi_t + G\left[\pi_t + \epsilon_t^i\right] \\
&= \frac{1-G}{G}\left(f_{t-1}\pi_{t+j} - f_{t-1}\pi_{t+j}\right) + \epsilon_{t+j,t} \qquad (7)
\end{aligned}
$$

Alike to the sticky-information model, if $G = 1$, then the model collapses into a FIRE model, implying no noise. A very nice feature of using Coibion and Gorodnichenko (2015b), is that both equations, (3) and (7), can be estimated identically using OLS.

iii. Diagnostic Expectations Model

Bordalo et al. (2019) propose a model that integrates information rigidities, specifically noisy-information, with a ”stereotypes” model (Bordalo et al., 2016) inspired by the Kahneman and Tversky (1973) representative heuristic. They argue that while information rigidities better capture the under-reactions in consensus or mean forecasts, they fail to account for the systematic over-reactions observed in individual-level forecasts. Instead, they offer a model of beleifs in which agent's learn from a noisy signal and put wieght on what they believe to be more likely (”representative”).

Beginning with inflation, again following an AR(1) process, we describe the signal received by an agent i, as $s_t^i = \pi_t + \epsilon_t^i$. Following Bordalo et al. (2019), we define the representative heuristic model or ”diagnostic Kalman filter” as $G_\theta = 1 - \theta$. Combining this with an individual-level equation (6) we obtain:

$$f_t^i \pi_t = f_{t-1}^i \pi_t + G_\theta \left(s_t^i - f_{t-1}^i \pi_t \right) \tag{8}$$

Next we can define $f_t^i \pi_{t+j}$ and $f_{t-1}^i \pi_{t+j}$ in a similar fashion as before:

$$f_t^i \pi_{t+j} = \left(1 - G_\theta \rho^j\right) f_{t-1}^i \pi_{t+j} + G_\theta \left(\pi_{t+j} + \epsilon_t^i\right) \tag{9}$$

$$f_{t-1}^i \pi_{t+j} = \rho^j f_{t-1}^i \pi_{t+j} \tag{10}$$

Once again, following Coibion and Gorodnichenko (2015b), we can rewrite these equations in terms of ex-post forecast errors and ex-ante forecast revisions:

$$\begin{aligned} FE_{t+j,t}^i &= \pi_{t+j} - \left(1 - G_\theta \rho^j\right) f_{t-1}^i \pi_{t+j} + G_\theta \left(\pi_{t+j} + \epsilon_t^i\right) \\ &= \pi_{t+j} - \left(1 - (1+\theta)\rho^j\right) f_{t-1}^i \pi_{t+j} + (1+\theta)\left(\pi_t + \epsilon_t^i\right) \\ &= -\frac{\theta\,(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2} \left(f_t^i \pi_{t+j} - f_{t-1}^i \pi_{t+j}\right) + \epsilon_{t+j|t}^i \end{aligned} \tag{11}$$

Where θ characterizes the extent to which expectations deviate from rationality based on ”representativeness”. That is, θ is the percentage amount a forecaster overreacts to news/information compared to a rational forecaster (with noise). For $\theta = 0$, the equation collapses to a rational noisy-information model, i.e., equation (10). If, however, the coefficient on revisions is non-zero, θ is greater than zero and reflects ”overreaction” by rational forecasters. Similar to estimation in mean forecasts, the error term for equation (11) is a FIRE error term, which again is orthogonal to information in time t. Thus we can use pooled OLS to estimate our coefficients[3].

III. Data

i. *Real-time Data*

Inflation ”actuals” are calculated using vintages from the Philadelphia Fed’s Real-Time Data Set for Macroeconomists (1947III - 2022IV). The rationale for using this data is to closely match an agent’s or forecaster’s information set during the period in which they make their forecast. Thus, $t + j$ actuals are calculated from $t - 1$ vintages, i.e., CPI_{t+j}^{t-1} divided by CPI_{t-1}^{t-1} (with annual growth rates equal to CPI_{t+3}^{t-1} divided by CPI_{t-1}^{t-1}). Unfortunately, vintages only date back to 1994III, thus actuals in $t \leq$ 1994III are calculated using 1994III vintage data. The first-order autoregressive coefficient and real root for CPI actuals can be found in Appendix A along with persistence figures for each time horizon.

ii. *SPF Data*

Forecast data comes from the Philadelphia Fed’s Survey of Professional Forecasters (SPF; 1981III - 2022IV). This survey is conducted on a quarterly basis, in which individual forecasters make forecasts on a range of economic and financial indicators. For CPI, forecasters make forecasts for quarters $t - 1$ (forecasters know the $t - 1$ outcome at time t) to $t + 4$. In each time period used for this thesis (1981III - 2022IV), forecasts across horizons are winsorized to 5 IQR, and forecasters with fewer than 10 observations are excluded[4].

A total of 136 unique forecasters contribute, each providing an average of 37.1 forecasts[5]. However, there is no substantial arguement for forecasters systematically dropping out. Thus, observations with missing forecasts are excluded, and an unbalanced panel is used.

Along with individual-level forecasts, forecasts are mean aggregated across time periods to

[3] We also estimate forecaster fixed-effects and explore other possible threats to identification in the robustness checks of the results section.

[4] This is consistent with previous literature, as seen in Bordalo et al. (2020)

[5] Typically, a forecaster will stagger their forecasts rather than provide a continuous 37 forecasts

Table 1. Summary Statistics

Panel A: Individual-level Forecasts

	*E*rrors (1)		*R*evisions (2)				
t	Mean	SD	Mean	SD	Upward	Downward	No Revision
0	0.01	0.42	-0.01	0.39	0.39	0.41	-
1	0.01	0.85	-0.02	0.52	0.37	0.40	0.03
2	0.00	1.21	-0.04	0.62	0.36	0.42	0.02
3	-0.02	1.55	-0.05	0.71	0.36	0.42	0.02

Panel B: Mean Forecasts

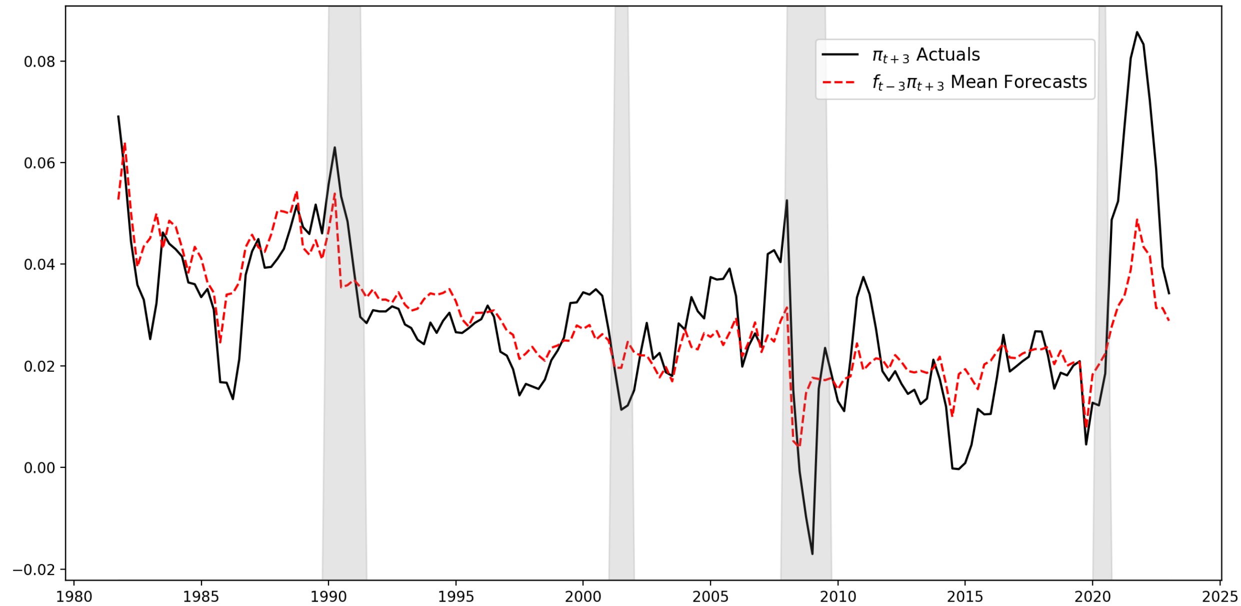
	(1)		(2)			
t	Mean	SD	Mean	SD	Upward	Downward
0	0.00	0.33	-0.01	0.28	0.46	0.54
1	-0.00	0.73	-0.03	0.37	0.46	0.54
2	-0.03	1.07	-0.05	0.43	0.45	0.55
3	-0.07	1.39	-0.07	0.49	0.44	0.56

Notes: Panel A describes the summary statistics for indivivdual-level forecasts across horizons 0-3, where horizon 3 are annualized inflation forecasts. Without loss of generality, annualized, $t+3$, inflation growth is calculated by $\left(f_t\pi_{t+2}+1\right)\left(f_t\pi_{t+3}+1\right)/400-1$. Revisions (2) are calculated by $f_t\pi_{t+3}-f_{t-1}\pi_{t+3}$ and errors (1) are calculated by $\pi_{t+3}-f_{t-1}\pi_{t+3}$. All values are percentages. Upward, Downward, and No Revision describe the share of direction for which a revision is made. Panel B describes the summary statistics for mean forecasts across horizons 0-3. Note mean forecasts will not have a no revision share as mean forecasts come from a mean aggregation of individual-level forecasts.

obtain a mean forecast data set. Table 1 provides the summary statistics for errors and revisions in horizons 0-3 for both individual-level and mean forecasts. A brief look at mean revisions and errors reveals a common pattern. Both forecasting errors and revisions are the lowest in the earliest horizon and become larger in magnitude as the forecasting horizon progresses. These statistics are slightly deviated from those of existing literature. Compared to Bordalo et al. (2020), both our mean errors and revisions are much closer to 0 with a larger variance. However, if we restrict our sample to pre-COVID, 1981III-2020II, the summary statistics are near identical to existing literature[6]. This striking feature will be examined more closely in the subsequent section. Figure 1 further underscores the distinction between samples.

[6]See Appendix B

Figure 1. Annual CPI Actuals and Mean Forecasts



Notes: This figure plots $t+3$ (annual) inflation actuals (solid black) and $t+3$ mean forecasts (dashed red).

IV. Results

i. *Mean Forecasts*

Following Coibion and Gorodnichenko (2015a), the relationship between $t+j$ mean ex-post forecast errors on $t+j$ mean ex-ante forecast revisions can be described in regression form as:

$$\pi_{t+3} - f_t\pi_{t+3} = \alpha + \beta\left(f_t\pi_{t+3} - f_{t-1}\pi_{t+2}\right) + \epsilon_{t+3|t} \tag{12}$$

As mentioned previously, this approach provides a simple method to test the null $\beta = 0$, i.e., the FIRE model, using OLS. Panel A of Table 2 reports equation (12) results for horizons 0-3 for our full sample. Estimates across forecast horizons are all significantly greater than 0, thereby rejecting FIRE. Additionally, every horizon reports positive coefficients, indicating systematic underpredictions by forecasters. Using estimates in horizon $t+3$, we can compute the structural parameters of the sticky- and noisy-information models: $\hat{\lambda} = \hat{\beta}/\left(1+\hat{\beta}\right) = 0.41$, $\hat{G} = 1/\left(1+\hat{\beta}\right) = 0.59$. These estimates are comparable to those of Coibion and Gorodnichenko

Table 2. Mean CG Regression Results

Panel A: Full Sample 1981III - 2022IV

	Forecast Horizon			
$\pi_{t+3} - f_t\pi_{t+3}$	0	1	2	3
α	0.000	0.000	0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.002)
$f_t\pi_{t+3} - f_{t-1}\pi_{t+3}$	0.621***	0.584***	0.565*	0.692*
	(0.137)	(0.240)	(0.330)	(0.368)

Panel B: Restricted Sample 1981III - 2020II

	Forecast Horizon			
$\pi_{t+3} - f_t\pi_{t+3}$	0	1	2	3
α	-0.000	-0.000	-0.002	-0.002**
	(0.000)	(0.001)	(0.001)	(0.001)
$f_t\pi_{t+3} - f_{t-1}\pi_{t+3}$	0.533***	0.295	0.120	0.194
	(0.154)	(0.181)	(0.158)	(0.184)

Notes: Both panels report OLS estimates using heteroscedasticity and autocorrelation robust standard errors with automatic lag selection (Newey & West, 1994). Let *, **, *** denote 0.10, 0.05, and 0.01 significant levels respectively.

(2015a) with a slight decrease in the "degree of information rigidities". However, if we compare these results to those of Bordalo et al. (2020), point estimates in Panel A are nearly double those of Bordalo et al. (2020). These differences are likely the result of policy regime shifts and large variance resulting from the COVID-19 pandemic[7]. Nevertheless, interpretation of these parameter estimates indicates significant information rigidities in forecasts, with forecasters updating their information set every 5-6 months according to the sticky-information model.

Panel B of Table 2 reports OLS results for horizons 0-3 in the restricted sample. These estimates are more comparable to those found in Angeletos et al. (2021), Bordalo et al. (2020), and Kohlhas and Walther (2021). Again, using the results of horizon $t+3$, we can again derive the structural parameter estimates of the sticky-information model and the noisy-information model: $\hat{\lambda} = \hat{\beta}/(1+\hat{\beta}) = 0.16$, $\hat{G} = 1/(1+\hat{\beta}) = 0.84$. Intuitively, these results make sense; when policy

[7]These differences in restricted samples are also reported in Angeletos et al. (2021) who utilize samples before and after Volker disinflation and oil shocks

Table 3. Individual-level CG Regression Results

Panel A: Full Sample 1981III - 2022IV				
	Forecast Horizon			
$\pi_{t+3} - f_t^i \pi_{t+3}$	0	1	2	3
α	-0.001***	-0.004***	-0.009***	-0.015***
	(0.000)	(0.000)	(0.000)	(0.000)
$f_t^i \pi_{t+3} - f_{t-1}^i \pi_{t+3}$	0.066**	0.021	-0.055	-0.042
	(0.030)	(0.044)	(0.051)	(0.052)

Panel B: Restricted Sample 1981III - 2020II				
	Forecast Horizon			
$\pi_{t+3} - f_t^i \pi_{t+3}$	0	1	2	3
α	-0.002***	-0.005***	-0.009***	-0.016***
	(0.000)	(0.000)	(0.000)	(0.000)
$f_t^i \pi_{t+3} - f_{t-1}^i \pi_{t+3}$	0.010	-0.178***	-0.308***	-0.315***
	(0.030)	(0.076)	(0.033)	(0.032)

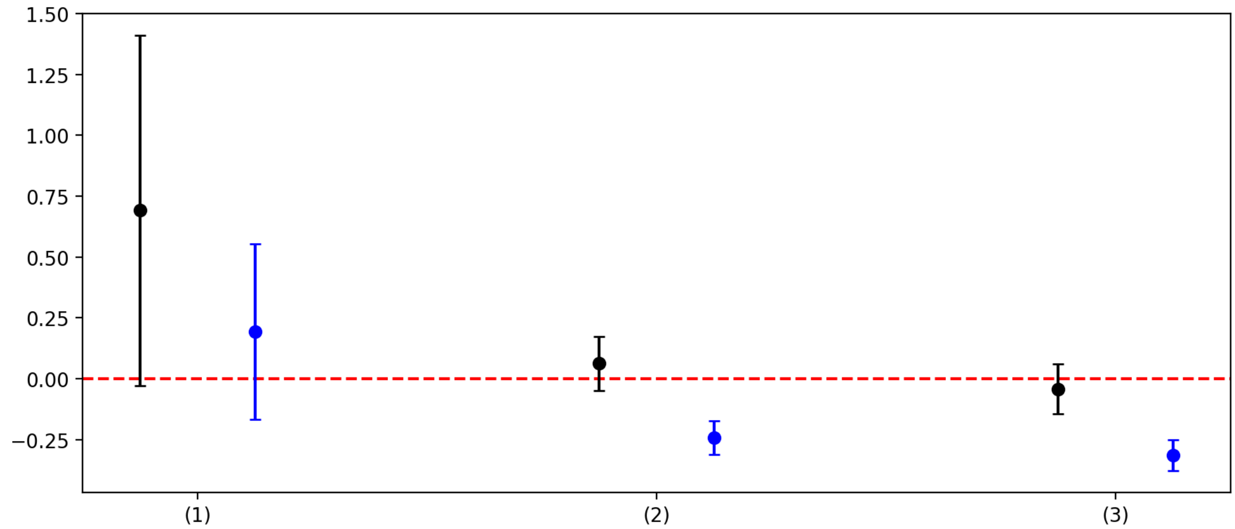
Notes: Both panels report the forecaster-level estimates with fixed-effects using clustered standard errors. Let *, **, *** denote 0.10, 0.05, and 0.01 significant levels respectively.

regime is ”calm”, forecasts are more certain, i.e., less information rigidities. Likewise, when policy regime changes dramatically and/or the state of the economy becomes uncertain, agents' forecasts are built on uncertain information, hence information rigidities are larger.

ii. Individual-level Forecasts

Unlike the sticky- and noisy-information models, estimating the structural parameters of the diagnostic model cannot be done using mean forecast revisions and errors. Instead, Table 3 reports equation 13 results with forecaster fixed-effects. Estimates in Panel A are not significantly different from zero in horizons 1-3, and therefore, we cannot reject the null in this case. However, this changes when focusing on the restricted sample (Panel B), in which estimates in 3 out of 4 horizons are significantly less than zero. This signifies overreaction among individual forecasters, the opposite of what was seen in mean forecasts.

Figure 2. Full and Restricted Sample Estimates



Notes: Column (1) shows the 95% CI for mean OLS estimates of $t + 3$ revisions for the full sample, 1981III - 2022IV (Black), and the restricted sample, 1981III - 2020II (Blue). Column (2) shows the 95% CI for the pooled OLS estimates for both samples. Column (3) shows the 95% CI for forecaster fixed-effects estimates for both samples. The horizontal red line represents the null, i.e., FIRE.

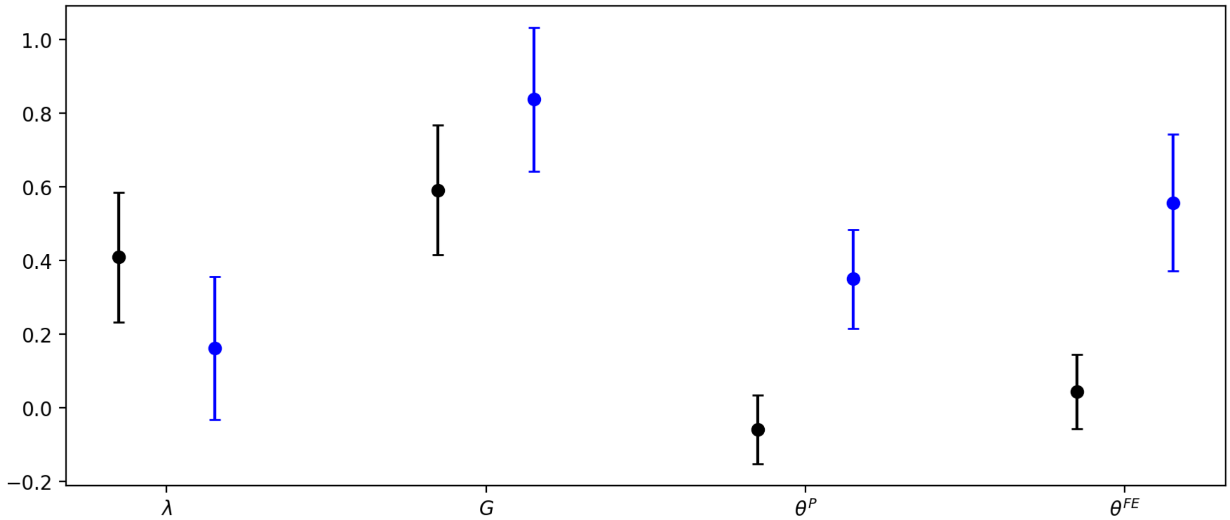
$$\pi_{t+3} - f_t^i \pi_{t+3} = \alpha + \beta \left(f_t^i \pi_{t+3} - f_{t-1}^i \pi_{t+2} \right) + \epsilon_{t+3}^i \tag{13}$$

Inverting the $t+3$ estimate, we get the structural parameter of the diagnostic model: $\theta = 0.56$. This estimate, again, is slightly higher than that of Bordalo et al. (2020), but identical to the average θ for all variables. The intuition behind the value of θ is that forecasters overreact to new information 56 percent more than a rational forecaster with noise would.

Figure 2 plots the 95% confidence intervals for coefficient estimates in each specification and sample. As expected, the variance in the full sample (black) is higher compared to the restricted sample (blue) across all estimation specifications. Focusing on the restricted sample, the significance of individual-level estimates becomes different from null in both pooled and fixed-effects specifications[8]. Bordalo et al. (2020) argue that $\theta > 0$ reconciles the distortion between individual and mean CG estimates by assuming each agent has a private signal in which no other forecaster knows. Thus the mean across forecasts has the potential to be negative. Figure 3

[8]See Appendix C for pooled OLS estimates for both samples

Figure 3. Full and Restricted Sample Model Parameters



Notes: Column (1) shows the 95% CI for the structural parameter of the Sticky-information model, λ, using $t+3$ OLS mean coefficients in the full sample, 1981III - 2022IV (Black), and the restricted sample, 1981III - 2020II (Blue). Column (2) shows the 95% CI for the structural parameter of the Noisy-information model, G, using $t+3$ OLS mean coefficients in both samples. Columns (3) and (4) show the 95% CI for the structural parameter of the Diagnostic Expectations model, θ, for pooled and fixed-effects $t+3$ coefficients in both samples.

shows the structural parameter estimates for all models, including θ estimates for both pooled and fixed-effects specifications.

iii. Robustness Checks

There are various concerns regarding the identification of β. Coibion and Gorodnichenko (2015a) and Mankiw et al. (2003) demonstrate that several variables exhibit explanatory power over mean ex-post forecast errors, including lagged annual inflation, lagged short-term interest rates, lagged change in the log of oil prices, and lagged unemployment rates. However, Coibion and Gorodnichenko (2015a) finds that the explanatory power of these variables becomes insignificant when the predictive variable is changed to ex-ante forecast revisions.

Unfortunately, due to time constraints, lagged short-term interest rates and lagged unemployment rates were not verified with the current sample. Nevertheless, Appendix D reports the coefficient estimates for mean forecasts, in both samples, with one-quarter lagged annual inflation and one-quarter lagged change in the log of oil prices (Spot West Texas Intermediate) as controls. Estimates of the control in the full sample are significantly different for both lagged inflation and lagged change in the log of oil prices. Additionally, significance in revisions is lost in both cases. In the restricted sample, we find that lagged inflation is not significant but we lose the predictive power of revisions. For lagged change in the log of oil prices, the estimate on revisions changes in both magnitude, significance, and sign. Additionally, our control is significantly different zero. These results are concerning in the context of our models, as no other variables should be able to predict forecast errors. However, it should be noted that for both regressions with controls, our error term is not orthogonal. However, this is addressed in the proceeding robustness check.

Another concern is heterogeneity in forecaster priors given agent's receive a noisy signal. To address this, we follow the approach of Bordalo et al. (2020) by regressing different horizons onto each other at the individual-level. Appendix E reports the coefficients of $t + 2$ errors on $t + 3$ revisions. For individual-level revisions, the coefficient remains negative and significant in the restricted sample. However, this changes dramatically in the full sample, where the coefficient on revisions loses significance and changes sign. This is indicative of significant noise in agent signals, most likely the result of the COVID-19 shock.

This is different if we test for heterogeneity in priors using mean forecasts. For this test, we follow Coibion and Gorodnichenko (2015a) who use lagged inflation as a control with 2 lags of log change in oil price as instruments to control for the now time variant error term. Appendix E reports the estimates for our previous OLS result and 2SLS IV results, one without lagged inflation as a control and another with (both using 2 lags of log change in oil price as instruments). In each specification across samples, the coefficient on revisions is not significant. In addition, the magnitudes of coefficients in the IV specification are much larger than that of standard OLS. These estimates are comparable to those found in Coibion and Gorodnichenko (2015a). Also notice, that our lagged inflation control is not significantly different from zero, indicating that heterogeneity in signals does not play a signficant role when forecasts are mean aggregated.

Lastly, it is evident that shocks play a significant role in shaping expectations. The substantial differences between the full and restricted samples underscore these dynamics, as illustrated by our analysis of oil prices in Appendix D. Given this, it should be noted that our coefficients in the full sample are not identified. Several papers, including those by Angeletos et al. (2021) and Coibion and Gorodnichenko (2012), attempt to understand the relationship between expectations and shocks. Specifically, Angeletos et al. (2021) argue against low-dimensional Autoregressive Moving Average Model (AMAR), favoring for the more flexible local projection method of Jordà (2005). Their research attempts to reconcile conflicting findings from prior literature and different samples, revealing that forecasters tend to underreact in the initial periods following a shock but then overshoot thereafter. This further underscores the necessity for investigation into the behavioral patterns of expectations, particularly in ”unfamiliar and extreme” events.

V. Conclusion

Expectations form the bedrock of economic decision-making, exerting profound influence over behaviors, prices, and policy outcomes. This thesis embarked on a comprehensive exploration, revisiting and evaluating expectation formation models within the context of contemporary inflation data. Examination of mean forecasts revealed compelling evidence of consistent underpredictions by forecasters, evidenced by positive coefficients observed across forecast horizons. The rejection of the null hypothesis in favor of the Full-information Rational Expectations (FIRE) model underscored the presence of information rigidities, with forecasters updating their information sets at intervals of approximately 5-6 months. A deeper analysis of individual-level forecasts uncovered intriguing patterns of overreaction among individual forecasters, in contrast to the trends observed in mean forecasts. Estimation of structural parameters in the diagnostic expectations model further shed light on the extent of ”representativeness” in individual forecasts.

While these results aligned with previous literature largely under the restricted sample, deviations were noted under the full sample. This is likely the result of the COVID-19 shock, which is further underscored in previous literature. We further confirm the significance of noise in individual-level forecasts resulting in heterogeneity in agent priors. Indeed, these deviations in estimates across ”calm” and more ”shock-specific” samples are found in prior literature including Angeletos et al. (2021). While results of our control table are concerning, it should be taken with a grain of salt. Introducing lagged variables in our regression induces a time component in our error term that likely effects results. Given the significance of the COVID-19 shock, exploring these deviations in a more robust manner could yield valuable insights.

In essence, expectations play a pivotal role in understanding both macro and microeconomic dynamics. This thesis aimed to scrutinize the current biases of inflation expectations in light of contemporary data and assess the explanatory power of expectation models. The deviations identified in estimates within the full sample open up an intriguing avenue for future research, emphasizing the necessity of considering heterogeneity and idiosyncratic behaviors in expectation formation.

References

Andolfatto, D., Hendry, S., & Moran, K. (2008). Are inflation expectations rational? *Journal of Monetary Economics*, *55*(2), 406–422. https://doi.org/10.1016/j.jmoneco.2007.07.004

Angeletos, G.-M., Huo, Z., & Sastry, K. (2021). Imperfect Macroeconomic Expectations: Evidence and theory. *Nber Macroeconomics Annual*, *35*, 1–86. https://doi.org/10.1086/712313

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2016). Stereotypes. *Quarterly Journal of Economics*, *131*(4). https://doi.org/10.1093/qje/qjw029

Bordalo, P., Gennaioli, N., La Porta, R., & Shleifer, A. (2019). Diagnostic expectations and stock returns. *The Journal of Finance*, *74*(6), 2839–2874. https://doi.org/10.1111/jofi.12833

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020). Overreaction in Macroeconomic Expectations. *American Economic Review*, *110*(9), 2748–2782. https://doi.org/10.1257/aer.20181219

Calvo, G. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, *12*(3). https://doi.org/10.1016/0304-3932(83)90060-0.

Capistrán, C., & Timmermann, A. (2009). Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking*, *41*(2-3), 365–396. https://doi.org/10.1111/j.1538-4616.2009.00209.x

Carlson, J. A. (1977). A study of price forecasts. *RePEc: Research Papers in Economics*, 27–56. https://www.nber.org/chapters/c10501.pdf

Chen, J., Górnicka, L., & Žďárek, V. (2022). Biases in Survey Inflation Expectations: Evidence from the Euro Area. *IMF working paper*, *2022*(205), 1. https://doi.org/10.5089/9798400204401.001

Coibion, O., & Gorodnichenko, Y. (2012). What Can Survey Forecasts Tell Us about Information Rigidities? *Journal of Political Economy*, *120*(1), 116–159. https://doi.org/10.1086/665662

Coibion, O., & Gorodnichenko, Y. (2015a). Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation. *American Economic Journal: Macroeconomics*, *7*(1), 197–232. https://doi.org/10.1257/mac.20130306

Coibion, O., & Gorodnichenko, Y. (2015b). Information Rigidity and the expectations formation Process: a simple framework and new facts. *The American Economic Review*, *105*(8), 2644–2678. https://doi.org/10.1257/aer.20110306

Coibion, O., Gorodnichenko, Y., & Kamdar, R. (2018). The formation of expectations, inflation, and the Phillips curve. *Journal of Economic Literature*, *56*(4), 1447–1491. https://doi.org/10.1257/jel.20171300

Coibion, O., Gorodnichenko, Y., & Kumar, S. (2018). How Do Firms Form Their Expectations? New Survey Evidence. *American Economic Review*, *108*(9), 2671–2713. https://doi.org/10.1257/aer.20151299

Coibion, O., Gorodnichenko, Y., & Ropele, T. (2019). Inflation Expectations and firm Decisions: new causal evidence*. *The Quarterly Journal of Economics*, *135*(1), 165–219. https://doi.org/10.1093/qje/qjz029

Grant, A., & Thomas, L. B. (1999). Inflationary expectations and rationality revisited. *Economics Letters*, *62*(3), 331–338. https://doi.org/10.1016/s0165-1765(98)00244-4

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *The American Economic Review*, *95*(1), 161–182. https://doi.org/10.1257/0002828053828518

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237–251. https://doi.org/10.1037/h0034747

Kohlhas, A., & Walther, A. (2021). Asymmetric Attention. *American Economic Review*, *111*(9), 2879–2925. https://doi.org/10.1257/aer.20191432

Lucas, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, *1*, 19–46. https://doi.org/10.1016/s0167-2231(76)80003-6

Malmendier, U., & Nagel, S. (2015). Learning from Inflation Experiences *. *The Quarterly Journal of Economics*, *131*(1), 53–87. https://doi.org/10.1093/qje/qjv037

Mankiw, N. G., & Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics*, *117*(4), 1295–1328. https://doi.org/10.1162/003355302320935034

Mankiw, N. G., & Reis, R. (2006). Pervasive stickiness. *The American Economic Review*, *96*(2), 164–169. https://doi.org/10.1257/000282806777211937

Mankiw, N. G., Reis, R., & Wolfers, J. (2003). Disagreement about Inflation Expectations. *NBER*, *18*, 209–248. https://doi.org/10.1086/ma.18.3585256

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, *29*(3), 315. https://doi.org/10.2307/1909635

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, *61*(4), 631–653. https://doi.org/10.2307/2297912

Paquet, A. (1992). Inflationary expectations and rationality. *Economics Letters*, *40*(3), 303–308. https://doi.org/10.1016/0165-1765(92)90009-n

Pesando, J. E. (1975). A note on the rationality of the Livingston price expectations. *Journal of Political Economy*, *83*(4), 849–858. https://doi.org/10.1086/260359

Sargent, T. J., & Lucas, R. E. (1979). After Keynesian Macroeconomics. *Federal Reserve Bank of Minneapolic Quarterly Review*, *3*(2), 1–16. https://doi.org/10.21034/qr.321

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690. https://doi.org/10.1016/s0304-3932(03)00029-1

Woodford, M. (2003). *Imperfect common knowledge and the effects of monetary policy*. Princeton University Press.

Zarnowitz, V., & Braun, P. (1992). Twenty-two years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance. *RePEc: Research Papers in Economics*, 11–94. https://econpapers.repec.org/paper/nbrnberwo/3965.htm

Appendix

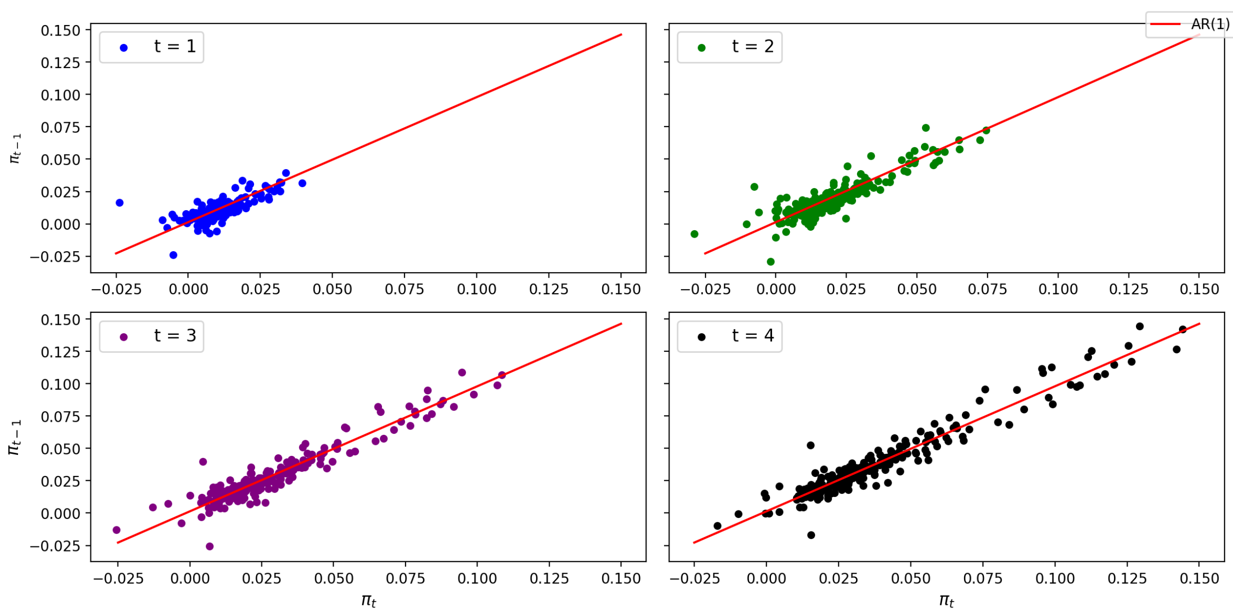
Appendix A: Autoregression Results & Figures

Table 1. AR(1) Coefficient & Root

	ϕ (1)	ρ (2)
$\pi_t = \alpha + \phi\pi_{t-1}$	0.966 (0.018)	1.035

Notes: Column (1) reports the coefficient of the AR(1) regression for annual inflation using conditional MLE. Column (2) reports the corresponding AR(1) real root.

Figure 1. Persistence for $t+1$ to $t+3$ Actuals



Notes: Each panel plots growth rates for actuals in each time horizon. The red line in each panel is the corresponding AR(1) model.

Appendix B: 1981III - 2020II Summary Statistics

Table 2. Summary Statistics

Panel A: Individual-level Forecasts

	*E*rrors (1)		*R*evisions (2)				
t	Mean	SD	Mean	SD	Upward	Downward	No Revision
0	-0.02	0.39	-0.04	0.37	0.36	0.44	-
1	-0.07	0.75	-0.07	0.49	0.34	0.42	0.04
2	-0.15	1.01	-0.09	0.58	0.34	0.41	0.03
3	-0.23	1.22	-0.11	0.67	0.34	0.44	0.02

Panel B: Mean Forecasts

	(1)		(2)			
t	Mean	SD	Mean	SD	Upward	Downward
0	-0.02	0.30	-0.03	0.26	0.43	0.57
1	-0.08	0.64	-0.07	0.34	0.43	0.57
2	-0.16	0.87	-0.09	0.40	0.42	0.58
3	-0.27	1.06	-0.12	0.46	0.41	0.59

Notes: See Table 1 for details of calculations

Appendix C: Individual-level Pooled OLS Results

Table 3. Individual-level CG Regression Results

Panel A: Full Sample 1981III - 2022IV

	Forecast Horizon			
$\pi_{t+3} - f_t^i \pi_{t+3}$	0	1	2	3
α	0.000*	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$f_t \pi_{t+3} - f_{t-1} \pi_{t+3}$	0.077**	0.048	0.022	0.062
	(0.028)	(0.043)	(0.053)	(0.057)

Panel B: Restricted Sample 1981III - 2020II

	Forecast Horizon			
$\pi_{t+3} - f_t \pi_{t+3}$	0	1	2	3
α	-0.000**	-0.000***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
$f_t \pi_{t+3} - f_{t-1} \pi_{t+3}$	0.0128	-0.147***	-0.253***	-0.242***
	(0.029)	(0.036)	(0.035)	(0.036)

Notes: Both panels report the pooled OLS estimates for mean forecasts using clustered standard errors. Let *, **, *** denote 0.10, 0.05, and 0.01 significant levels respectively.

Appendix D: Mean Estimates with Controls

Table 4. Mean CG Regression Results with Controls

Panel A: Full Sample 1981III - 2022IV			
$\pi_{t+3} - f_t\pi_{t+3}$	No Control	Lagged Annual Inf.	Lagged Change in Log of Oil Prices
α	-0.000	-0.011	-0.001
	(0.002)	(0.004)	(0.002)
$f_t\pi_{t+3} - f_{t-1}\pi_{t+3}$	0.692*	0.228	0.250
	(0.368)	(0.256)	(0.418)
γ_{t-1}	-	0.354*	0.026***
		(0.157)	(0.006)

Panel B: Restricted Sample 1981III - 2020II			
$\pi_{t+3} - f_t\pi_{t+3}$	No Control	Lagged Annual Inf.	Lagged Change in Log of Oil Prices
α	-0.002*	-0.007*	-0.003**
	(0.001)	(0.003)	(0.001)
$f_t\pi_{t+3} - f_{t-1}\pi_{t+3}$	0.194	0.089	-0.233
	(0.184)	(0.265)	(0.324)
γ_{t-1}	-	0.157	0.023***
		(0.125)	(0.008)

Notes: Both panels report OLS estimates using heteroscedasticity and autocorrelation robust standard errors with automatic lag selection (Newey and West (1994)). Oil prices are spot prices from West Texas Intermediate FRED data. Let *, **, *** denote 0.10, 0.05, and 0.01 significant levels respectively.

Appendix E: Heterogeneity in Signals

Table 5. Individual-level CG estimates (Different Horizons)

$\pi_{t+2} - f_t^i \pi_{t+2}$	Full Sample	Restricted Sample
α	0.000	-0.002***
	(0.000)	(0.000)
$f_t^i \pi_{t+3} - f_{t-1}^i \pi_{t+3}$	0.112	-0.256***
	(0.069)	(0.043)

Notes: Panel A reports OLS estimates using heteroscedasticity and autocorrelation robust standard errors with automatic lag selection (Newey and West (1994)). Panel B reports pooled OLS estimates for mean forecasts using clustered standard errors. Let *, **, *** denote 0.10, 0.05, and 0.01 significant levels respectively.

Table 6. Mean CG IV 2SLS Results

Panel A: Full Sample 1981III - 2022IV			
		IV	
$\pi_{t+3} - f_t \pi_{t+3}$	OLS	(1)	(2)
α	0.000	0.001	0.000
	(0.002)	(0.002)	(0.001)
$f_t \pi_{t+3} - f_{t-1} \pi_{t+3}$	0.692	1.367	1.383
	(0.368)	(1.060)	(0.929)
π_{t+3}^{t-1}			0.015
			(0.016)

Panel A: Full Sample 1981III - 2022IV			
		IV	
$\pi_{t+3} - f_t \pi_{t+3}$	OLS	(1)	(2)
α	-0.002*	-0.000	0.002
	(0.001)	(0.001)	(0.004)
$f_t \pi_{t+3} - f_{t-1} \pi_{t+3}$	0.194	0.687	0.666
	(0.184)	(0.614)	(0.626)
π_{t+3}^{t-1}			-0.005
			(0.009)

Notes: Both panels report 2SLS IV regression estimates using 2 lags of log oil price changes. All estimates are computed using heteroscedasticity and autocorrelation robust standard errors with automatic lag selection (Newey and West (1994)). Let *, **, *** denote 0.10, 0.05, and 0.01 significant levels respectively.

Appendix F: All Time Horizons

Figure 2. Full Sample (1981III-2022IV) Estimates (All Horizons)

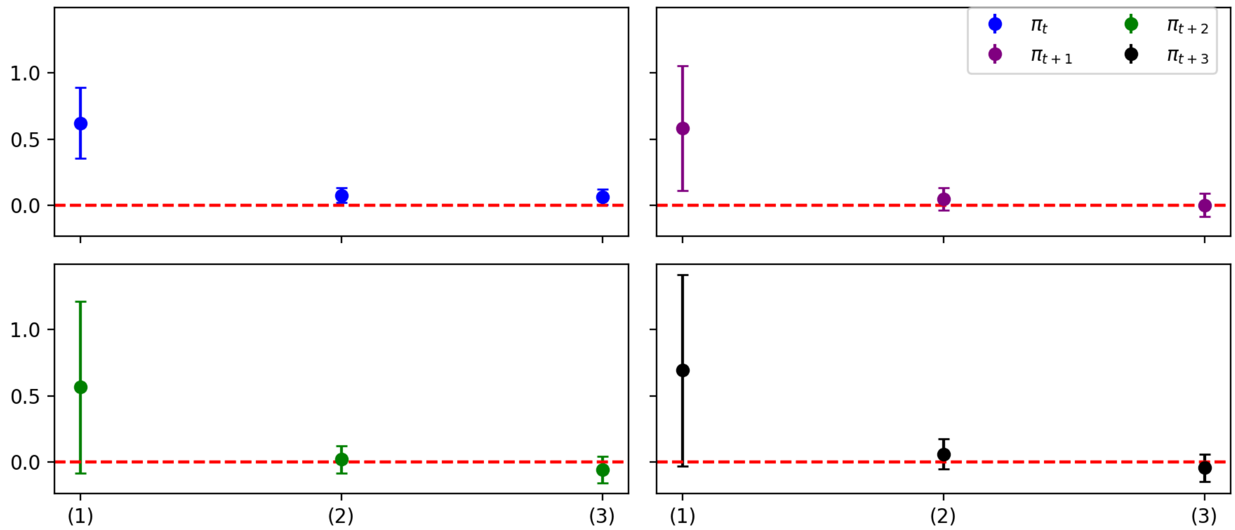


Figure 3. Restricted Sample (1981III-2020II) Estimates (All Horizons)