Information Rigidity and Economic Uncertainty: A New Theory and Stylized Facts*

Job Market Paper Draft†

Romina Kazandjian‡

December 2, 2019

I propose a structural micro-founded sticky-noisy information model with high- and low-uncertainty regimes. Agents first appraise the state of uncertainty and only spend resources to update their inflation expectations if they perceive uncertainty as sufficiently high. Time-varying uncertainty affects expectation formation through two direct channels: 1) the *wake-up call effect*, which causes agents to pay more attention, increasing their quantity of information; and 2) the *wait-and-see effect*, which decreases their quality of information and prompts them to put less weight on new noisier information. Using structural estimation of alternative models with information frictions, I find that accounting for the indirect *state-dependence* channel, the proposed innovation of the model, better explains the observed information rigidity, since it considers the interaction between the two direct effects. A substantial amount of information rigidity is due to inattention, leaving ample room for policymakers to employ frequent, direct, and simple forward guidance to pierce the veil of inattention.

Keywords: Attention, inflation expectations, information rigidity, state-dependence, survey forecasts, uncertainty

JEL Codes: C3, C5, D8, E3, E5

*Special thanks to dissertation advisor Xuguang Simon Sheng and dissertation committee members Evan Kraft, Tucker McElroy, and Jeff Fuhrer, as well as to Ekaterina Peneva, Sandra Rivera, Saad Ahmad, Maria Luengo-Prado, Geoff Tootell, Daniel Cooper, Bent Sorenson, Jenny Tang, Slavik Sheremirov, Anat Bracha, Joseph Guse, and seminar participants at the Federal Reserve Bank of Boston, the 2018 SEA Annual Meetings, the 2019 Young Scholars Initiative North America Convening, and the 2019 EEA Annual Meetings for helpful comments and suggestions. Thanks to the Federal Reserve Bank of Boston, the US International Trade Commission, and the AEA for financial support through the AEA Summer Economics Fellows Program.

†For most recent draft see: `www.rominakazandjian.com/research`.

‡Ph.D. Candidate, American University, Department of Economics, 4400 Massachusetts Avenue NW, Washington DC, 20016; +1-516-782-4578; `romina.kazandjian@student.american.edu`; `www.rominakazandjian.com`.

1 Introduction

How economic agents form expectations about the future is of great macroeconomic importance because it influences decision-making and thus, economic behavior. Based on their expectations about the future, consumers make decisions about how much to save and consume, firms make decisions about setting prices and hiring workers, investors make decisions about how much risk they are willing to bear to achieve their target returns, and central bankers decide on appropriate monetary policy strategies. Specifically, agents' inflation expectations impact business cycle dynamics and the effectiveness of monetary policy. The Phillips curve describing the relationship between inflation and unemployment features inflation expectations as one of the main drivers of inflation, which underscores the importance of well-anchored expectations as a primary goal of monetary policy. This study furthers this objective by shedding new light on how economic uncertainty affects people's inflation expectations.

Especially since the Great Recession, the economics profession has been publicly criticized for its unrealistic simplifying assumptions, prominent among which is the assumption of rational expectations, whereby economic agents form identical expectations using their full information of the true underlying model of the economy. The full-information rational expectations hypothesis has for a long time been subject to both theoretical and empirical rejections (Caskey, 1985; Frankel and Froot, 1987; Sargent, 1993; Branch, 2004; Pesaran and Weale, 2006; Coibion and Gorodnichenko, 2012, 2015; Dovern et al., 2015; Fuhrer, 2018). Since Lucas (1972) and Kydland and Prescott (1982), a growing research field has attempted to modify the full-information rational expectations assumption by introducing information imperfections to economic models. Coibion and Gorodnichenko (2015) have framed this research agenda around the concept of *information rigidity*, which is the existence of information frictions despite agents having rational expectations, or at least taking into account all currently available to them information. Among others, in this seminal paper the authors assert that information rigidity is state-dependent; yet, there are few existing state-dependent models Reis (2006); Van Nieuwerburgh and Veldkamp (2006); Gorodnichenko (2008); Woodford (2009); Maćkowiak and Wiederholt (2012); Nimark (2014) and to my knowledge, none that are based on regime-switching economic uncertainty.

Hence, this research contributes to our understanding of how agents form expectations by introducing the indirect effect of uncertainty on information rigidity. Specifically, I propose a state-dependent structural micro-founded sticky-noisy information model with high-and low-uncertainty regimes. Agents are both rationally inattentive and faced with noisy information. The main innovation is that the attention that agents allocate to acquiring and processing information (quantity of information), as well as the noisiness in the information they glean (quality of information) are driven by time-varying economic uncertainty. There are interactions between the attention and noisiness effects, which moderate their influence on aggregate information rigidity. Which of these opposite effects dominates is ambiguous *a priori* and depends on the parameters of the model. I answer this question by conducting structural estimation using Simulated Method of Moments (SMM) of the proposed state-dependent model and the alternative models with information frictions.

The proposed theoretical model is micro-founded: agents solve an optimization problem in order to minimize the cost of attention. In the solution, attention is a positive function of economic uncertainty because collecting and analyzing information is costly in terms of time and resources. When economic uncertainty is low, it may be optimal for agents *not* to update their expectations every period. Since macroeconomic conditions are less volatile, the cost in terms of forecast accuracy of *not* updating information frequently is relatively lower, so agents optimally pay less attention. In contrast, in high-uncertainty states, this cost increases, as previous forecasts can quickly become obsolete in light of volatile economic conditions. Thus, agents have an incentive to pay closer attention to macroeconomic conditions in these periods. This is the attention or *wake-up call effect*. At the same time, when uncertainty is relatively low, the signals agents glean have relatively higher precision and are optimally trusted more. Conversely, greater volatility makes it harder to make accurate predictions. With higher uncertainty, the lower precision of information signals prompts agents to optimally put less weight on the incoming information and have less faith in it. This is the noisiness or *wait-and-see effect*.[1] In high- (low-) uncertainty periods, there is a higher (lower) demand for updated accurate information (attention effect) but lower (higher) supply (noisiness effect). Therefore, the attention and the noisiness effect directly influence aggregate information rigidity in opposite directions.

To capture the attention effect, I use the sticky information model (Mankiw and Reis, 2002; Reis, 2006; Maćkowiak and Wiederholt, 2015), whereby agents have to pay fixed costs to acquire and process information in order to update their forecasts, so that every period a fraction of agents is rationally inattentive. I introduce the noisiness effect by unifying

[1] The *wait-and-see* notion dates back to Bernanke (1983), who originally proposed a theory of irreversible choice under uncertainty. According to it, uncertainty increases the value of waiting to make lasting investment decisions until the uncertainty is resolved and reliable information is obtained, so that short-run investment cycles form.

this framework with the noisy information model (Woodford, 2002; Sims, 2003; Maćkowiak and Wiederholt, 2009), which posits that agents continuously update their information but face a signal-extraction problem due to noise in the signal they glean. The proposed state-dependent model improves upon existing hybrid sticky-noisy information models (Andrade and Le Bihan, 2013; Giacomini et al., 2020), which consider the effects of rational inattention and noisy information as independent drivers of information rigidity, thus overestimating it. Whereas the attention and noisiness channels directly contribute to information rigidity, state-dependence indirectly moderates the interactions between these two direct effects on aggregate information rigidity. To my knowledge, these interactions are entirely ignored in the existing literature. Introducing state-dependence to this theoretical setup is thus an innovation and a major contribution to the literature. This research promises to add granularity to our understanding of the expectation formation process of economic agents and hence, its impact on macroeconomic dynamics in the future, especially during current and forthcoming episodes of heightened uncertainty.

The proposed model posits that economic agents form expectations in two stages. First, before deciding whether to engage in a comprehensive forecast re-estimation, agents conduct a preliminary step of assessing the hidden state of economic uncertainty. Using and interpreting free public information, agents calculate the probabilities of being in high- or low-uncertainty regimes, which is approximated with a Markov-switching model. Based on this inference and the solution of the agent's optimization problem, agents update their forecasts with probability proportional to the probability to be in the high-uncertainty state. In the second stage of the model, only the attentive agents update their forecasts, while the rest simply carry forward their previous predictions. This stage is approximated with a Kalman filter as in a noisy information model: attentive agents optimally assign weights to the incoming noisy information and their previous estimate according to the Kalman gain.[2] I do not assert that economic agents *actually* run a Markov-switching model or a Kalman filter to update their forecasts. Instead, these algorithms merely approximate the agents' decision-making process, so that deciding whether uncertainty is high or low is less costly than re-estimating a full forecast every period.

Structural estimation of alternative models with information frictions using SMM demon-

[2]The Kalman gain represents the relative importance given to the error in the incoming data versus that of the existing estimate. If the Kalman gain is relatively high, more weight is placed on the incoming observations and the existing estimates are unstable, whereas if it is relatively low, the incoming observations are perceived to be less accurate and the filter puts more weight on the existing estimates, which are then more stable. The specification of the Kalman gain is presented in equation (4.10).

strates that the proposed state-dependent model is superior to the hybrid and sticky information models, while preserving some of the desirable features of the noisy information model. Moreover, the proposed model is more appealing than the noisy information model due to its ability to account for inattention observed in survey data. Hence, all three channels, the direct attention and noisiness and the indirect state-dependence effects need to be included in models with information frictions, so that they match survey expectations. Structural estimations further suggest that the attention effect clearly dominates the noisiness effect: more than half of the estimated information rigidity is due to inattention, which is good news for policymakers. It leaves ample room for monetary policy, which can affect the quantity of information available to economic agents by employing frequent, direct, and simple forward guidance. For instance, a creative recent initiative by the Bank of Jamaica anchors expectations by raising awareness about its inflation-targeting policy through reggae music.

Next, Section 2 describes the empirical data and Section 3 provides stylized facts and reduced-form empirical tests to motivate the theory. Prominently, Section 4 outlines the proposed theoretical model and Section 5 describes its structural estimation. Section 6 compares the estimation results of alternative models with information frictions. Section 7 provides results of the relative contributions of the direct attention and noisiness, and the indirect state-dependence channels to information rigidity. Finally, Section 8 offers concluding remarks and discusses the relevance of these findings to monetary policy.

2 Data

The survey data of quarterly forecasts of the GDP deflator[3] by individual professional forecasters are from the US Survey of Professional Forecasters (SPF), collected by the Federal Reserve Bank of Philadelphia[4] in the period Q4 1968 - Q2 2019. The sample size varies between 9 and 131 individual forecasters, with an average of 46 forecasters every quarter and includes 443 different individual forecasters in total, 70 of whom report forecasts in at least 40 of the 203 total periods. All of the presented stylized facts and empirical tests are robust to constraining the sample of forecasters to this core sample, so the subsequent analysis uses the full sample. The panel data set is unbalanced, as forecasters enter, exit, and merge

[3]Forecasts are of the level of chain-weighted GDP price index (GDP implicit deflator), seasonally adjusted with varying base year. Prior to 1992, the forecasts are for the GNP implicit deflator.

[4]Prior to Q2 1990, the SPF was conducted by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER).

during the survey period. Figure A.1 presents the frequency of forecaster participation in the sample and Figure A.2 shows time series of realized second- and final-release values of inflation vs. mean SPF forecasts.

The use of survey data on professional forecasters is dictated by data availability and the fact that as some of the most informed agents in the economy whose principal job is to make economic predictions, they serve as a conservative benchmark, or a lower bound, for the aggregate degree of information rigidity in the economy. Moreover, as documented by Carroll (2003), the predictions of professional forecasters spread throughout the economy and influence the expectations of firms and consumers, which are theoretically more relevant on the macro level. Specifically, the use of forecasts of GDP deflator inflation are used due to the much longer available time series dating back to Q4 1968, as compared to other inflation indicators (CPI starting in Q3 1981 and PCE starting in Q1 2007).

The surveys are conducted with a variety of financial and non-financial service providers for their forecasts of a plethora of macroeconomic variables. The SPF data includes industry classifications for about 53 percent of forecasts, of which 42 percent are financial service providers (insurance, investment banking, commercial banking, etc.) and 58 percent are non-financial service providers (manufacturers, universities, forecasting firms, etc.). The surveying process is the same for every respondent: the survey questionnaire includes recent historical data from the Bureau of Economic Analysis (BEA) *advance report* including the first estimate of GDP (and components) for the previous quarter, which is released at the end of the first month of each quarter and the most recent report of other governmental statistical agencies. Forecasters submit their predictions by the second to third week of the middle month of each quarter, so their information sets include the data reported in the *advance report*. More information about the SPF is available at `www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters`.

The SPF forecasts used in this study are of the "fixed-horizon" data structure, whereby forecasters give predictions on a rolling basis for the current and the next four quarters. The summary statistics and empirical results presented here have been checked for robustness using forecast data collected by *Consensus Economics, Inc.* in the period October 1989 - December 2014. This data is monthly and of the alternative "fixed-event" type: every month, forecasters are asked to submit their annual forecasts for the current and next calendar years. However, the "fixed-horizon" data structure of the SPF is preferred because unlike the "fixed-event" data, it isolates time variation and removes the seasonal effect of diminishing horizons on agents' information sets.

I use one-quarter ahead individual forecasts of GDP deflator inflation, given by the following equation:

$$\hat{y}_{t+1|t} = \left[\left(\frac{P_{i,t+1}}{P_{i,t}} \right)^4 - 1 \right] \times 100. \tag{2.1}$$

I focus on one-quarter ahead forecasts, since they represent short-run expectations that are more liable to the influence of spikes in contemporaneous economic uncertainty than medium- and long-run predictions. Moreover, as pointed out by Ryngaert (2017), forecast errors at horizons greater than one present an endogeneity problem, so that the empirical tests of information rigidity presented in Section 3 could not be estimated by OLS for these longer horizons. I also duplicate all summary statistics and empirical results using nowcasts, the agents' expectations about GDP deflator inflation for the current period formed in the current period.[5] Data on the realizations of GDP deflator inflation used in the calculation of forecast and errors (actual minus forecasted values) are obtained from the Federal Reserve Bank of Philadelphia's real-time data set, including first, second, third, and most recent (final) release of data.

In line with Bernanke (1983), Romer (1990), and Bloom (2009), uncertainty is construed as second-moment shocks to economic activity. Specifically, two main indicators are used to approximate economic uncertainty. First, the economic policy uncertainty (EPU) index is based on newspaper coverage frequency and specifically proxies for changes in policy-related economic uncertainty. It is compiled by Baker et al. (2016) and is constructed as a weighted average of three components: 1) newspaper coverage of economic policy-related uncertainty; 2) expiration of federal tax code provisions; and 3) disagreement among forecasters. The EPU is available online at `www.policyuncertainty.com`. Since the baseline EPU is only available since 1985, I use the historical EPU index, which is based on newspaper coverage of six rather than ten large newspapers as in the baseline, to extend the time series to cover the sample period Q4 1968 - Q2 2019. I normalize the historical EPU to the same mean and variance as the baseline EPU index for the period when they overlap (January 1985 - December 2014), using the methodology of Bloom (2009). Second, as a robustness check, I use the VXO, which is the Chicago Board Options Exchange index of implied volatility based on trading of S&P100 options for a 30-day horizon. It is used as a proxy for the market's short-term expectation of stock market volatility. The VXO is obtained from the Federal Reserve Economic Data (FRED) and is available from 1986. Prior values are calculated from

[5] Results not shown but available upon request.

actual monthly returns volatilities (monthly standard deviations, again normalized to the same mean and variance as the VXO index for the period when they overlap, as per Bloom (2009)).

The EPU and VXO indices have significant correlation of 0.37; yet, each introduces its own independent variation. For instance, the VXO is more reflective of episodes of short-term financial-sector uncertainty, such as the Asian Financial Crisis and the Lehman Brothers collapse, whereas the EPU is more sensitive to wars, elections, political disagreement over taxation and spending, etc., which involve major political concerns, in addition to stock market volatility. The EPU is preferred to the VXO because it is 'more exogenous' from the expectation formation process of professional forecasters, many of whom are financial-sector institutions. I urge caution in interpreting the presented results for the period prior to 1985-86 due to the less reliable quality of the EPU and VXO data. Quarterly values of the extended EPU and VXO indices are 3-month averages of monthly values. Finally, for robustness, I also use a dummy indicator for recessionary periods, as defined by the National Bureau of Economic Research (NBER) and obtained from FRED. Quarterly values of the recession indicator are 3-month averages of monthly values.

3 Motivation

3.1 Stylized Facts

The following observations document the presence of uncertainty-based state-dependence in the SPF data and motivate the introduction of regime-switching in the agents' expectation formation process. Table 1 presents descriptive statistics on the mean forecast error (actual realization - mean forecast),[6], empirical attention λ_t (the proportion of forecasters updating their forecasts each period), and forecaster disagreement (cross-sectional standard deviation of individual forecasts). Figure A.4 visually shows these stylized facts.

The mean forecast errors are larger in absolute value during HU than LU periods using all releases of realized inflation data. The difference in means between LU and HU periods is significant at least at the 10 percent level for all forecast errors. The negative values during HU periods suggest that on average, forecasters overestimate inflation during high-

[6] I present stylized facts for the forecast errors using first, second, third, and final release of actual data from the Federal Reserve Bank of Philadelphia's real-time data set.

Table 1: Summary Statistics Motivating State-dependence

Panel A: Mean Forecast Errors					
	Full Sample	LU	HU	Abs. Difference	AR(1) Coeff.
First Release	-0.090	0.032	-0.367	0.399*	0.341***
	(0.098)	(0.132)	(0.109)	(0.213)	(0.067)
Second Release	-0.021	0.112	-0.337	0.449*	0.383***
	(0.105)	(0.141)	(0.113)	(0.229)	(0.066)
Third Release	0.004	0.139	-0.313	0.452**	0.370***
	(0.104)	(0.139)	(0.115)	(0.226)	(0.066)
Final Release	0.011	0.121	-0.241	0.363*	0.543***
	(0.090)	(0.121)	(0.102)	(0.195)	(0.059)
Panel B: Empirical Attention λ_t (Fraction of Nowcast and Forecast Updaters)					
	Full Sample	LU	HU	Abs. Difference	AR(1) Coeff.
Mean λ_t	0.931	0.921	0.953	0.031**	0.996***
	(0.006)	(0.008)	(0.004)	(0.013)	(0.006)
Panel C: Forecaster Disagreement					
	Full Sample	LU	HU	Abs. Difference	AR(1) Coeff.
Standard Dev.	1.026	1.167	0.706	0.461***	0.877***
	(0.049)	(0.065)	(0.039)	(0.101)	(0.033)
Observations	199-202	139-140	59-62		

Standard errors in parentheses *** $p< 0.01$, ** $p<0.05$, * $p<0.1$.

The table shows the means of the full sample, LU and HU sub-samples with standard errors in parenthesis. Absolute differences are estimated with two-tailed means t-tests during LU vs. HU periods. AR(1) coefficients are obtained from regressions on 1-quarter lag. Mean forecast errors are obtained by subtracting the individual quarterly forecasts from the actual realizations and averaging across forecasters. Empirical attention λ_t is calculated as the proportion of forecasters who update their forecasts and nowcasts out of all forecasters in the sample. Forecaster disagreement refers to the cross-sectional standard deviation of forecasts.

volatility episodes. Moreover, their mean forecast errors during LU periods are much smaller in absolute value and positive, implying that forecasters underestimate these variables during less volatile times, albeit to a lesser degree. These findings suggest that when economic uncertainty increases forecast accuracy decreases, despite agents paying more attention and updating their forecasts at higher rates. The significant AR(1) coefficients demonstrate the predictability of mean forecast errors. Figure A.5 confirms that mean forecast errors are fairly persistent in the inflation expectations of professional forecasters.

Similarly, empirical attention λ_t[7] is higher during HU episodes, confirming that agents indeed pay more attention to inflation during periods of heightened uncertainty. This result

[7] Missing values of forecasts and nowcasts of GDP deflator inflation are interpreted as "not updating." Whenever forecasters report a forecast and nowcast that are different from the previous period's, this is interpreted as "updating." Hence, in the first period when a forecaster joins the survey, she is counted as "updating." Removing the effect of new entrants does not substantially change the findings. Forecasters are considered "attentive" only when they update both their forecast and nowcast to avoid measurement errors and focus on intentional forecast updating resulting from forecast modeling, which produces both forecasts and nowcasts.

is significant at the 5 percent level. Figures A.4 and A.7 present graphically the significant difference in the means of attention during LU and HU periods. Figure A.7 also presents its standard deviation during the LU and HU regimes. Attention is less volatile during HU periods, since forecasters prevalently pay more attention when uncertainty is relatively high. Figure A.3 confirms this by showing that the distribution of attention during HU periods includes many more forecasters paying close to full attention, compared to LU periods and the full sample. Attention also varies over time and with the state of economic uncertainty. It is never complete for long periods of time ($\lambda_t \neq 1$), as shown in Figure A.6. The fact that attention is not complete fits with the sticky information model; however, the empirical observation that it varies over time nonrandomly (high and significant AR(1) coefficient in Table 1) does not. On the other hand, the observation that attention is not always complete cannot be reconciled with the noisy information model. Hence, neither the canonical sticky, nor the basic noisy information model on its own is able to account for these stylized facts.

Finally, the finding that forecaster disagreement is lower during HU than LU periods is driven by the sample period. As shown in Figure A.8, disagreement is higher prior to the late 1980s and declines thereafter, as noted by Coibion and Gorodnichenko (2012). This is due to the Great Moderation when inflation volatility markedly declined; decisive monetary policy and improved central bank communication anchored price expectations to an (implicit) target. Due to the limited coverage of the EPU index prior to 1985, the early part of the sample is also a period that the Markov-switching model in Appendix B designates as LU. Two-tailed t-tests during the sub-period until 1990 shows disagreement lower in HU periods than LU periods, while the result gets reversed in the sub-period after 1990.[8] In the later part of the period, there is evidence of state-dependent forecaster disagreement that is also counter-cyclical, as in Ilut and Schneider (2014); Lahiri and Sheng (2008).Yet, the overall trend of declining disagreement is dominant, which has produced this puzzling finding. The fact that disagreement varies with the state of economic uncertainty is explained by the sticky information model but not by the basic noisy information model.[9]

[8]These findings are not significant due to the fewer observations in each sub-period, 84 before 1990 and 118 afterward.

[9]The basic noisy information model can be altered to generate time-varying disagreement by introducing heterogeneity in the noise among agents or conditional time-variance of the noisy signal that is correlated with the magnitude of the shocks (Andrade and Le Bihan, 2013).

3.2 Basic Empirical Tests

3.2.1 Preliminary OLS Regressions

In order to further motivate the state-dependent model of expectation formation, I examine the relationship of the empirical degree of attention and the unobserved state of economic uncertainty as ($\lambda_t = f(\sigma_t^2)$) by regressing the empirical degree of attention λ_t on the proxies of uncertainty x_t. The following equation is estimated using OLS regressions with heteroskedasticity-robust Huber-White standard errors:

$$\lambda_t = \alpha + \beta x_t + \delta_t + \epsilon_t, \tag{3.1}$$

where the dependent variable λ_t is the empirical attention and $0 \leq \lambda_t \leq 1$. It is calculated as the proportion of forecasters who revise their forecasts from the previous period *t - 1* in all forecasters who make a prediction at time *t*. The independent variable x_t is a signal for the hidden state of economic uncertainty σ_t^2 at time *t*. The extended EPU index is used as the empirical proxy for uncertainty. In addition, the extended VXO index is also used as a robustness check. Regressions (5) - (8) include quarter dummies δ_t to capture any seasonality effect. Regressions (2), (4), (6), and (8) also include a constant α. Finally, ϵ_t is the error term in the regressions.

Table 2: Preliminary OLS Regressions

	Empirical Attention λ_t in GDP Deflator Inflation Forecasts							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extended EPU	0.00825***	0.000745***			0.00630***	0.000751***		
	(0.000223)	(0.000219)			(0.000339)	(0.000224)		
Extended VXO			0.0428***	0.000801			0.0328***	0.000733
			(0.00136)	(0.000583)			(0.00226)	(0.000613)
Constant		0.855***		0.915***		0.842***		0.906***
		(0.0264)		(0.0145)		(0.0293)		(0.0189)
Observations	203	203	203	203	203	203	203	203
Quarter Dummies	NO	NO	NO	NO	YES	YES	YES	YES
R-squared	0.901	0.103	0.895	0.004	0.929	0.112	0.920	0.013
Adj. R-squared	0.901	0.0981	0.895	-0.0009	0.928	0.0939	0.918	-0.0071

Robust standard errors in parentheses *** p< 0.01, ** p<0.05, * p<0.1.

The table presents results from OLS regressions of empirical attention λ_t on the proxies of economic uncertainty, the extended EPU and VXO indices, with heteroskedasticity-robust Huber-White standard errors. Empirical attention λ_t is calculated as the proportion of forecasters who update their forecasts of GDP deflator inflation out of all forecasters in the sample for each quarter.

These regressions should be considered as evidence of association, rather than causation, between empirical attention λ_t and economic uncertainty, as approximated by the two uncertainty indices. Whereas there is little reason to worry about reverse causality, since it is clearly economic uncertainty affecting the degree of attention and not vice versa, the regressions may still be afflicted by endogeneity due to omitted variable bias. There may be omitted variables that are correlated with both the uncertainty proxy and the error term of the regressions.Yet, Table 2 presents an initial step in empirically establishing the relationship between attention λ_t and uncertainty.

The results from the OLS regressions show that there is highly statistically significant (at the 1 percent level) positive relationship between empirical attention in the forecasts of GDP deflator inflation and the extended EPU index. The relationship is of the same level of significance using the extended VXO index as the independent variable. However, the coefficients are of lower magnitude when a constant is included in the regressions and even lose significance in the model with the extended VXO index as the uncertainty proxy. The regressions including quarter dummies also show slightly smaller coefficients but of the same levels of significance. Despite the limitation of these simple OLS regressions, the results from regressions (1) and (5) using the EPU index as the uncertainty proxy suggest that if uncertainty increases by one standard deviation, attention rises by 2.7-3.5 standard deviations, *ceteris paribus*. Since the sample size varies between 9 and 131 individual forecasters every quarter, with an average of 46 forecasters, this implies that a one-standard deviation increase in EPU uncertainty could make 2 to 39 professional forecasters attentive, 10-14 on average, which is an economically significant effect.

3.2.2 Bi-variate VAR

A simple recursive bi-variate vector autoregression (VAR) is estimated in order to address some of the concerns around the previous OLS regression analysis. The VAR approach allows for symmetric treatment of the included variables in a structural sense by estimating an equation for each variable explaining its evolution based on its own lags and the lags of the other variable in the system. Two VAR models are estimated with the following Cholesky identification ordering $[\text{EPU}, \lambda_t]'$ and $[\text{VXO}, \lambda_t]'$. The empirical proxy for uncertainty is ordered first, since it affects attention contemporaneously, whereas attention can be considered to impact uncertainty only with a lag. As a robustness check on this imposed assumption, I also estimated the reverse ordering leading to very similar results. The estimated VAR models include one lag of all variables, as preferred by the Hannan-Quinn information criterion (HQIC) and the Schwarz's Bayesian information criterion (SBIC). To confirm the stationarity of the series, the augmented Dickey-Fuller test is performed and it strongly rejects the null hypothesis of a unit root in the attention measure and the EPU and VXO indices at the 1 percent significance level. The Phillips-Perron unit-root test using Newey-West standard errors to account for serial correlation confirms the rejection of the unit root null for all three variables at the 1 percent significance level. Finally, the results are also confirmed with the modified Dickey-Fuller GLS test proposed by Elliott et al. (1996), a t-test for a unit root in which the series has been transformed by a generalized least-squares regression.

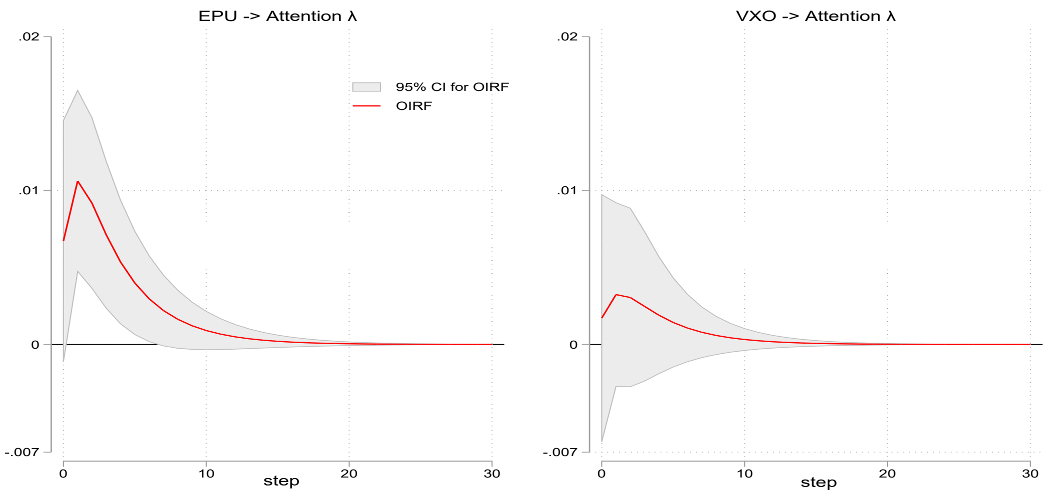

Figure 1: Cholesky Orthogonalized Impulse-Response Functions
The figure presents impulse-response functions orthogonalized using Cholesky decomposition, i.e. the impact of an uncertainty shock on empirical attention λ_t. Attention λ_t is calculated as the proportion of forecasters who update their forecasts of GDP deflator inflation out of all forecasters in the sample for each quarter. **Post-estimation results**: The estimated VAR models are stable, since all eigenvalues lie inside the unit circle. The null hypothesis that there is no residual first-order autocorrelation in the residuals of the VAR model including the EPU index cannot be rejected at the 10 percent level. The Jarque-Bera, skewness, and kurtosis statistics reject the null hypothesis that the disturbances in the VAR models are normally distributed, which may indicate some model misspecification. Wald tests of the null hypothesis that all endogenous variables at one lag are separately and jointly zero are strongly rejected in all equations. Finally, Granger causality Wald tests suggest that the null hypothesis that attention does not Granger-cause uncertainty can be rejected at the 10 percent for the VAR with the EPU index, while the null hypothesis that uncertainty does not Granger-cause attention can be rejected at the 1 percent level, which is in line with the hypothesized relationship. Neither of these Granger causality tests can reject the null for the VAR using the VXO index.

Figure 1 presents the impulse-response functions orthogonalized using Cholesky decompositions. The EPU uncertainty shock (one standard deviation increase in the EPU index) causes a significant increase in the empirical attention λ_t. Attention shoots up immediately in the first quarter after the uncertainty shock and then slowly dies down by the seventh quarter after the shock. In contrast, there is no significant effect on attention from a VXO uncertainty shock (one standard deviation increase in the VXO index). These results suggest that only an increase in the broader economic policy-related uncertainty has a positive, significant, and persistent effect on attention in GDP deflator inflation forecasts. An increase in financial sector-related uncertainty, as proxied by the VXO index, has no significant effect on attention in inflation forecasts, which may be due to the high-frequency volatility in financial data. This finding provides empirical confirmation of the choice of the EPU index, rather than the VXO index, as the baseline uncertainty proxy in the model. The impact of an

EPU uncertainty shock on attention is of similar magnitude as the results obtained from the OLS regressions in Section 3.2.1, which adds further validity to these findings. The results are in line with the hypothesized relationship, according to which the degree of attention is affected by economic uncertainty, suggesting that the expectation formation process of economic agents is state-dependent.

3.3 Empirical Test of Information Rigidity

To confirm the presence of information rigidity in the SPF data, I conduct a standard empirical test used in the literature of regressing the mean forecast error on its lag (Coibion and Gorodnichenko, 2012; Ryngaert, 2017). In Appendix C, I confirm these results with two additional tests, regressing the mean forecast error on the mean forecast revision (Coibion and Gorodnichenko, 2015) and regressing the mean forecast revision on its lag (Dovern et al., 2015; Nordhaus, 1987). I also test the central hypothesis of this study that information rigidity is characterized by uncertainty-based state-dependence by including interaction terms with uncertainty indicators in these regressions.

The main empirical approach used to estimate information rigidity expresses the predictability of the mean forecast error with respect to its lag, as developed by Coibion and Gorodnichenko (2012) and Ryngaert (2017):

$$MeanForecastError_{t+1|t} = \alpha^a + \beta^a MeanForecastError_{t|t-1} + \delta_t + \epsilon_t, \tag{3.2}$$

where α^a is a constant, $MeanForecastError_{t+1|t} = ActualValue_{t+1} - MeanForecast_{t+1|t}$, $MeanForecastError_{t|t-1} = ActualValue_t - MeanForecast_{t|t-1}, \delta_t$ are quarter dummies to account for any seasonality effect in the data, and ϵ_t is the error term in the regression. Since the error ϵ_t is orthogonal to information dated t and earlier, this equation can be estimated by OLS (Coibion and Gorodnichenko, 2012). This empirical test is performed on mean forecast data of GDP deflator inflation from the SPF for the entire sample Q4 1968 - Q2 2019. As a robustness check to control for the effect of major revisions in the realized data, mean forecast errors are calculated using the first, second, third, and final releases of the actual values, obtained from the Federal Reserve Bank of Philadelphia's real-time data set. The results are presented in Table 3.

As expected, all estimated β^a coefficients are positive and highly significant at the 1 percent level, indicating that there are significant information rigidities in the data. The magnitude of the coefficients generally increase with later releases of actual data. None of

Table 3: Information Rigidity Test Based on Forecast Error

	Forecast Error (First Release)			Forecast Error (Second Release)		
	(1)	(2)	(3)	(4)	(5)	(6)
Lag FE (First Release)	0.341***	0.338***	0.349***			
	(0.0966)	(0.0986)	(0.108)			
Lag FE (Second Release)				0.383***	0.383***	0.395***
				(0.0966)	(0.0971)	(0.106)
Constant		-0.0667	0.0466		-0.00824	0.0671
		(0.0962)	(0.168)		(0.0987)	(0.187)
Observations	199	199	199	197	197	197
Quarter Dummies	NO	NO	YES	NO	NO	YES
R-squared	0.115	0.113	0.142	0.147	0.147	0.179
Adjusted R-squared	0.110	0.109	0.124	0.143	0.143	0.162
	Forecast Error (Third Release)			Forecast Error (Final Release)		
	(7)	(8)	(9)	(10)	(11)	(12)
Lag FE (Third Release)	0.370***	0.370***	0.380***			
	(0.101)	(0.101)	(0.109)			
Lag FE (Final Release)				0.543***	0.543***	0.547***
				(0.0769)	(0.0769)	(0.0791)
Constant		-0.00253	0.0558		0.00299	0.180
		(0.0965)	(0.197)		(0.0757)	(0.159)
Observations	200	200	200	201	201	201
Quarter Dummies	NO	NO	YES	NO	NO	YES
R-squared	0.137	0.137	0.161	0.296	0.296	0.306
Adjusted R-squared	0.132	0.132	0.144	0.293	0.292	0.292

Robust standard errors in parentheses *** p< 0.01, ** p<0.05, * p<0.1.

The table presents results from OLS regressions of mean forecast errors on their lags, following Coibion and Gorodnichenko (2012) and Ryngaert (2017). The mean forecast errors are obtained by subtracting the individual quarterly forecasts from the first-, second-, third-, and final-release actual data and averaging across forecasters. All regressions include heteroskedasticity-robust Huber-White standard errors.

the constants in the regressions are significant, suggesting no bias in the forecasts. Overall, the specifications including constants and quarter dummies are preferred as a baseline due to their higher adjusted R-squared, although all specifications presented in this study are robust to excluding the constants and quarter dummies. The results using the final actual data release to calculate forecast errors are comparable to the findings of Ryngaert (2017) of $\beta^a = 0.53^{***}$ for one-quarter ahead forecasts. However, these results are lower than Coibion and Gorodnichenko (2012)'s $\beta^a = 0.88^{***}$ for four-quarters ahead forecast data, which could be explained by the fact that information rigidity increases over longer prediction horizons.

Table 4: Information Rigidity Test Based on Forecast Error: Interactions

	Forecast Error (First Release)			Forecast Error (Second Release)		
	(1)	(2)	(3)	(4)	(5)	(6)
Lag FE (First Release)	0.385***	0.436***	0.772***			
	(0.116)	(0.136)	(0.216)			
Lag FE (Second Release)				0.427***	0.469***	0.719***
				(0.111)	(0.132)	(0.210)
Lag FE×HU Indicator	-0.424**	-0.503**	-0.00467**	-0.447**	-0.465*	-0.00369*
	(0.176)	(0.236)	(0.00184)	(0.183)	(0.250)	(0.00193)
HU Dummy	-0.408**			-0.416**		
	(0.175)			(0.175)		
Prob(HU)		-0.412*			-0.399*	
		(0.230)			(0.228)	
EPU Index			-0.00494**			-0.00509**
			(0.00203)			(0.00213)
Constant	0.134	0.141	0.502*	0.142	0.147	0.539*
	(0.189)	(0.203)	(0.301)	(0.207)	(0.222)	(0.325)
Observations	199	199	199	197	197	197
Quarter Dummies	YES	YES	YES	YES	YES	YES
R-squared	0.170	0.167	0.177	0.206	0.199	0.205
Adjusted R-squared	0.144	0.141	0.151	0.181	0.173	0.180
	Forecast Error (Third Release)			Forecast Error (Final Release)		
	(7)	(8)	(9)	(10)	(11)	(12)
Lag FE (Third Release)	0.407***	0.447***	0.672***			
	(0.114)	(0.136)	(0.221)			
Lag FE (Final Release)				0.563***	0.597***	0.730***
				(0.0835)	(0.0954)	(0.210)
Lag FE×HU Indicator	-0.437**	-0.449*	-0.00339	-0.274*	-0.386*	-0.00249
	(0.193)	(0.267)	(0.00207)	(0.161)	(0.206)	(0.00196)
HU Dummy	-0.413**			-0.144		
	(0.170)			(0.140)		
Prob(HU)		-0.426*			-0.224	
		(0.222)			(0.188)	
EPU Index			-0.00552***			-0.00407*
			(0.00209)			(0.00207)
Constant	0.140	0.157	0.580*	0.144	0.162	0.515*
	(0.213)	(0.227)	(0.321)	(0.175)	(0.189)	(0.298)
Observations	200	200	200	201	201	201
Quarter Dummies	YES	YES	YES	YES	YES	YES
R-squared	0.188	0.181	0.188	0.307	0.310	0.316
Adjusted R-squared	0.162	0.155	0.163	0.285	0.289	0.295

Robust standard errors in parentheses *** p< 0.01, ** p<0.05, * p<0.1.

The table presents results from OLS regressions of mean forecast errors on their lags, following Coibion and Gorodnichenko (2012) and Ryngaert (2017). High uncertainty (HU) indicators include: HU Dummy = dummy for periods with probability of HU regime at least 50 percent (see Appendix B); Prob(HU) = probability of HU regime (see Appendix B); EPU Index = uncertainty proxy. The mean forecast errors are obtained by subtracting the individual quarterly forecasts from the releases of actual data and averaging across forecasters. All regressions include heteroskedasticity-robust Huber-White standard errors.

To test the central hypothesis of this study that information rigidity is characterized by uncertainty-based state-dependence, Table 4 adds to the baseline specifications from Table 3 interactions of the explanatory variable (lagged mean forecast error) with various measures of economic uncertainty, including a high-uncertainty (HU) dummy for quarters with probability of HU regime at least 50 percent and the continuous probability of HU regime itself, Prob(HU), both as estimated by a Markov-switching dynamic autoregression model in Appendix B, as well as the extended EPU index on which this Markov-switching dynamic autoregression is based.

The β^a coefficients using all releases of actual data are of even higher magnitude when the interaction terms are included, as compared to the baseline results from Table 3. The coefficients on the interaction terms are all negative and in the majority of specifications, at least significant at the 10 percent level. This indicates that whereas information rigidity overall may be higher than previously estimated, it significantly decreases with rising economic uncertainty. In half of the models in Table 4, the coefficients of the interaction terms are even larger in absolute value than the coefficients of the lagged forecast error, suggesting that information rigidity in HU periods falls between about 49 percent (Model (10), $-0.274/0.563$) and 118 percent (Model (1), $-0.424/0.385$). This finding that information rigidity is decreasing in economic uncertainty is in line with the results of Baker et al. (2019), who find significantly negative coefficients on interaction terms with dummies for periods with natural disasters.[10]

The results from the two additional tests of information rigidity regressing the mean forecast error on the mean forecast revision (Coibion and Gorodnichenko, 2015) and regressing the mean forecast revision on its lag (Dovern et al., 2015; Nordhaus, 1987) in Appendix C offer some additional support to the notions that information rigidity is present even in short-term forecasts and that it decreases rather substantially with rises in economic uncertainty. The effect of state-dependence on information rigidity is significant and merits more detailed investigation. The results give a preliminary indication that the hypothesized attention or *wake-up call* effect may be stronger than the alternative noisiness or *wait-and-see* effect of uncertainty on information rigidity. Thus, as uncertainty increases, forecasters pay significantly more attention and update their forecasts at a more frequent rate to reflect new developments in the data. Doing so outweighs the effect of increased noisiness in the new incoming information, which would otherwise cause forecasters to optimally put less weight

[10] Baker et al. (2019) include interactions in the empirical test regressing the mean forecast error on the mean forecast revision (Coibion and Gorodnichenko, 2015) and in the test regressing the mean forecast revision on the lagged mean forecast revision (Dovern et al., 2015; Nordhaus, 1987).

on it in updating their predictions. Which effect on information rigidity dominates is further examined in Section 5 and Section 7, which present results from structural estimation and regressions of simulated forecasts from the theoretical model that is described next.

4 Theoretical Model

Collecting and especially analyzing information in order to update one's expectation is costly in terms of time, resources, and expertise. In normal times when economic uncertainty is relatively low, it is optimal for agents *not* to update their expectations every period. In contrast, during high-uncertainty episodes, the cost of *not* updating one's information in terms of forecast accuracy can dramatically increase, as one's previous prediction quickly becomes obsolete. The notion that attention λ_t is a positive function of economic uncertainty is derived from the solution of an optimization problem that agents solve in order to minimize the cost of attention in terms of forecast accuracy (mean squared error, MSE).

Hence, in this model, economic agents form their expectations in two stages. First, before deciding whether to engage in a comprehensive forecast re-estimation, agents conduct a preliminary step of assessing the hidden state of economic uncertainty σ_t^2, which can be in a low-uncertainty (LU) or a high-uncertainty (HU) regime: $\sigma_t^2 \in \{\sigma_{LU}^2, \sigma_{HU}^2\}$. Using easily obtainable common public information, agents calculate the probabilities of each regime and decide in which of the two states of economic uncertainty they are more likely to find themselves. Based on this inference, on average agents update their forecasts with probability $\lambda_{t|t}$, the expected degree of attention. According to the solution of the optimization problem, attention is proportional to the probability to be in the HU regime, conditional on the observed data: $\lambda_{t|t} \propto \pi_{HU|t}$. The innovation of the model is that the degree of attention is state-dependent and evolves with the states of economic uncertainty.

In the second stage of the expectation formation process, only the fraction of agents λ_t who are attentive update their individual time-t forecasts $\hat{y}_{t+1|t}$ for the macroeconomic variable being forecasted y at period $t+1$. The remaining $(1-\lambda_t)$ of agents do not update their information at time t and instead simply carry forward their forecasts from the previous period $t-1$, $\hat{y}_{t|t-1}$, according to the state equation. The agents who update their predictions use the Kalman filter, as in a traditional noisy information model. Using this filter, they assign optimal weights to their idiosyncratic incoming noisy information and their previous estimate depending on the estimated precision of the new data and the perceived accuracy of the existing estimate. The expected Kalman gain, the weight agents put on new information, $k_{t|t} \not\propto \pi_{HU|t}$, since the higher the probability to be in the HU state, the noisier the new incoming information. Hence, in the proposed state-dependent sticky-noisy information model, both the attention effect and the noise effect are driven by time-varying economic uncertainty and they influence aggregate information rigidity in the economy in opposite directions. Which effect dominates is ambiguous *a priori* and depends on the parameters of the model. The decision problem of the individual agent is illustrated in Figure 2.

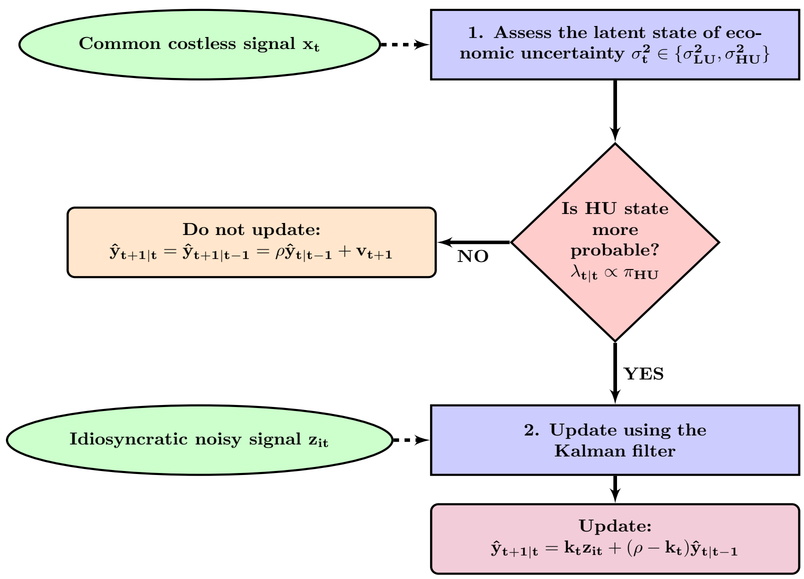

Figure 2: Agent's Problem in State-dependent Sticky-Noisy Information Model

Thus, agents update their information sets according to a hybrid sticky-noisy information model á la Andrade and Le Bihan (2013). However, here the probability of updating, or attention λ_t, is dependent on the hidden state of economic uncertainty, as perceived by the forecasters. It is important to note that I do not assert that economic agents *actually* run a Markov-switching model or a Kalman filter to update their forecasts and that the former is less computationally costly than the latter. Instead, these algorithms approximate the agents' actual decision-making process. So, the preliminary stage of making inference about the unobserved state of economic uncertainty is computationally less expensive (assumed free in the model) than engaging in a full-fledged forecast updating exercise every period, which carries a nonzero cost for acquiring and processing information ($\Psi_t(\lambda_t)$ in the model) and relatively lower benefit in terms of forecast accuracy when uncertainty is sufficiently low. In practice, agents can use publicly available, free, and easily accessible information to decide the extent to which it is worth their time and resources to collect and analyze new information and to re-estimate their forecasts in any given period, which carries the nonzero cost $\Psi_t(\lambda_t)$.

4.1 Micro-foundation

The economic agent chooses her degree of attention λ_t to minimize the following value function, in mean squared error (MSE) units, w.r.t. $\lambda_t \in [0,1]$:

$$
\begin{aligned}
V(\lambda_t) &= MSE + \text{Cost of attention} \\
&= MSE + \Psi_t(\lambda_t) \qquad (4.1)
\end{aligned}
$$

The MSE of the individual forecast is the weighted sum of the MSE if the agent is updating and the MSE if the agent is not updating, weighted by their respective probabilities:

$$
MSE = \lambda_t \times \sigma^2_{t|t} + (1-\lambda_t) \times \sigma^2_{t|t-1}. \qquad (4.2)
$$

Substituting for $\sigma^2_{t|t}$, the MSE of the nowcast at time t from the second stage of the model in equation (D.13), where k_t is the Kalman gain:

$$
\begin{aligned}
MSE &= \lambda_t[(1-\rho^{-1}k_t)\sigma^2_{t|t-1}] + \sigma^2_{t|t-1} - \lambda_t\sigma^2_{t|t-1} \\
&= (1-\rho^{-1}\lambda_t k_t)\sigma^2_{t|t-1}. \qquad (4.3)
\end{aligned}
$$

The cost of attention $\Psi_t(\lambda_t)$ is assumed to take the form $\Psi_t(\lambda_t) = (\lambda_t k_t)^2$. As in Branch et al. (2009), a cost function that is quadratic in λ_t allows for increasing marginal costs, with the marginal cost tending to zero when $\lambda_t \to 0$. Hence, it is always optimal to choose $\lambda_t > 0$. This cost function is also quadratic in the Kalman gain k_t, which is the weight agents put on new information. Thus, the more attention agents pay and the more weight they put on the new incoming information, the higher the cost of attention. Converting this problem into the constrained optimization format, this becomes:

$$
V(\lambda_t) = \max -\{(1-\rho^{-1}\lambda_t k_t)\sigma^2_{t|t-1} + (\lambda_t k_t)^2\} \text{ s.t. } \lambda_t \leq 1 \text{ and } -\lambda_t \leq 0. \qquad (4.4)
$$

Appendix D presents the solution of the constrained optimization problem using the Karush-Kuhn-Tucker complementarity slackness conditions. The result is that

$$
\lambda_t = \frac{\rho^{-1}\sigma^2_{t|t-1}}{2k_t} \qquad (4.5)
$$

maximizes the objective function, unless uncertainty is above the two thresholds

$$\sigma_t^2 \geq 2\rho^2 - \sigma_{t|t-1}^2$$
$$\sigma_t^2 > \frac{(\sigma_{t|t-1}^2 + \sigma_t^2)^2 - 4\rho^2\sigma_{t|t-1}^2 + 4\rho^4}{4\rho^2},$$

in which case complete attention ($\lambda_t = 1$) is optimal.

4.2 First Stage: Uncertainty States via Markov-switching Model

The first stage of the expectation formation process involves agents assessing the latent state of economic uncertainty $\sigma_t^2 \in \{\sigma_{LU}^2, \sigma_{HU}^2\}$. This state is hidden, so agents have to use an observed common costless signal x_t about the state of economic uncertainty in order to estimate the probabilities of whether they are in a low-uncertainty (LU) or a high-uncertainty (HU) regime at time t. This process is approximated by a Markov-switching model. Hamilton (1994, Chapter 22) provides a detailed overview and discussion of Markov-switching models. In estimating the Markov-switching model in Section 5.1, I follow the methodology of Perlin (2015).

Agents observe a common costless signal $x_t = \mu_{\sigma_t} + \phi x_{t-1} + \epsilon_t$ with a state-dependent mean μ_{σ_t} and error term $\epsilon_t \sim i.i.d.\ \mathcal{N}(0,\ \sigma_{\sigma_t}^2)$, about the hidden state of economic uncertainty σ_t^2 at time t. The hidden state σ_t^2 is assumed to follow an irreducible aperiodic two-state Markov chain, where by definition $P\{\sigma_t^2|\sigma_{t-1}^2, \sigma_{t-2}^2, ..., \sigma_1^2\} = P\{\sigma_t^2|\sigma_{t-1}^2\}$. The transition of the states is a stochastic process; however, the dynamics of the switching process are known and driven by a matrix of transitional probabilities.

The Markov-switching model is estimated with maximum likelihood. Yet, since the states of uncertainty are unknown, the notation of the likelihood function becomes $f(x_t|\sigma_t^2 = j; \theta)$ for state j and conditional on a set of parameters $\theta = (\mu_L,\ \mu_H,\ \sigma_L^2,\ \sigma_H^2,\ p_{LL},\ p_{HH})$. The full log likelihood function of the model is a weighted sum of the likelihood in each state j with the weights equal to the state's probabilities. These probabilities are not observed; however, we can make inference about the probabilities based on the available information. Appendix D.2 explains how these conditional probabilities are estimated using Hamilton's iterative algorithm to obtain:

$$\pi_{j|t} = Pr(\sigma_t^2 = j|I_t) = \frac{f(x_t|\sigma_t^2 = j; I_{t-1})Pr(\sigma_t^2 = j|I_{t-1})}{\sum_{j\in\{L,H\}} f(x_t|\sigma_t^2 = j; I_{t-1})Pr(\sigma_t^2 = j|I_{t-1})}. \tag{4.6}$$

Using the set of conditional probabilities $\pi_{j|t}$, the log likelihood of the model can be calculated by maximum likelihood as a function of the parameters of the model that maximize:

$$\ln L = \sum_{t=1}^{T} \ln \sum_{j \in \{L,H\}} \Big(f(x_t | \sigma_t^2 = j; \theta) \pi_{j|t} \Big). \tag{4.7}$$

4.3 Second Stage: Forecast Updating via Kalman Filter

In this model, agents update their information sets according to a hybrid sticky-noisy information model. Only a proportion of the population λ_t updates its time-t forecasts $\hat{y}_{t+1|t}$ for variable y at time period $t+1$. The remaining population $(1-\lambda_t)$ does not update its information at time t and simply carries forward its forecast from the previous period $t-1$, $\hat{y}_{t|t-1}$, scaled by the persistence of the state process ρ and polluted with error, according to the state equation, which takes an AR(1) form:

$$y_t = \rho y_{t-1} + v_t, \text{ where } v_t \sim i.i.d.\ \mathcal{N}(0,\, \sigma_v^2) \text{ and } 0 < \rho < 1. \tag{4.8}$$

The agents who update their information at period t receive an idiosyncratic noisy information signal. They solve a signal-extraction problem using the Kalman filter. The observation equation is:

$$z_{it} = y_t + \eta_{it}, \tag{4.9}$$

where $\eta_{it} \sim i.i.d.\ \mathcal{N}(0,\, \sigma_t^2)$. The variance of the error η_{it}, σ_t^2, is stochastic and thus, the model is characterized by stochastic volatility. Furthermore, σ_t^2 is assumed to follow an irreducible aperiodic two-state Markov chain, where $\sigma_t^2 \in \{\sigma_{LU}^2, \sigma_{HU}^2\}$, the two states of economic uncertainty from Section 4.2.

Equations (4.8) and (4.9) together describe the information structure in state space form (ssf). The recursion is initialized with $\hat{y}_{1|0} = \mathbb{E}(y_1|z_{i0}) = 0$ with associated MSE $\sigma_{1|0}^2 = Var(y_1|z_{i0}) = \frac{\sigma_v^2}{1-\rho^2}$, which are just the unconditional mean and variance of y_1 using information from time $t = 0$ (Hamilton, 1994, Chapter 13).

For each agent i, the Kalman filter assigns different weights to the incoming noisy information and the previous estimate depending on their respective precisions. The Kalman gain k_t is a measure of how much the forecaster can trust her information signal. The more credible the signal, the more weight the forecaster will optimally put on it in updating her expectation. The Kalman gain k_t is defined as:

$$k_t = \frac{\rho \sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_t}. \tag{4.10}$$

Appendix D.3 provides the derivation of the attentive agents' forecasts for the state variable and its mean squared error (MSE) at time t for the period $t+1$, respectively, as:

$$\hat{y}_{t+1|t} = k_t z_{it} + (\rho - k_t)\hat{y}_{t|t-1} \tag{4.11}$$
$$\sigma^2_{t+1|t} = \rho(\rho - k_t)\sigma^2_{t|t-1} + \sigma^2_v. \tag{4.12}$$

By the law of large numbers, assuming the population of forecasters is large, a state-dependent fraction λ_t of the population is attentive each period and a proportion $(1-\lambda_t)$ is inattentive. Thus, the mean forecast of the entire population is the sum of the mean forecasts within each group of forecasters j, who all updated their information sets j periods ago, weighted by the respective proportion of each group in the total population.

Appendix D.3 shows the derivation of the mean forecast at time t for time $t+1$ of all agents i in group $j = 0$ w.l.o.g.:

$$\begin{aligned}\mathbb{E}_j[\bar{y}_{t+1|t-j}|j=0] = \bar{y}^{all}_{t+1|t} &= \lambda_t \bar{y}^{attentive}_{t+1|t} + (1-\lambda_t)\bar{y}^{inattentive}_{t+1|t} \\ &= \lambda_t k_t y_t + (\rho - \lambda_t k_t)\bar{y}_{t|t-1}. \end{aligned} \tag{4.13}$$

4.4 State-dependent Attention λ_t and Kalman Gain k_t

The innovation of the proposed state-dependent sticky-noisy information model is that the quantity of information represented by attention λ_t and the quality of information that agents glean reflected by the Kalman gain k_t are driven by time-varying economic uncertainty. In order to investigate the effect of time-varying uncertainty on λ_t and k_t, I define these variables in terms of uncertainty σ^2_t and provide an auxiliary assumption in this section.

Using the auxiliary assumption from Bloom (2009) that $\sigma_{HU} = 2 \times \sigma_{LU}$ [11], Appendix D.4 expresses attention λ_t and the Kalman gain k_t in terms of the filtered probabilities of HU

[11] Therefore, $\sigma^2_{HU} = 4 \times \sigma^2_{LU}$. As in Bloom (2009), I confirm the results for $\sigma_{HU} = 2 \times \sigma_{LU}$ are also valid for $\sigma_{HU} = 1.5 \times \sigma_{LU}$ and $\sigma_{HU} = 3 \times \sigma_{LU}$.

state:

$$\lambda_{t|t} = \frac{\sigma^2_{t|t-1} + \sigma^2_{LU}(1 + 3\pi_{HU|t})}{2\rho^2} \tag{4.14}$$

$$k_{t|t} = \frac{\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_{LU}(1 + 3\pi_{HU|t})} \tag{4.15}$$

Partial derivative analysis in Appendix D.4 shows that both attention $\lambda_{t|t}$ and Kalman gain $k_{t|t}$ are increasing functions of the previous MSE. Agents are expected to pay more attention and to have more faith in the new information when their previous forecasts have been more inaccurate. Moreover, attention $\lambda_{t|t}$ is an increasing function of LU-regime uncertainty and by construction, also HU-regime uncertainty, confirming that agents are expected to become more attentive during periods of greater volatility. Yet, the Kalman gain $k_{t|t}$ is a decreasing function of uncertainty, since agents optimally put less weight on new incoming information when it is more noisy. Similarly, whereas attention is an increasing function of the conditional probability of a HU state, so that so that $\lambda_{t|t} \propto \pi_{HU|t}$, the Kalman gain is a decreasing function of it: $k_{t|t} \not\propto \pi_{HU|t}$. In other words, the greater the probability of a perceived HU state, the greater the probability of updating a forecast but the lesser the faith in the new information. Finally, attention $\lambda_{t|t}$ is a decreasing function of the persistence of the state process ρ, while the Kalman gain $k_{t|t}$ is an increasing function of it, suggesting that more persistent processes cause agents to optimally become less attentive but they optimally trust new information more. Overall, all these relationships are of the expected signs.

5 Structural Estimation

In the proposed state-dependent sticky-noisy information model, both the attention and the noise effect are driven by time-varying economic uncertainty and they influence aggregate information rigidity in the economy in opposite directions. Which effect dominates is ambiguous *a priori* and depends on the parameters of the model. In this section, I conduct structural estimations of the proposed model, as well as the alternative theories by simulating the theoretical models and estimating the structural parameters of interest using Simulated Method of Moments (SMM). Since the proposed model nests within itself the hybrid, the noisy, and the sticky information models, in Section 6, I compare the ability of these alternative theories to match the data by shutting down the state-dependence, the state-dependence and inattention, and the state-dependence and noisiness channels, respectively. I further examine the relative importance of each channel in Section 7.

5.1 SMM Estimation

I conduct the structural estimation of the proposed theoretical model using SMM, which estimates the structural parameters by minimizing the difference between empirical moments and their theoretical counterparts, as in Duffie and Singleton (1993), Gourieroux and Monfort (1996), Ruge-Murcia (2012) and Giacomini et al. (2020). The estimated structural parameters are thus the most likely to have generated the moments in the survey data.

The structural parameters I estimate are $\theta = (\,\rho \;\sigma_v^2 \;\sigma_{LU}^2\,)'$: the persistence of the state process, the variance of the error of the state process, and the baseline uncertainty. I match three moments from the Survey of Professional Forecasters (SPF) data and the simulated theoretical model, including the consensus (mean) forecast, accuracy expressed as the root mean squared error (RMSE) of forecasts from realized inflation, and forecaster disagreement (cross-sectional standard deviation of individual forecasts) over the entire length of the sample, Q4 1968 - Q2 2019. The objective function that is minimized is the difference between the empirical moments and the average moments from the simulation performed with a draw of θ:

$$J(\theta) = [\hat{m} - m(\theta)]' W [\hat{m} - m(\theta)], \tag{5.1}$$

where $\hat{m}$ are the empirical consensus forecast, accuracy, and disagreement, while $m(\theta)$ are their simulated theoretical counterparts. I set the weighting matrix W to equal the identity matrix, so that all three moments are weighted equally.

To obtain the forecasts needed to calculate the theoretical moments, I simulate the proposed model as follows. In the state-dependent sticky-noisy information model, agents update their expectations infrequently and when they do so, they face a signal-extraction problem. Crucially, the rate of updating, attention λ_t, and the weight agents put on new incoming information, Kalman gain k_t are not constant but driven by the unobserved states of economic uncertainty $\sigma_t^2 \in \{\sigma_{LU}^2, \sigma_{HU}^2\}$. Summing equation (4.13) over all groups of forecasters j who update their predictions in the same period yields the mean forecast in the state-dependent sticky-noisy information model, presented in Appendix D.5.

I model the dynamics of individual one-quarter ahead forecasts for $N = 70$ agents in order to approximate the cross-section of the core sample of 70 professional forecasters who participate in the SPF at least 40 periods for the length of the sample period $T = 203$, obtaining a 203×70 matrix of individual simulated forecasts of GDP deflator inflation. In the first period, the $(1 - \lambda_t)$ fraction of agents who do not update their forecasts predict GDP deflator inflation equal to 2.149 percent, the mean of realized GDP deflator inflation in the previous 20 years.[12] In every period after that, a time-varying fraction λ_t of the forecasters chosen at random update their expectations and thus, these attentive agents have forecasts as in the noisy information model. The fraction of updaters λ_t for each quarter is calculated per equation (4.14). The remaining fraction $(1 - \lambda_t)$ of forecasters are inattentive and carry forward their forecast from the last period they updated their predictions, scaled by the persistence of the state process ρ and polluted with error v_{t+1}.

I set the starting values of θ_0 as follows: the AR(1) coefficient $\rho = 0.9$ and the process innovation variance $\sigma_v^2 = 0.5$, as in Ryngaert (2017), and the signal noise variance, or baseline time-varying economic uncertainty is set to $\sigma_{LU}^2 = 0.4432^2 = 0.1964$, as estimated by Bloom (2009).[13] The conditional probability of being in a high-uncertainty (HU) state $\pi_{HU|t}$ is estimated from the Markov-switching dynamic autoregression presented in Appendix B and Section 4.2, using the methodology of Perlin (2015). I perform $M = 1000$ rounds for each simulation, discarding the first 100, and report the average individual forecasts, which are then used to calculate the mean forecast, RMSE of the forecasts from realized inflation, and forecaster disagreement (cross-sectional standard deviation of individual forecasts) over the entire length of the sample.

5.2 Estimation Results

The estimated parameters θ from the SMM structural estimation of the proposed state-dependent sticky-noisy information model are presented in Table 5, Column (4). The structural estimation results of the alternative models with information frictions are discussed in more detail in Section 6.

The structural estimation sets the persistence of the state process $\rho = 0.999$. This is slightly higher than the AR(1) coefficient from the survey data of 0.896, even though this value is not used in the estimation. The variance of the error of the state process $\sigma_v^2 = 0.006$ is lower than the value calibrated by Ryngaert (2017) of 0.5. Finally, the estimated baseline uncertainty $\sigma_{LU}^2 = 0.172$ is not much different from the estimation in Bloom (2009) of 0.1964. Since both attention λ_t and the Kalman gain k_t are time-varying and driven by the state of economic uncertainty, as defined in equations (4.14) and (4.15) respectively, I calibrate the proposed theoretical model with the values of the estimated structural parameters and

[12] Setting the expectation of the inattentive agents in the first period to the average realized GDP deflator inflation of the previous 2, 5, or 10 years instead does not substantially change the results.

[13] The specific starting values do not impact the estimated structural parameters θ very much. The presented results use these starting values as found in the literature.

present their dynamics in Figure 3.

Table 5: Estimated Parameters θ using SMM

θ	Sticky Info (1)	Noisy Info (2)	Hybrid Info (3)	State-dependent Info (4)
λ	0.128		0.128*	$\lambda_t = f(\sigma^2_{LU})$
σ^2_t (σ^2_{LU} in (4))		0.517	0.191	0.172
σ^2_v		0.523	0.145	0.006
ρ		0.905	0.999	0.999

*Calibrated as in (1).
Results are based on SMM estimation, matching mean forecast error, RMSE, and forecaster disagreement in all models with information frictions except for (1). For the estimation of the sticky information model in (1), only two moments are used, consensus forecast error and accuracy.

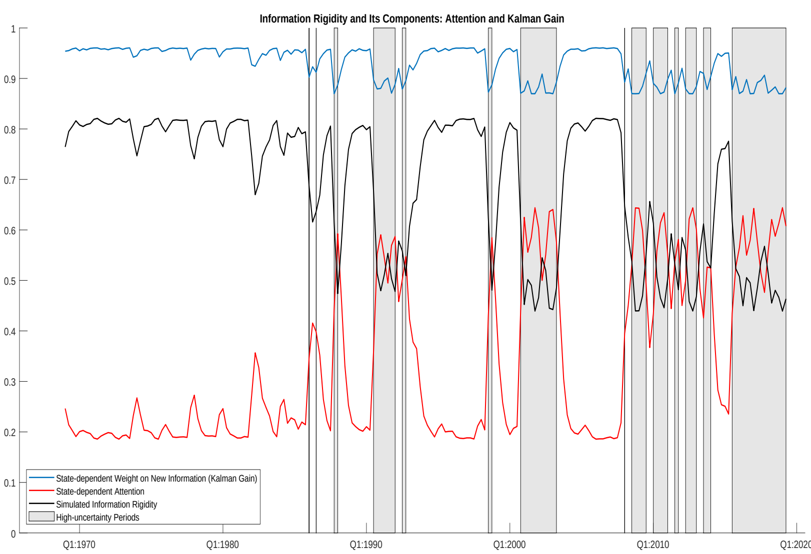

Figure 3: Information Rigidity and Components: Attention λ_t and Kalman Gain k_t
The figure presents the time series of the simulated attention λ_t and Kalman gain k_t at the estimated structural parameters θ using SMM and the aggregate information rigidity, calculated as $(\rho - \lambda_t k_t)$ in equation (7.1).

As expected, during high-uncertainty periods, agents optimally increase their quantity of information (attention increases) but since the quality of their information is lower, they optimally put less weight on it in forming their predictions (the Kalman gain decreases). Over the sample, attention varies between 0.19 and 0.64, which is within the range in the literature: 0.18 in Mankiw and Reis (2002), 0.25 in Coibion and Gorodnichenko (2012), 0.46 in Coibion and Gorodnichenko (2015) for models of rational inattention. On the other hand, the Kalman gain takes values between 0.87 and 0.96, which is relatively high compared to other studies: 0.5 in Coibion and Gorodnichenko (2015) and Dovern et al. (2015), 0.66 in Ryngaert (2017), 0.7 in Afrouzi (2017). These differences are not surprising, since in the proposed model there are interactions between the attention and noisiness channels, which moderate the magnitudes of both. This is evident in Figure 3, where also presented is aggregate information rigidity, calculated as $(\rho - \lambda_t k_t)$ in equation (7.1). It varies in the range 0.44-0.82, which is in line with the estimated aggregate information rigidity values.

This result implies that forecasters update their information sets every two to five quarters. Moreover, the dynamics of the simulated aggregate information rigidity over the sample uncover that the attention effect clearly dominates the noisiness effect: in high- (low)-uncertainty periods, the lower (higher) information rigidity attributed to higher (lower) attention clearly outweighs the higher (lower) information rigidity attributed to noisiness by the lower (higher) Kalman gain. As a result, aggregate information rigidity, which is influenced by these two effects in opposite directions, is lower (higher) in HU (LU) periods. This important finding is further scrutinized in Section 7 by conducting empirical tests on the proposed state-dependent model and the alternative models with information frictions. Before that, in Section 6, I analyze the results from the structural estimations of the alternative theories.

6 Alternative Models with Information Frictions

The proposed state-dependent sticky-noisy information model nests within itself the three prominent models in the field of information rigidity: 1) sticky information, 2) noisy information, and 3) hybrid sticky-noisy information models. Each of these existing models can be considered as a version of the proposed state-dependent sticky-noisy information model with the 1) state-dependence and noisiness, 2) state-dependence and inattention, and 3) state-dependence channels, respectively, shut down. Both the sticky and the noisy information models feature a single information friction and simply conflate information rigidity with inattention and noisiness, respectively. As a result, they are unable to match all the features of the data at once. Hybrid sticky-noisy information models have a notable advantage of including both attention and noisiness channels. Yet, existing hybrid models simply add these two frictions as if they are independent effects. The proposed state-dependent model improves on these models, since it allows for interaction between the attention and noisiness channels, which are both driven by time-varying economic uncertainty and that moderates the impact of both on aggregate information rigidity.

In this section, I describe in more detail the structural estimation of the existing alternative models with information frictions and compare their predictions to the observed SPF survey forecasts and the realized values of the predicted variable, GDP deflator inflation. The estimated parameters θ for all discussed models with information frictions are presented in Table 5.

6.1 Sticky Information Model

First, assuming no noisiness and no state-dependence but inattention yields the sticky information model. According to it, agents update their forecasts with a constant Poisson probability λ: each agent is equally likely to update every period. By the law of large numbers, assuming the population of forecasters is large, a random constant fraction λ of agents are attentive each period, while the remaining $(1 - \lambda)$ simply carry forward their predictions from their last update (Mankiw and Reis, 2002). The attentive agents glean the true value of the forecasted variable, unpolluted by noise. The mean forecast of the sticky information model follows equation (D.16) but with a constant λ, as shown in Appendix D.5.

I model the dynamics of *mean* expectations in the sticky information model during the length of the sample period in the SPF survey data $T = 203$, yielding a 203×1 vector of mean forecasts. As in the proposed state-dependent model, in the first period, the $(1 - \lambda)$ fraction of inattentive agents predict GDP deflator inflation equal to 2.149 percent, the mean of realized GDP deflator inflation in the previous 20 years. In every period after that, the contemporaneous mean forecast is the weighted average of the realized inflation and the previous average inflation forecast, weighted by λ and $(1 - \lambda)$, respectively.

In the SMM structural estimation, I estimate structural parameter $\theta = \lambda$, minimizing the difference between empirical and theoretical moments, including the consensus forecast and accuracy expressed as the root mean squared error (RMSE) of forecasts from realized inflation over the entire length of the sample, Q4 1968 - Q2 2019. I set the starting values of $\theta_0 = \lambda = 0.25$, as in Coibion and Gorodnichenko (2012). The value of the structural parameter that is most likely to have produced the observed moments is $\lambda = 0.128$. This value is below the range estimated in the literature and also lower than the empirical estimate of attention as the fraction of forecasters revising their predictions every period, $\lambda = 0.931$, found in Table 1. The estimated degree of attention implies that agents only revise their predictions every seven to eight quarters, which is clearly at odds with the survey data for professional forecasters. This inconsistency suggests that the attention effect alone is unable to explain the observed data.

6.2 Noisy Information Model

Second, shutting down both the direct inattention and the indirect state-dependence channels yields the noisy information model. According to it, all agents are attentive every period but they face a signal extraction problem (Woodford, 2002; Sims, 2003). The noisy information setup is described in detail in Section 4.3. The mean forecast is derived in Appendix D.5.

I simulate the dynamics of *individual* expectations in the noisy information model using a Kalman filter according to equations (4.9) and (4.8) for $N = 70$ agents in order to approximate the cross-section of the core sample of 70 professional forecasters who participate in the SPF at least 40 periods for the length of the sample period $T = 203$. This results in a 203×70 matrix of individual simulated forecasts of GDP deflator inflation.

In the SMM structural estimation, I estimate structural parameters $\theta = (\, \rho \ \sigma_v^2 \ \sigma_t^2 \,)'$: the persistence of the state process, the variance of the error of the state process, and the signal noise variance. I use the same moments as in the estimation of the proposed state-dependent model in Section 5: the consensus forecast, accuracy (RMSE), and forecaster disagreement over the entire length of the sample, Q4 1968 - Q2 2019. I set the starting values of θ_0 as follows: the AR(1) coefficient $\rho = 0.9$ and the process innovation variance $\sigma_v^2 = 0.5$, as in Ryngaert (2017), and the signal noise variance is set to $\sigma_t^2 = 0.1964$, as the value of baseline uncertainty estimated by Bloom (2009). I perform $M = 1000$ rounds for each simulation, discarding the first 100, and report the average individual forecasts, which are then used to calculate the three moments over the entire length of the sample.

The estimated parameters in Table 5 of $\theta = (\, 0.517 \ 0.523 \ 0.905 \,)'$ imply a Kalman gain $k = 0.54$, or agents updating their forecast about every two quarters. This is much closer to the results in existing studies and the observed information rigidity than the results from the sticky information model. However, since the noisy information model cannot account for any inattention observed in the data, it is also not satisfactory as a modeling device for the expectation formation process of economic agents.

6.3 Hybrid Sticky-Noisy Information Model

Third, assuming no state-dependence yields the hybrid sticky-noisy information model, which includes both inattention and noisiness but considers them independent information frictions. The hybrid model simply adds the sticky information model 'on top of' the noisy information model without considering the interactions between these two information frictions. Agents update their expectations infrequently and when they do so, they face a signal-extraction problem (Andrade and Le Bihan, 2013). Substituting equation (D.31) into equation (D.27) yields the mean forecast in the hybrid information model, presented in Appendix D.5.

I model the dynamics of *individual* expectations in the hybrid sticky-noisy information model for $N = 70$ agents, the core SPF sample, during the length of the sample period $T = 203$, obtaining a 203×70 matrix of individual simulated forecasts of GDP deflator inflation. As in the sticky and state-dependent information models, in the first period, inattentive agents predict GDP deflator inflation equal to 2.149 percent, the mean of realized GDP deflator inflation in the previous 20 years. In every period after that, constant fraction λ of the forecasters chosen at random update their expectations and have forecasts as in the noisy information model. Fraction $(1 - \lambda)$ are inattentive and carry forward their last predictions, scaled by the persistence of the state process ρ and polluted with error.

In the SMM structural estimation, I estimate the same structural parameters as in the estimation of the noisy information model $\theta = \left(\begin{smallmatrix} \rho & \sigma_v^2 & \sigma_t^2 \end{smallmatrix} \right)'$. As in Giacomini et al. (2020), I do not estimate but simply calibrate $\lambda = 0.128$, as per the sticky information model estimation. This is an appropriate calibration, since in the hybrid information model, there is no interaction between the attention and noisiness effects. I use the same moments as in the estimation of the proposed state-dependent model and the noisy information model: the consensus forecast, accuracy (RMSE), and forecaster disagreement over the entire length of the sample, Q4 1968 - Q2 2019. I set the starting values of θ_0 as in the noisy information model. I perform $M = 1000$ rounds for each simulation, discarding the first 100, and report the average individual forecasts, which are then used to calculate the three moments over the entire length of the sample.

The estimated parameters in Table 5 of $\theta = \left(\begin{smallmatrix} 0.191 & 0.145 & 0.999 \end{smallmatrix} \right)'$ imply a Kalman gain $k = 0.57$, which is broadly in line with the literature. However, since the hybrid model adds the noisiness and attention information frictions together, aggregate information rigidity can be estimated by substituting the time-varying λ_t and k_t with constant ones in equation (7.1). Aggregate information rigidity thus equals $(\rho - \lambda k) = 0.926$, implying that agents update their predictions every 13 quarters, which is much higher than both the range found in the literature and observed in survey data. This illustrates the principal shortcoming of the hybrid sticky-noisy information model: whereas it correctly takes into account both inattention and noisiness, since it simply adds the information frictions together with no regard for their interactions, it ends up predicting information rigidity that is implausibly high.

6.4 State-dependent Sticky-Noisy Information Model

Finally, the proposed state-dependent sticky-noisy information model nests within itself the three aforementioned models with information frictions, thereby including all three posited channels affecting information rigidity: the direct attention and noisiness, and the indirect state-dependence channels. Specifically, time-varying uncertainty drives both attention and noisiness, thereby integrating them and accounting for the interactions between them. This greatly improves upon existing hybrid sticky-noisy information models (Andrade and Le Bihan, 2013; Giacomini et al., 2020), which simply add the two frictions as independent effects. The proposed model is thus superior to the hybrid information model and theoretically more appealing than the noisy information model due to its inclusion of rational inattention observed in survey data. The structural estimation of the proposed model is discussed in detail in Section 5.1.

Figure 4 presents time series of the inflation forecasts from the four theoretical models with information frictions, simulated at the values of the structural parameters estimated using SMM, along with mean SPF survey data, and realized GDP deflator inflation. As the figures show, the proposed state-dependent sticky-noisy information model performs relatively well compared to the sticky, noisy, and hybrid sticky-noisy information models. Inattention has a smoothing effect on the estimated forecasts, while noisiness closely tracks the oscillations of the survey forecasts. The sticky and especially the hybrid information models produce forecasts that deviate the most from the SPF expectations. The simulated predictions from the proposed state-dependent model are closest to those from the noisy information model. This visually demonstrates the finding that the proposed state-dependent model is superior to the hybrid and sticky information models, while preserving some of the desirable features of the noisy information model. Yet, the proposed model is more appealing than the noisy information model due to its ability to account for observed inattention.

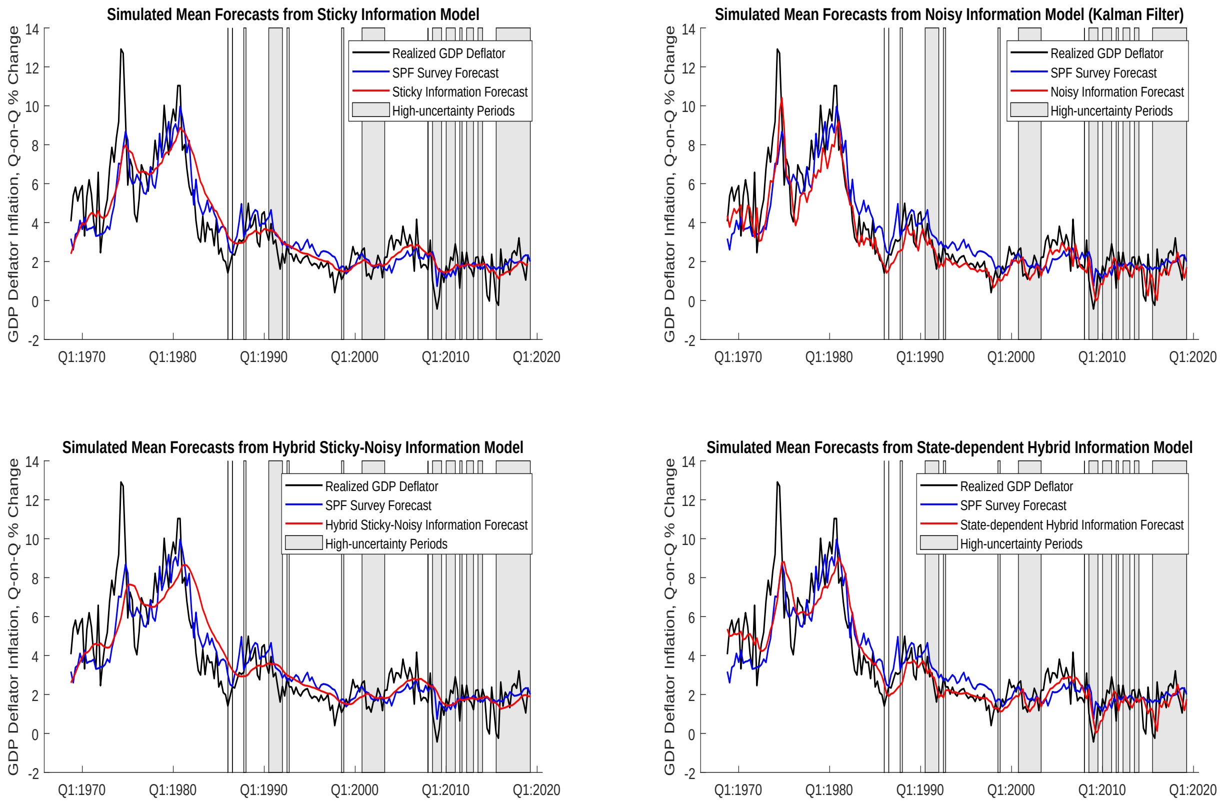

Figure 4: Simulated Forecasts: Alternative Models with Information Frictions
The figure presents time series of the mean SPF survey data, realized GDP deflator inflation, and simulated mean forecast data from the sticky, noisy, hybrid sticky-noisy, and state-dependent sticky-noisy information models at the values of the structural parameters estimated using SMM. The shaded regions indicate time periods to which the Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the high-uncertainty (HU) regime.

7 Results: Contributions to Information Rigidity

To further investigate the important finding from Section 5.2 that the attention dominates the noisiness effect when the two are allowed to interact in the proposed state-dependent model, I use empirical tests to estimate the relative contribution of each of the three channels affecting information rigidity, the direct attention and noisiness and the indirect state-dependence effects. Specifically, again I conduct the empirical test estimating information rigidity by regressing the mean forecast error on its lag from Section 3.3, as in Coibion and Gorodnichenko (2012) and Ryngaert (2017). However, this time I compare the regression coefficient using SPF data to the ones obtained using simulated forecasts at the values of the estimated structural parameters according to the sticky, noisy, hybrid sticky-noisy, and state-dependent sticky-noisy information models.

First, the theoretical foundation of this empirical test is obtained by subtracting both sides of equation (4.13) from the realized inflation y_{t+1}, which yields the following specification:

$$\begin{aligned} y_{t+1} - \bar{y}_{t+1|t} &= \rho y_t + v_{t+1} - [\lambda_t k_t y_t + (\rho - \lambda_t k_t)\bar{y}_{t|t-1}] \\ &= (\rho - \lambda_t k_t)(y_t - \bar{y}_{t|t-1}) + v_{t+1}, \end{aligned} \tag{7.1}$$

where the regression coefficient is $(\rho - \lambda_t k_t)$ and the error is v_{t+1}. In other words, the coefficient of information rigidity depends on the parameters of the model, attention λ_t (attention channel) and Kalman gain k_t (noisiness channel), both of which in turn are driven by economic uncertainty as per equations (4.14) and (4.15) (state-dependence channel), as well as the persistence of the state process ρ. To be closer to the theory, the subsequent regressions do not include constant terms or quarter dummies, as in Section 3.3. However, the inclusion of constant terms and quarter dummies does not substantially change the presented results.

Table 6: Information Rigidity Based on Forecast Error: SPF and Simulated Forecasts

	SPF FE (1)	Sticky FE (2)	Noisy FE (3)	Hybrid FE (4)	State-dependent FE (5)
Lag SPF FE	0.543***				
	(0.0769)				
Lag Sticky FE		0.681***			
		(0.0707)			
Lag Noisy FE			0.323***		
			(0.0906)		
Lag Hybrid FE				0.712***	
				(0.0696)	
Lag State-dependent FE					0.496***
					(0.0877)
Observations	201	201	201	201	201
R-squared	0.296	0.467	0.104	0.510	0.246
Adjusted R-squared	0.293	0.465	0.0995	0.507	0.242

Robust standard errors in parentheses *** p< 0.01, ** p<0.05, * p<0.1.

The table presents the results from OLS regressions of mean forecast errors on their lags (Coibion and Gorodnichenko, 2012; Ryngaert, 2017) using survey data and simulated data from the sticky, noisy, hybrid sticky-noisy, and state-dependent sticky-noisy information models at the values of the structural parameters estimated using SMM. The theoretical models with information frictions are simulated according to Section 6. All regressions include heteroskedasticity-robust Huber-White standard errors.

According to the results presented in Table 6, the simulated forecasts from the proposed state-dependent sticky-noisy information model yield a regression coefficient that is closest to the one from the survey data, explaining 91 percent of the information rigidity that is present in the inflation expectations of professional forecasters. As previously noted, both the sticky and the hybrid information models overestimate information rigidity, while the noisy information model underestimates it. Thus, all three channels, the direct inattention and noisiness and the indirect state-dependence effects need to be included in theoretical models, so that they match the information rigidity observed in survey inflation expectations. As noted previously, this has been a major shortcoming of existing hybrid sticky-noisy information models (Andrade and Le Bihan, 2013; Giacomini et al., 2020).

Following Ryngaert (2017), I estimate the relative importance of the three channels on information rigidity, since the hybrid model is the state-dependent model with state-dependence shut down, the sticky information model is the hybrid model with noisiness shut down, and the noisy information model is the hybrid model with inattention shut down. To calculate the relative importance of the inattention channel, I divide the coefficient estimate from the sticky information model by the coefficient estimate from the hybrid model. To obtain the relative contribution of the noisiness effect, I likewise divide the coefficient estimate from the noisy information model by the coefficient estimate from the hybrid model. This back-of-the-envelope calculation suggests that the inattention and noisiness frictions account for 95.7 percent and 45.4 percent of the information rigidity estimated in the inflation expectations of professional forecasters.[14]. Adding the contributions of these two direct channels sums up to 141.1 percent, which confirms that there are sizable additional offsetting interactions between the attention and noisiness channels and that in fact, they are not independent, as assumed by the hybrid sticky-noisy information model.

To my knowledge, this is the first study to estimate the relative contributions to information rigidity of inattention and noisiness, so this is another important finding. The fact that more than half of the estimated information rigidity is due to inattention is good news for monetary policymakers. It leaves ample room for monetary policy, which can affect the quantity of information available to economic agents by employing frequent, direct, and simple forward guidance.

The offsetting effect of the state-dependence channel accounts for -43.6 percent of the overall information rigidity, calculated as the residual of the coefficient estimate from the hybrid model divided by the coefficient estimate from the state-dependent model.[15] This value almost exactly offsets the overestimation of information rigidity by the hybrid model. Hence, it is important to consider state-dependence in models with information frictions

[14] $(0.681/0.712) \times 100 = 95.7\%; (0.323/0.712) \times 100 = 45.4\%$.
[15] $[1 - (0.712/0.496)] \times 100 = -43.6\%$.

because it accounts for an indirect channel through which the direct attention and noisiness effects interact to affect aggregate information rigidity.

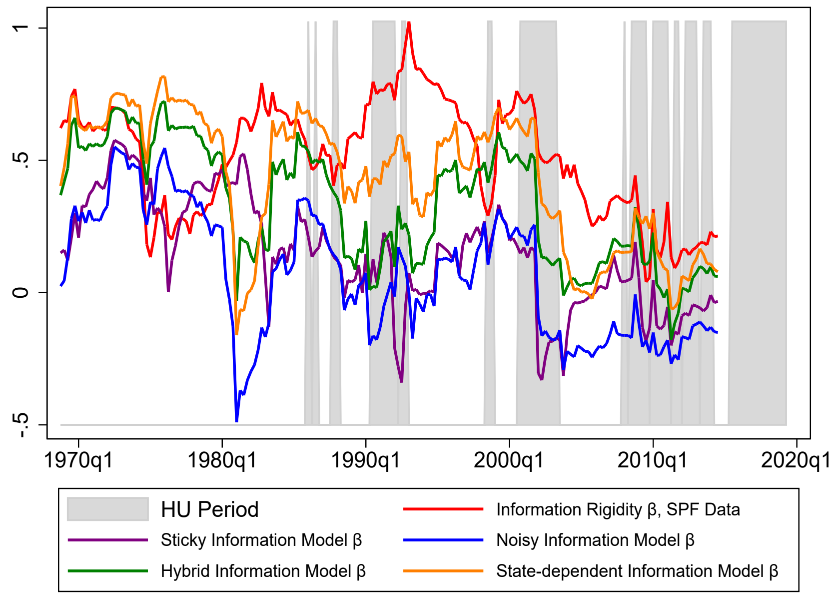

Figure 5: Information Rigidity in SPF vs. Simulated Forecasts
This figure presents coefficients from 5-year rolling regressions of the mean forecast error on its lag, a standard test of information rigidity (Coibion and Gorodnichenko, 2012; Ryngaert, 2017). I estimate the coefficients of information rigidity using SPF data and data from simulations of the sticky, noisy, hybrid sticky-noisy, and the proposed state-dependent sticky-noisy information models at the values of the structural parameters estimated using SMM. The shaded regions indicate time periods to which a Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the high-uncertainty (HU) state.

Finally, to address concerns that the estimated information rigidity coefficients are time-varying in the proposed state-dependent sticky-noisy information model: $\beta_t = (\rho - \lambda_t k_t)$, I also estimate the regression in equation (7.1) using 5-year rolling regressions[16] of the mean forecast error on its lag. For comparison, I conduct similar estimation for the sticky, noisy, and hybrid sticky-noisy information models, simulated at the values of the structural parameters estimated using SMM. The results are presented in Figure 5. As expected from the preceding discussion, information rigidity obtained from the state-dependent sticky-noisy information model is the most successful out of the four alternative models at matching the dynamics of information rigidity in the survey inflation expectations, especially during recent high-uncertainty periods. The three alternative models with information frictions consistently offer worse explanatory power in accounting for the observed information rigidity than the proposed model.

[16]Similar results were obtained from 8-year rolling regressions and 10-year rolling regressions.

8 Policy Relevance and Conclusion

"Monetary policymakers must understand the determinants of inflation in order to attain their inflation goal." (Tootell (1998), p. 21)

The most common framework for analyzing inflation dynamics is the Phillips curve, although it is often criticized. The traditional Phillips curve is backward-looking and features adaptive expectations, as in Friedman (1968) and Phelps (1968). More recently, the micro-founded New-Keynesian Phillips curve includes forward-looking expectations, which allows monetary policy an additional channel through which to affect inflation (Woodford, 2003). A hybrid Phillips curve combines both adaptive and forward-looking expectations, as in Galı and Gertler (1999), Galı et al. (2001), and Galı et al. (2003). The hybrid model is widely used, as in Berganza et al. (2018), Blanchard et al. (2015), IMF (2013), and Mikolajun and Lodge (2016). Albuquerque and Baumann (2017) show that if some firms are backward-looking and set prices based on past values, while others are forward-looking and maximize profits, the hybrid Phillips curve controls for both. According to Forbes (2018), inflation expectations are significantly more correlated with core inflation in the last decade, whereas domestic slack and lagged inflation are less correlated with it. Mikolajun and Lodge (2016) and Bems et al. (2018) find that survey measures of forward-looking inflation expectations are one of the main drivers of inflation in both advanced economies and emerging markets and they perform better at explaining inflation than globalization proxies. Fuhrer (2017) demonstrates that survey expectations improve the performance of standard macroeconomic relationships, as well as DSGE models, and largely eliminate the need for lagged dependent variables reflecting habits, price indexation or autocorrelated structural shocks.

Recently, the Phillips curve has become flatter. Some attribute this to the forces of globalization (Borio and Filardo, 2007; Ahmad and Civelli, 2016; Forbes, 2018). Others find that credible monetary policy has stabilized inflation expectations and trend inflation (Mishkin, 2009). The flatter Phillips curve makes it harder for monetary policy to bring inflation back to target once it deviates, which highlights the importance of well-anchored expectations around the central bank's target (Bean, 2006; Gnan and Valderrama, 2006). Thus, Gnan and Valderrama (2006) have called for "the stabilization of inflation expectations as a primary goal of monetary policy" (p. 37). Yet, as Bean (2006) acknowledges, we know relatively little about how people form their inflation expectations.

This study sheds new light on how economic agents form inflation expectations. Specifically, I propose a micro-founded state-dependent sticky-noisy information model with high-and low-uncertainty regimes. Like existing hybrid sticky-noisy information theories, in the proposed model agents are both rationally inattentive and faced with noisy information. The principal innovation of this model is that the quantity of information represented by attention λ_t and the quality of information that agents glean reflected by the Kalman gain k_t are driven by time-varying uncertainty. As a result, there are interactions between the attention and noisiness effects, which moderate their impact on aggregate information rigidity.

The results using data of professional forecasters include stylized facts demonstrating that forecast errors, attention, and forecaster disagreement (in the latter part of the sample) are state-dependent. Regression results show a highly statistically significant relationship between attention and uncertainty and results from bi-variate VAR confirm that an uncertainty shock causes a significant increase in attention. The standard empirical tests estimating information rigidity yield coefficients that are in line with the literature. Regression analysis including interactions of uncertainty indicators suggests that information rigidity declines with increasing uncertainty. Structural estimation of alternative models with information frictions using Simulated Method of Moments (SMM) demonstrates that the proposed state-dependent model is superior to the hybrid and sticky information models, while preserving some of the desirable features of the noisy information model. The proposed model is also more appealing than the noisy information model due to its ability to account for the observed rational inattention. Hence, all three channels, the direct attention and noisiness and the indirect state-dependence effects, need to be included in theoretical models with information frictions, so that they match the observed information rigidity.

The structural estimation further demonstrates that the attention effect clearly dominates the noisiness effect: in high- (low-) uncertainty periods, the lower (higher) information rigidity attributed to higher (lower) attention outweighs the higher (lower) information rigidity attributed to noisiness by the lower (higher) Kalman gain. As a result, aggregate information rigidity, which is influenced by these two effects in opposite directions, is lower (higher) in high- (low-) uncertainty periods. The fact that more than half of the estimated information rigidity is due to inattention is good news for monetary policymakers. It leaves ample room for monetary policy, which can affect the quantity of information available to economic agents by employing frequent, direct, and simple forward guidance. Some creative recent initiatives include the Bank of Jamaica anchoring expectations by raising awareness about its inflation-targeting policy through reggae music (Whelan, 2019) and the Bank of Finland using Twitter to explain recent European Central Bank rate decisions and what they entail for average citizens in plain Finnish (Weber, 2019).

References

Afrouzi, Hassan (2017). "Strategic inattention, inflation dynamics and the non-neutrality of money." Technical report.

Ahmad, Saad and Andrea Civelli (2016). "Globalization and inflation: A threshold investigation." *Journal of Macroeconomics* 48, 283–304.

Albuquerque, Bruno and Ursel Baumann (2017). "Will us inflation awake from the dead? the role of slack and non-linearities in the phillips curve." *Journal of Policy Modeling* 39(2), 247–271.

Andrade, Philippe and Hervé Le Bihan (2013). "Inattentive professional forecasters." *Journal of Monetary Economics* 60(8), 967–982.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2016). "Measuring economic policy uncertainty." *The Quarterly Journal of Economics* 131(4), 1593–1636.

Baker, Scott R., Tucker S. McElroy, and Xuguang S. Sheng (2019). "Expectation formation following large unexpected shocks." *Review of Economics and Statistics*, 1–45.

Bean, Charles Richard (2006). "Globalisation and inflation." Technical report, Bank of England Quarterly Bulletin, Winter.

Bems, Rudolfs, Francesca G. Caselli, Francesco Grigoli, Bertrand Gruss, and Weicheng Lian (2018). "Is inflation domestic or global? evidence from emerging markets." Technical report, International Monetary Fund Working Paper No. 18/241.

Berganza, Juan Carlos, Pedro del Ro, and Fructuoso Borrallo (2018). *Determinants and Implications of Low Global Inflation Rates*, pp. 183–216. Springer, Cham.

Bernanke, Ben S. (1983). "Irreversibility, uncertainty, and cyclical investment." *The Quarterly Journal of Economics* 98(1), 85–106.

Blanchard, Olivier, Eugenio Cerutti, and Lawrence Summers (2015). "Inflation and activity–two explorations and their monetary policy implications." Technical report, National Bureau of Economic Research.

Bloom, Nicholas (2009). "The impact of uncertainty shocks." *Econometrica* 77(3), 623–685.

Borio, Claudio E.V. and Andrew J. Filardo (2007). "Globalisation and inflation: New cross-country evidence on the global determinants of domestic inflation." Technical report, BIS working paper.

Branch, William A. (2004). "The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations." *The Economic Journal* 114(497), 592–621.

Branch, William A., John Carlson, George W. Evans, and Bruce McGough (2009). "Monetary policy, endogenous inattention and the volatility tradeoff." *The Economic Journal* 119(534), 123–157.

Carroll, Christopher D. (2003). "Macroeconomic expectations of households and professional forecasters." *The Quarterly Journal of Economics* 118(1), 269–298.

Caskey, John (1985). "Modeling the formation of price expectations: A Bayesian approach." *The American Economic Review* 75(4), 768–776.

Coibion, Olivier and Yuriy Gorodnichenko (2012). "What can survey forecasts tell us about informational rigidities?" *Journal of Political Economy* 120(1), 116–159.

Coibion, Olivier and Yuriy Gorodnichenko (2015). "Information rigidity and the expectations formation process: A simple framework and new facts." *The American Economic Review* 105(8), 2644–2678.

Diebold, Francis X., Frank Schorfheide, and Minchul Shin (2016). "Real-time forecast evaluation of dsge models with stochastic volatility." Technical Report w22615, National Bureau of Economic Research.

Dovern, Jonas, Ulrich Fritsche, Prakash Loungani, and Natalia Tamirisa (2015). "Information rigidities: Comparing average and individual forecasts for a large international panel." *International Journal of Forecasting* 31(1), 144–154.

Duffie, Darrell and Kenneth Singleton (1993). "Simulated moments estimation of markov models of asset prices." *Econometrica* 61(4), 929–52.

Elliott, Graham, Thomas J Rothenberg, and James H Stock (1996). "Efficient tests for an autoregressive unit root." *Econometrica* 64, 813–836.

Forbes, Kristin J. (2018). "Has globalization changed the inflation process." Technical report, Working Paper.

Frankel, Jeffrey A. and Kenneth A. Froot (1987). "Using standard survey data to test standard propositions regarding exchange rate expectations." *The American Economic Review* 77(1), 133–153.

Friedman, Milton (1968). "The role of monetary policy." *The American Economic Review* 58(1), 1–17.

Fuhrer, Jeff (2017). "Expectations as a source of macroeconomic persistence: Evidence from survey expectations in a dynamic macro model." *Journal of Monetary Economics* 86, 22–35.

Fuhrer, Jeff (2018). "Intrinsic expectations persistence: Evidence from professional and household survey expectations." Technical Report 18-9, Federal Reserve Band of Boston Working Paper.

Galı, Jordi and Mark Gertler (1999). "Inflation dynamics: A structural econometric analysis." *Journal of Monetary Economics* 44(2), 195–222.

Galı, Jordi, Mark Gertler, and J. David Lopez-Salido (2001). "European inflation dynamics." *European economic review* 45(7), 1237–1270.

Galı, Jordi, Mark Gertler, and J. David Lopez-Salido (2003). "Erratum to 'european inflation dynamics'." *European economic review* 47(4), 759–760.

Giacomini, Raffaella, Vasiliki Skreta, and Javier Turén (2020). "Heterogeneity, inattention and bayesian updates." *American Economic Journal: Macroeconomics*.

Gnan, Ernest and Maria Teresa Valderrama (2006). "Globalization, inflation and monetary policy." *Monetary Policy & the Economy Q* 4, 37–54.

Gorodnichenko, Yuriy (2008). "Endogenous information, menu costs and inflation persistence." Technical Report w14184, National Bureau of Economic Research.

Gourieroux, Christian and Alain Monfort (1996). *Simulation-based Econometric Methods*. Oxford University Press.

Hamilton, James D. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle." *Econometrica* 57(2), 357–384.

Hamilton, James Douglas (1994). *Time series analysis*. Princeton: Princeton University Press.

Ilut, Cosmin L. and Martin Schneider (2014). "Ambiguous business cycles." *The American Economic Review* 104(8), 2368–2399.

IMF (2013). "The dog that didn't bark: Has inflation been muzzled or was it just sleeping?" Technical report, World Economic Outlook, Chapter 3, April.

Krolzig, Hans-Martin (2013). *Markov-switching vector autoregressions: Modelling, statistical inference, and application to business cycle analysis*, Volume 454. Springer Science & Business Media.

Kydland, Finn E. and Edward C. Prescott (1982). "Time to build and aggregate fluctuations." *Econometrica* 50(6), 1345–1370.

Lahiri, Kajal and Xuguang Simon Sheng (2008). "Evolution of forecast disagreement in a Bayesian learning model." *Journal of Econometrics* 144(2), 325–340.

Lucas, Robert E. (1972). "Expectations and the neutrality of money." *Journal of Economic Theory* 4(2), 103–124.

Maćkowiak, Bartosz and Mirko Wiederholt (2009). "Optimal sticky prices under rational inattention." *The American Economic Review* 99(3), 769–803.

Maćkowiak, Bartosz and Mirko Wiederholt (2012). "Information processing and limited liability." *The American Economic Review* 102(3), 30–34.

Maćkowiak, Bartosz and Mirko Wiederholt (2015). "Business cycle dynamics under rational inattention." *The Review of Economic Studies* 82(4), 1502–1532.

Mankiw, N. Gregory and Ricardo Reis (2002). "Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve." *The Quarterly Journal of Economics* 117(4), 1295–1328.

Mikolajun, Irena and David Lodge (2016). "Advanced economy inflation: the role of global factors." Technical report, ECB Working Paper No. 1948.

Mishkin, Frederic S. (2009). "Globalization, macroeconomic performance, and monetary policy." *Journal of Money, Credit and Banking* 41, 187–196.

Nimark, Kristoffer (2014). "Man-bites-dog business cycles." *The American Economic Review* 104(8), 2320–2367.

Nordhaus, William D. (1987). "Forecasting efficiency: concepts and applications." *The Review of Economics and Statistics* 69(4), 667674.

Perlin, Marcelo (2015). "Ms regress - the matlab package for markov regime switching models." Technical report, SSRN: http://ssrn.com/abstract=1714016 or http://dx.doi.org/10.2139/ssrn.1714016.

Pesaran, M. Hashem and Martin Weale (2006). *Survey Expectations*. North-Holland Press.

Phelps, Edmund S. (1968). "Money-wage dynamics and labor-market equilibrium." *Journal of political economy* 76(4, Part 2), 678–711.

Reis, Ricardo (2006). "Inattentive producers." *The Review of Economic Studies* 73(3), 793–821.

Romer, Christina D. (1990). "The great crash and the onset of the great depression." *The Quarterly Journal of Economics* 105(3), 597–624.

Ruge-Murcia, Francisco (2012). "Estimating nonlinear dsge models by the simulated method of moments: With an application to business cycles." *Journal of Economic Dynamics and Control* 36(6), 914–938.

Ryngaert, Jane (2017). "What do (and don't) forecasters know about u.s. inflation." Technical report, Job Market Paper Draft.

Sargent, Thomas J. (1993). *Bounded Rationality in Macroeconomics*. Oxford: Oxford University Press.

Sims, Christopher A. (2003). "Implications of rational inattention." *Journal of Monetary Economics* 50(3), 665–690.

Tootell, Geoffrey (1998). "Globalization and u.s. inflation." *New England Economic Review* Jul/Aug, 21–33.

Van Nieuwerburgh, Stijn and Laura Veldkamp (2006). "Learning asymmetries in real business cycles." *Journal of Monetary Economics* 53(4), 753–772.

Weber, Michael (2019). "What's (not) up with inflation? - transcript." *The Brookings Institution*.

Whelan, Robbie (2019). "Keep de rates dem low-jamaica sets inflation fight to reggae beat." *The Wall Street Journal*.

Woodford, Michael (2002). *Imperfect Common Knowledge and the Effects of Monetary Policy.* Princeton: Princeton University Press.

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

Woodford, Michael (2009). "Information-constrained state-dependent pricing." *Journal of Monetary Economics* (S), S100–S124.

Appendix A Additional Figures

Figure A.1: Frequency of Forecaster Participation

The figure shows the frequency of forecaster participation in the sample. Individual quarterly forecasts are obtained from the U.S. Survey of Professional Forecasters (SPF) in the period Q4 1968 - Q2 2019. The sample size varies between 9 and 131 individual forecasters, with an average of 46 forecasters, every quarter and includes 443 different individual forecasters in total, 70 of whom, pictured in red, report forecasts in at least 40 periods. The panel data set is unbalanced, as forecasters enter, exit, and merge during the survey period.

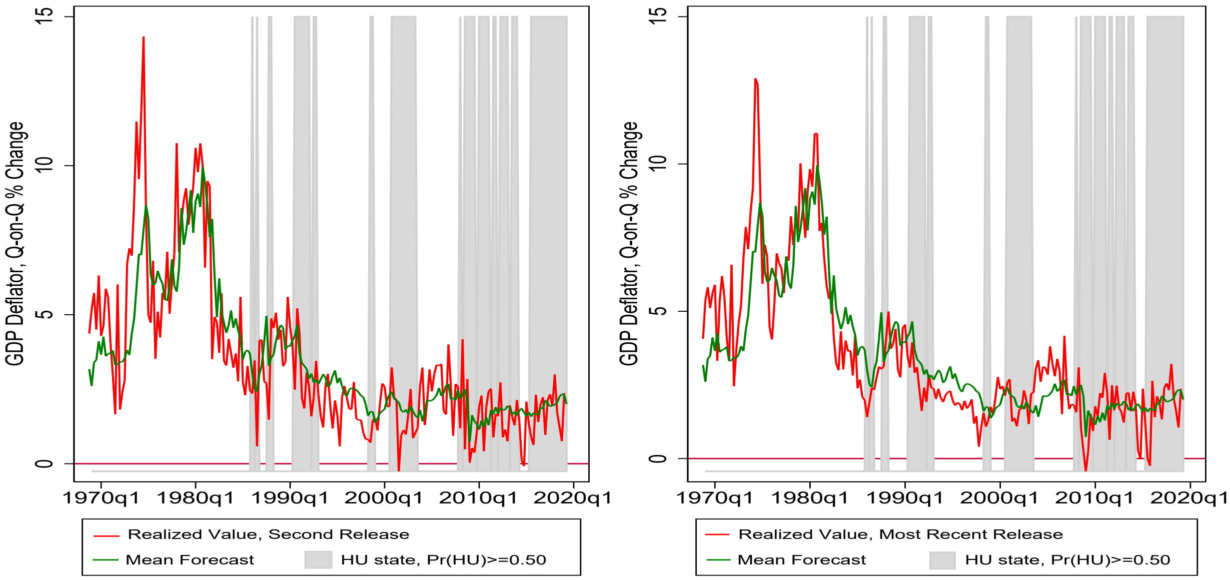

Figure A.2: Inflation Realizations vs. Mean SPF Forecasts

The figure shows the time series of realized second- and most recent-release values vs. the cross-sectional mean of individual forecasts of GDP deflator inflation for each quarter. The releases of actual data are from the Federal Reserve Bank of Philadelphia's real-time data set. Individual quarterly forecasts are obtained from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF) in the period Q4 1968 - Q2 2019. The shaded regions indicate periods to which the Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the high-uncertainty (HU) regime.

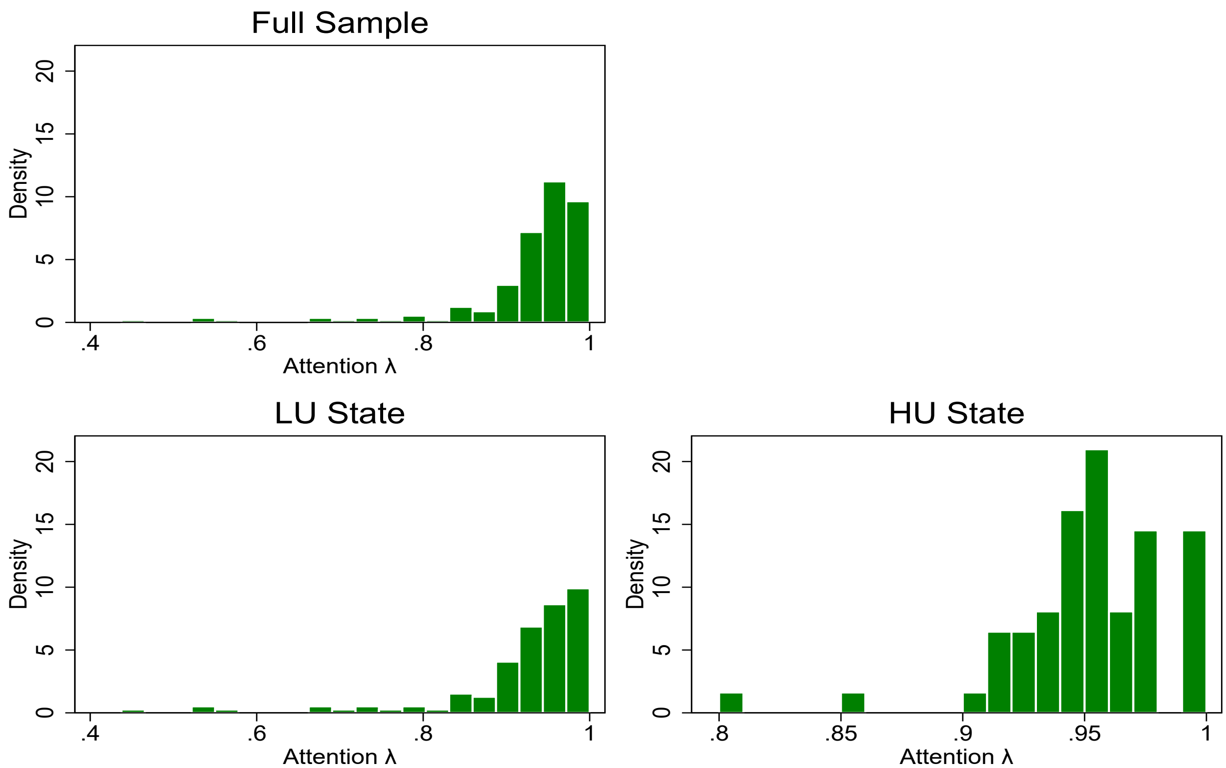

Figure A.3: Histograms of Empirical Attention λ_t
The figure presents the distribution of empirical attention λ_t in GDP deflator inflation forecasts in the full sample, as well as during high-uncertainty (HU) and low-uncertainty (LU) periods. The degree of attention λ_t is expressed as the proportion of forecasters who update their forecasts and nowcasts of GDP deflator inflation out of all forecasters in the sample for each quarter. Missing values of the nowcasts and forecasts are interpreted as "not updating." Whenever forecasters report a nowcast and a forecast that are different from the ones they had made the previous period, this is interpreted as "updating." Individual quarterly forecasts are obtained from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF) in the period Q4 1968 - Q2 2019. HU periods are periods to which the Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the HU state. LU periods are all the remaining periods in the sample.

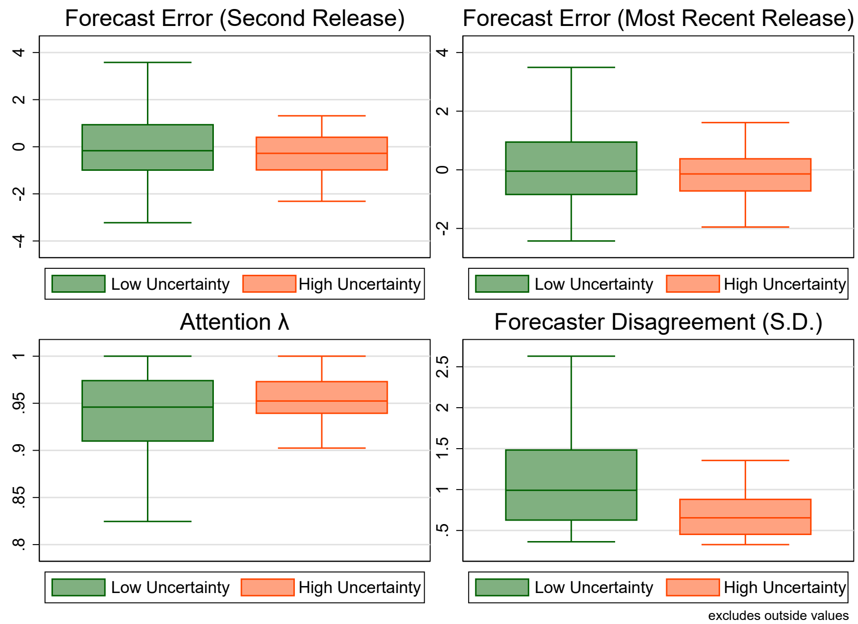

Figure A.4: State-dependence in Forecast Errors, Empirical Attention, and Disagreement
The figure presents box plots of the means and standard deviations of mean forecast errors (using second and final release of realized data), attention λ_t, and forecaster disagreement during the LU and HU sub-samples, as presented in Table 1. Mean forecast errors are obtained by subtracting the individual quarterly forecasts from the actual data and averaging across forecasters. Attention λ_t is calculated as the proportion of forecasters who update their forecasts and nowcasts out of all forecasters in the sample each quarter. Forecaster disagreement refers to the cross-sectional standard deviation of forecasts each quarter. All graphs exclude outside values.

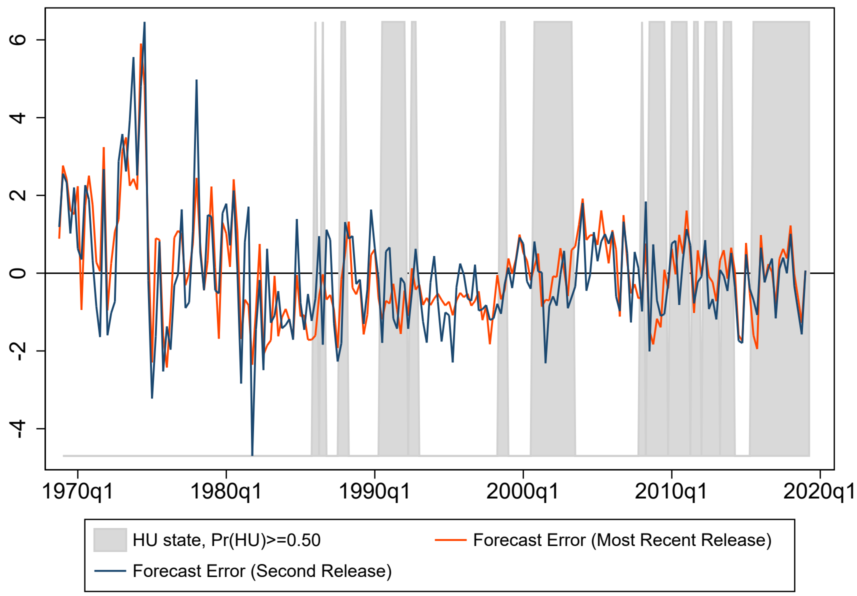

Figure A.5: Mean Forecast Errors

The figure presents time series of the mean forecast errors of forecasts of GDP deflator inflation using the second and the most recent release of realized data. Mean forecast errors are obtained by subtracting the individual quarterly forecasts from the second-release and most recent-release actual data and averaging across forecasters. The releases of actual data are from the Federal Reserve Bank of Philadelphia's real-time data set. Individual quarterly forecasts are obtained from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF) in the period Q4 1968 - Q2 2019. The shaded regions indicate periods to which the Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the high-uncertainty (HU) regime.

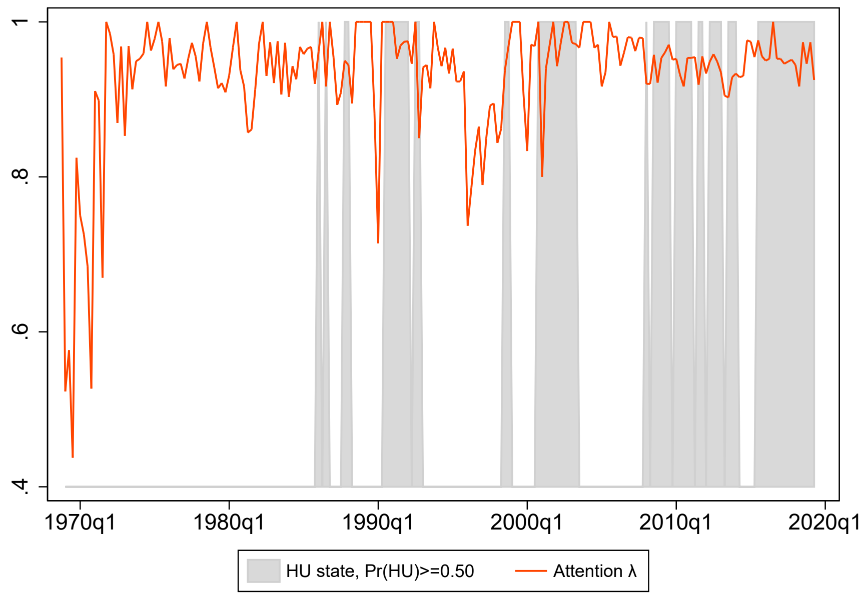

Figure A.6: Empirical Attention λ_t

The figure presents time series of the empirical degree of attention λ_t, expressed as the proportion of forecasters who update their forecasts and nowcasts of GDP deflator inflation out of all forecasters in the sample for each quarter. Missing values of the nowcasts and forecasts are interpreted as "not updating." Whenever forecasters report a nowcast and forecast that are different from the ones they had made the previous period, this is interpreted as "updating." Individual quarterly forecasts are obtained from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF) in the period Q4 1968 - Q2 2019. The shaded regions indicate periods to which the Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the high-uncertainty (HU) regime.

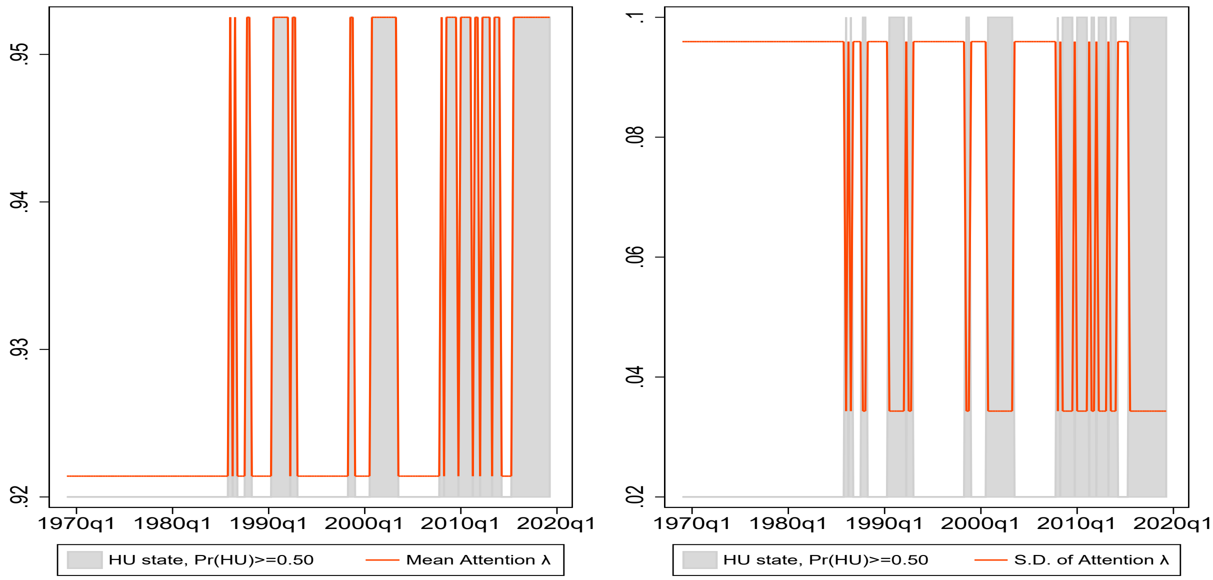

Figure A.7: Empirical Attention λ_t: Mean and Standard Deviation

This figure presents the empirical observations noted in Table 1 about the mean and standard deviation of the empirical attention λ_t across all LU vs. all HU periods. Empirical attention λ_t is expressed as the proportion of forecasters who update their forecasts and nowcasts of GDP deflator inflation out of all forecasters in the sample for each quarter. Missing values of the nowcasts and forecasts are interpreted as "not updating." Whenever forecasters report a nowcast and forecast that are different from the ones they had made the previous period, this is interpreted as "updating." Individual quarterly forecasts are obtained from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF) in the period Q4 1968 - Q2 2019. The shaded regions indicate periods to which the Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the high-uncertainty (HU) regime.

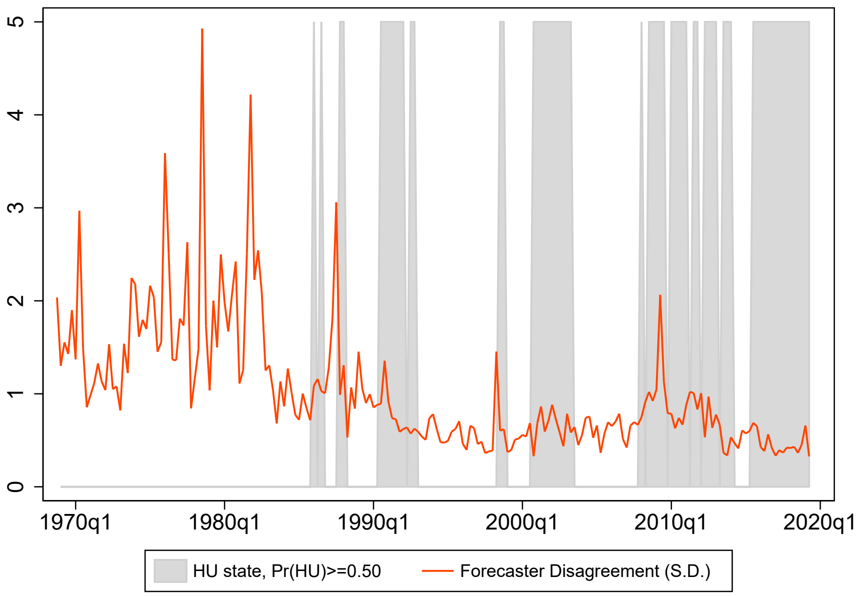

Figure A.8: Forecaster Disagreement

The figure presents time series of forecaster disagreement, the cross-sectional standard deviation of forecasts for each quarter. Individual quarterly forecasts are obtained from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF) in the period Q4 1968 - Q2 2019. The shaded regions indicate periods to which the Markov-switching dynamic autoregression from Appendix B assigns a probability of at least 50 percent of being in the high-uncertainty (HU) regime.

Appendix B Uncertainty-based State-dependence

A Markov-switching dynamic autoregression model[17] is used to estimate the probabilities of the latent or hidden state of economic uncertainty being low uncertainty (LU) or high uncertainty (HU), $\sigma_t^2 \in \{\sigma_{LU}^2, \sigma_{HU}^2\}$, for each time t in the sample, conditional on an "observed" time-t signal x_t for the state of economic uncertainty and parameters θ, $\pi_{j|t} = P\{\sigma_t^2 = \sigma_j^2 | x_t, \theta\}$, where $j \in \{LU, HU\}$. Empirical uncertainty indicators are used to approximate this "observed" variable x_t (the common costless signal in equation (D.1)). Section 4.2 incorporates the Markov-switching model into the proposed theoretical hybrid sticky-noisy information model.

Similarly to one of the proposed volatility specifications in Diebold et al. (2016), volatility follows a Markov-switching dynamic autoregression model:

$$x_t = \mu_{\sigma_t} + \phi x_{t-1} + \epsilon_t, \tag{B.1}$$

with a state-dependent mean μ_{σ_t} and error term $\epsilon_t \sim i.i.d.\ \mathcal{N}(0,\, \sigma_{\sigma_t}^2)$ and state-independent coefficient ϕ. The "observed" signal x_t about the hidden state σ_t^2 at time t is approximated with the extended EPU and VXO indices.

Table B1 presents results of Markov-switching dynamic autoregressions with the EPU and VXO indices as the "observed" signals for the hidden state of economic uncertainty. All specification include regime-switching mean and variance of the error term and non-switching coefficient on the AR(1) term, as specified in equation (B.1). Model (1) in Table B1, which uses the EPU index, is taken as the baseline. The EPU is preferred to the VXO because it is 'more exogenous' from the expectation formation process of professional forecasters, many of whom are financial-sector institutions. Approximating uncertainty with the EPU is less likely to involve confounding variable bias than the VXO, which may be jointly influenced with the degree of attention financial analysts pay to macroeconomic conditions, by other variables. Instead, the VXO is used as a robustness check on the baseline model, which does not yield very different state transition probabilities. Moreover, based on the SBIC, the baseline model is preferred to models with switching coefficients on the AR(1) term, models with non-switching variances of the error terms, models with three instead of two uncertainty states, and Markov-switching mean models (Krolzig, 2013). The baseline results are neither significantly affected by the extension of the baseline EPU index with the historical EPU index, nor by the conversion from monthly to quarterly frequency.

[17] Markov-switching models, also referred to as Hidden Markov Models in the machine learning literature, are explained in detail in Hamilton (1994, Chapter 22); Hamilton (1989) provides an application.

Table B1: Markov-switching Autoregression Based on EPU and VXO

	(1) Extended EPU		(2) Extended VXO	
	LU	HU	LU	HU
μ_{σ_t}	34.29***	59.90***	5.668***	10.344***
	(4.734)	(8.077)	(0.474)	(1.137)
Lag. Uncertainty Proxy x_{t-1}	0.580***		0.612***	
	(0.0549)		(0.0213)	
State-dep. St. Dev. σ_{σ_t}	14.781	30.686	1.662	5.570
	(1.074)	(3.385)	(0.331)	(1.139)
$P(\sigma_t^2 = \sigma_{LU}^2 \mid \sigma_{t-1}^2 = \sigma_{LU}^2)$	0.944		0.843	
	(0.033)		(0.051)	
$P(\sigma_t^2 = \sigma_{HU}^2 \mid \sigma_{t-1}^2 = \sigma_{HU}^2)$		0.896		0.780
		(0.068)		(0.144)
Expected Duration	17.937	9.637	6.386	4.546
	(10.588)	(6.330)	(2.096)	(2.980)
Observations	202		202	
SBIC	9.150		5.512	

Robust standard errors in parentheses *** p< 0.01, ** p<0.05, * p<0.1.

The table presents results of Markov-switching dynamic autoregressions of the EPU and VXO indices. All specifications include switching constant and variance of the error term and non-switching coefficient on the AR(1) term. The probabilities listed are the transitional probabilities that the LU and HU will persist from one period to the next. The expected duration is in quarters. The EPU index is compiled by Baker et al. (2016). The VXO is obtained from the Federal Reserve Economic Data (FRED). Quarterly values are 3-month averages of monthly values.

According to the baseline model, the probability of a LU regime of continuing from one period to the next is 0.944 and the probability of a HU regime of continuing from one period to the next is 0.896. The average duration of a LU period is about 18 quarters and the average duration of a HU period is about 10 quarters. For comparison, Hamilton (1989) applies a Markov-switching model to U.S. GNP data and estimates transitional probabilities of an expansion equal to 0.90 and of a recession equal to 0.75. These transitional probabilities correspond to average durations of 30 months for expansions and 12 months for recessions. As discussed below, these durations do not match the ones estimated here; however, they do not have to. Periods of high economic uncertainty are correlated with recessionary periods but the correlation is not very strong, since economic uncertainty may increase earlier and persist longer than periods of actual decline in real economic activity, or it may vary within recession periods.

Figure B.1 examines the fit of the baseline model by presenting the fitted values of the EPU index, the actual values, and the residuals. The fitted values are obtained using a smoothing algorithm based on all sample data. The model shows a relatively good fit and for the most

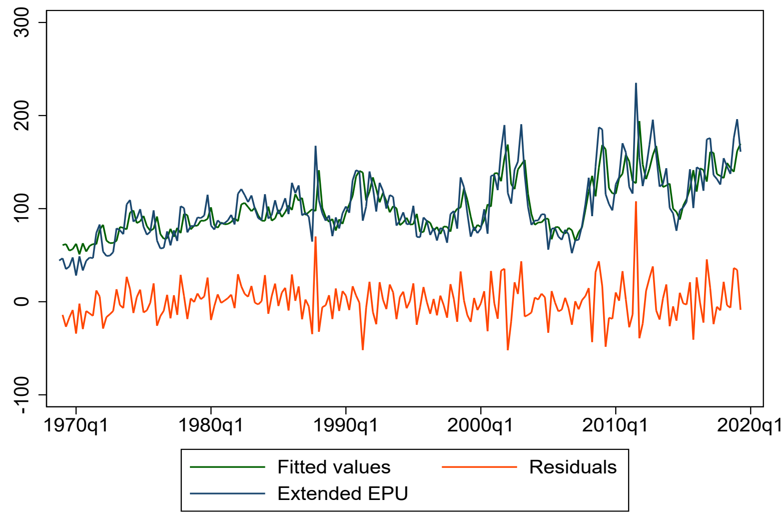

Figure B.1: Fit of Baseline Markov-switching Autoregression Model
The figure shows fitted values of the EPU index, actual values, and residuals of the baseline Markov-switching dynamic autoregression model. The fitted values are obtained using a smoothing algorithm based on all sample data. The EPU index is compiled by Baker et al. (2016). Quarterly values are 3-month averages of monthly values.

part, the residuals do not seem to account for much of the dependent variable variation. This result is robust to obtaining the fitted values and residuals using only information from previous periods, as well as from previous and contemporaneous periods.

Using the baseline Markov-switching dynamic autoregression model, the probability of being in the HU regime is compared to recessionary periods, as defined by the National Bureau of Economic Research (NBER) in Figure B.2. The probability of being in the HU regime does not exactly match the NBER-designated recessions and it does not have to, since economic uncertainty may increase earlier and persist long after the defined periods of decline in actual economic activity, or it may vary significantly within such episodes. Instead, as argued by Bloom (2009), recessions may be caused by economic uncertainty shocks, among other factors.

Based on the estimated probabilities using the baseline model, the threshold probability, above which the state is considered more likely to be the HU regime, is designated as $\pi_{HU|t} = P\{\sigma_t^2 = \sigma_{HU}^2\} \geq 0.50$. Figure B.3 presents these estimated HU periods using the baseline Markov-switching autoregression model as the shaded regions, along with the time series of the two uncertainty proxies, the EPU and VXO indices. Despite the fact that the HU periods are estimated only on the basis of the EPU index and do not take the VXO under consideration, they still seem to capture well most of the increases in economic uncertainty, as represented by both of the empirical proxies.

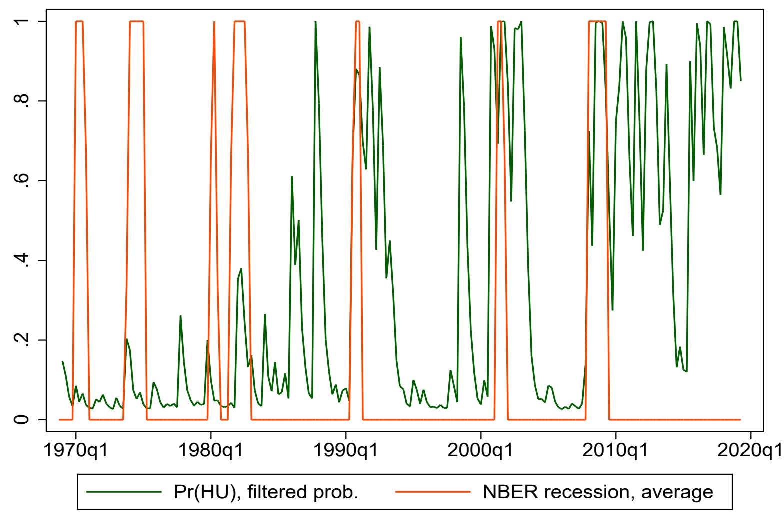

Figure B.2: Estimated HU Regime Probabilities vs. Recessions (NBER)
The figure presents the probability of being in the HU regime, estimated using the baseline Markov-switching autoregression model compared to recessionary periods, as defined by the National Bureau of Economic Research (NBER). The dummy recession indicator is obtained from the Federal Reserve Economic Data (FRED). Quarterly values of the recession indicator are 3-month averages of monthly values.

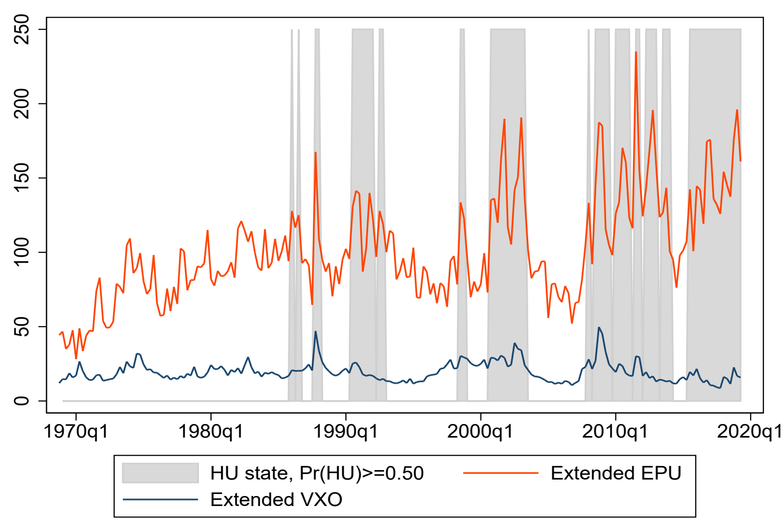

Figure B.3: Estimated HU Periods vs. Uncertainty Proxies
The figure presents estimated HU periods, in which the conditional probability of being in the HU regime $\pi_{HU|t} = P\{\sigma_t^2 = \sigma_{HU}^2\} \geq 0.50$, using the baseline Markov-switching autoregression model estimated on the basis of the EPU index (shaded regions), along with the time series of the EPU and VXO indices. The EPU index is compiled by Baker et al. (2016). The VXO is obtained from the Federal Reserve Economic Data (FRED). Quarterly values are 3-month averages of monthly values.

Appendix C Additional Tests of Information Rigidity

The second empirical test of information rigidity expresses the predictability of the average ex-post forecast error with respect to the mean ex-ante forecast revision, as derived by Coibion and Gorodnichenko (2015):

$$MeanForecastError_{t+1|t} = \alpha^b + \beta^b MeanForecastRevision_{t+1|t} + \delta_t + \epsilon_t, \tag{C.1}$$

where α^b is a constant, $MeanForecastError_{t+1|t} = ActualValue_{t+1} - MeanForecast_{t+1|t}$, $MeanForecastRevision_{t+1|t} = MeanForecast_{t+1|t} - MeanForecast_{t|t-1}$, δ_t are quarter dummies to account for any seasonality effect in the data, and ϵ_t is the error term in the regression.

Table C1: Information Rigidity Test Based on Forecast Error and Forecast Revision: Interactions

	FE (First Release)				FE (Second Release)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Revision	0.485*	0.454	0.480	0.685	0.740**	0.712**	0.733*	0.975
	(0.281)	(0.307)	(0.365)	(0.686)	(0.297)	(0.323)	(0.387)	(0.718)
Rev×HU Indicator		-0.146	-0.197	-0.00271		-0.200	-0.208	-0.00316
		(0.415)	(0.604)	(0.00583)		(0.504)	(0.697)	(0.00632)
HU Dummy		-0.369**				-0.407**		
		(0.176)				(0.185)		
Prob(HU)			-0.415*				-0.462*	
			(0.239)				(0.245)	
EPU Index				-0.00559**				-0.00646***
				(0.00222)				(0.00235)
Constant	0.137	0.248	0.274	0.705**	0.223	0.344	0.376	0.879**
	(0.192)	(0.213)	(0.228)	(0.323)	(0.204)	(0.226)	(0.241)	(0.342)
Observations	200	200	200	200	199	199	199	199
Quarter Dummies	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.054	0.068	0.064	0.073	0.084	0.098	0.095	0.106
Adjusted R-squared	0.0345	0.0386	0.0354	0.0438	0.0653	0.0702	0.0672	0.0781
	FE (Third Release)				FE (Final Release)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Revision	0.688**	0.660**	0.670*	0.877	0.558***	0.564**	0.658**	0.706
	(0.305)	(0.332)	(0.397)	(0.731)	(0.214)	(0.232)	(0.280)	(0.698)
Rev×HU Indicator		-0.197	-0.163	-0.00275		-0.386	-0.687	-0.00245
		(0.532)	(0.725)	(0.00648)		(0.415)	(0.567)	(0.00627)
HU Dummy		-0.395**				-0.318**		
		(0.183)				(0.159)		
Prob(HU)			-0.459*				-0.469**	
			(0.243)				(0.213)	
EPU Index				-0.00663***				-0.00772***
				(0.00236)				(0.00219)
Constant	0.221	0.332	0.369	0.883***	0.159	0.241	0.292	0.948***
	(0.214)	(0.233)	(0.246)	(0.339)	(0.176)	(0.195)	(0.209)	(0.313)
Observations	200	200	200	200	201	201	201	201
Quarter Dummies	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.073	0.087	0.085	0.096	0.053	0.065	0.071	0.096
Adjusted R-squared	0.0543	0.0586	0.0562	0.0677	0.0332	0.0360	0.0422	0.0678

Robust standard errors in parentheses *** p< 0.01, ** p<0.05, * p<0.1.

The table presents results from OLS regressions of mean forecast errors on mean forecast revisions, following Coibion and Gorodnichenko (2015). HU Dummy = dummy for periods with probability of HU regime at least 50 percent (see Appendix B); Prob(HU) = probability of HU regime (see Appendix B); EPU Index = uncertainty proxy. The mean forecast errors are obtained by subtracting the individual quarterly forecasts from the first-, second-, third-, and final-release actual data and averaging across forecasters. The mean forecast revisions are obtained by subtracting the one-quarter lagged mean forecast from the contemporaneous mean forecast. All regressions include quarter dummies and heteroskedasticity-robust Huber-White standard errors.

The third empirical test for estimating information rigidity in inflation expectations regresses the contemporaneous mean forecast revision on its one-quarter lag. This empirical test, offered by Dovern et al. (2015) and Nordhaus (1987), avoids the issue of realized data revisions and takes the form:

$$MeanForecastRevision_{t+1|t} = \alpha^c + \beta^c MeanForecastRevision_{t|t-1} + \delta_t + \epsilon_t, \tag{C.2}$$

where $MeanForecastRevision_{t+1|t} = MeanForecast_{t+1|t} - MeanForecast_{t|t-1}$, $MeanForecastRevision_{t|t-1} = MeanForecast_{t|t-1} - MeanForecast_{t-1|t-2}$, α^c is a constant, δ_t are quarter dummies to account for any seasonality effect in the data, and ϵ_t is the error term in the regression.Dovern et al. (2015) demonstrate that Nordhaus (1987)'s rationality test of regressing the contemporaneous mean forecast revision on its lag can be used to derive similar implications as the Coibion and Gorodnichenko (2015) regressions of average forecast errors on mean forecast revisions.

Table C2: Information Rigidity Test Based on Forecast Revision: Interactions

	Forecast Revision			
	(1)	(2)	(3)	(4)
Lag Revision	0.181	0.239*	0.286**	0.175
	(0.119)	(0.132)	(0.130)	(0.391)
Lag Rev×HU Indicator		-0.481**	-0.547*	-3.95e-05
		(0.231)	(0.318)	(0.00444)
HU Dummy		-0.165***		
		(0.0618)		
Prob(HU)			-0.213***	
			(0.0771)	
EPU Index				-0.00245***
				(0.000845)
Constant	-0.0724	-0.0358	-0.0115	0.185
	(0.0692)	(0.0783)	(0.0801)	(0.118)
Observations	201	201	201	201
Quarter Dummies	YES	YES	YES	YES
R-squared	0.039	0.085	0.083	0.070
Adjusted R-squared	0.0198	0.0563	0.0544	0.0409

Robust standard errors in parentheses *** p< 0.01, ** p<0.05, * p<0.1.

The table presents results from OLS regressions of mean forecast revisions on one-quarter lagged mean forecast revisions, following Dovern et al. (2015) and Nordhaus (1987). High uncertainty (HU) indicators include: HU Dummy = dummy for periods with probability of HU regime at least 50 percent (see Appendix B); Prob(HU) = probability of HU regime (see Appendix B); EPU Index = uncertainty proxy. The mean forecast revisions are obtained by subtracting the mean one-quarter lagged forecast from the contemporaneous mean forecast. All regressions include quarter dummies and heteroskedasticity-robust Huber-White standard errors.

Appendix D Derivations

D.1 Micro-foundation Solution

Lagrangian:

$$L(\lambda_t, \mu_1, \mu_2) = -[(1 - \rho^{-1}\lambda_t k_t)\sigma^2_{t|t-1} + (\lambda_t k_t)^2] + \mu_1(1 - \lambda_t) + \mu_2(0 + \lambda_t)$$

The Karush-Kuhn-Tucker conditions are:

1. FOC:
$$\begin{aligned}\frac{\partial L}{\partial \lambda_t} &= -(-\rho^{-1}k_t\sigma^2_{t|t-1} + 2\lambda_t k_t^2) - \mu_1 + \mu_2 = 0 \\ &= \rho^{-1}k_t\sigma^2_{t|t-1} - 2\lambda_t k_t^2 - \mu_1 + \mu_2 = 0\end{aligned}$$

2. Constraints:
$$\begin{aligned}\lambda_t &\leq 1 \\ -\lambda_t &\leq 0\end{aligned}$$

3. Complementary slackness conditions:
$$\begin{aligned}\mu_1, \mu_2 &\geq 0 \\ \mu_1(1 - \lambda_t) &= 0 \\ \mu_2\lambda_t &= 0\end{aligned}$$

Four cases:

1. $\mu_1 = 0, \mu_2 = 0$
2. $\mu_1 \neq 0, \mu_2 = 0$
3. $\mu_1 = 0, \mu_2 \neq 0$
4. $\mu_1 \neq 0, \mu_2 \neq 0$

1. $\mu_1 = 0, \mu_2 = 0$: both constraints are nonbinding
   From FOC:

$$\begin{aligned} \rho^{-1} k_t \sigma^2_{t|t-1} - 2\lambda_t k_t^2 = 0 \\ 2\lambda_t k_t^2 = \rho^{-1} k_t \sigma^2_{t|t-1} \\ \lambda_t = \frac{\rho^{-1}\sigma^2_{t|t-1}}{2k_t} \end{aligned}$$

   The objective function with $\lambda_t = \frac{\rho^{-1}\sigma^2_{t|t-1}}{2k_t}$ becomes:

$$\begin{aligned} &= -\left[\sigma^2_{t|t-1} - \frac{\rho^{-1} \times \rho^{-1}\sigma^2_{t|t-1} \times k_t\sigma^2_{t|t-1}}{2k_t} + \left(\frac{\rho^{-1}\sigma^2_{t|t-1} \times k_t}{2k_t}\right)^2\right] \\ &= -\left[\sigma^2_{t|t-1} - \frac{\rho^{-2}(\sigma^2_{t|t-1})^2}{2} + \frac{\rho^{-2}(\sigma^2_{t|t-1})^2}{4}\right] \\ &= -\left[\sigma^2_{t|t-1} - \frac{2\rho^{-2}(\sigma^2_{t|t-1})^2 - \rho^{-2}(\sigma^2_{t|t-1})^2}{4}\right] \\ &= -\left[\sigma^2_{t|t-1} - \frac{(\sigma^2_{t|t-1})^2}{4\rho^2}\right]. \end{aligned}$$

2. $\mu_1 \neq 0, \mu_2 = 0$: $(1 - \lambda_t) = 0$ is binding
   From Complementary slackness conditions:

$$\begin{aligned} (1 - \lambda_t) = 0 \\ -\lambda_t = -1 \\ \lambda_t = 1 \end{aligned}$$

   From FOC:

$$\begin{aligned} \rho^{-1} k_t \sigma^2_{t|t-1} - 2k_t^2 - \mu_1 = 0 \\ \mu_1 = \rho^{-1} k_t \sigma^2_{t|t-1} - 2k_t^2 \\ \mu_1 = k_t(\rho^{-1}\sigma^2_{t|t-1} - 2k_t) \end{aligned}$$

   and since $0 \leq k_t \leq 1 \to \mu_1 \geq 0$ only when $\rho^{-1}\sigma^2_{t|t-1} \geq 2k_t$.

$$\begin{aligned} \text{Thus, } 2k_t \leq \rho^{-1}\sigma^2_{t|t-1} \\ k_t \leq \frac{\rho^{-1}\sigma^2_{t|t-1}}{2}. \end{aligned}$$

Substitute for k_t from equation (4.10):

$$\begin{aligned}
\frac{\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1}+\sigma^2_t} &\leq \frac{\rho^{-1}\sigma^2_{t|t-1}}{2} \\
\frac{\rho}{\sigma^2_{t|t-1}+\sigma^2_t} &\leq \frac{1}{2\rho} \\
\sigma^2_{t|t-1}+\sigma^2_t &\geq 2\rho^2 \\
\sigma^2_t &\geq 2\rho^2-\sigma^2_{t|t-1}.
\end{aligned}$$

Attention is complete ($\lambda_t = 1$) only when uncertainty σ^2_t is equal to or above this threshold. Otherwise, the solution of Case 2 is not feasible. The objective function with $\lambda_t = 1$ becomes:

$$-\left[\sigma^2_{t|t-1}-\rho^{-1}k_t\sigma^2_{t|t-1}+k^2_t\right].$$

It is optimal for agents to pay complete attention ($\lambda_t = 1$) only when the value of the objective function in Case 2 is less than the value of the objective function in Case 1. Specifically:

$$\begin{aligned}
\sigma^2_{t|t-1}-\rho^{-1}k_t\sigma^2_{t|t-1}+k^2_t &< \sigma^2_{t|t-1}-\frac{\left(\sigma^2_{t|t-1}\right)^2}{4\rho^2} \\
-\rho^{-1}k_t\sigma^2_{t|t-1}+k^2_t &< -\frac{\left(\sigma^2_{t|t-1}\right)^2}{4\rho^2}.
\end{aligned}$$

Substitute for k_t:

$$\begin{aligned}
-\rho^{-1}\frac{\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1}+\sigma^2_t}\sigma^2_{t|t-1}+\left(\frac{\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1}+\sigma^2_t}\right)^2 &< -\frac{\left(\sigma^2_{t|t-1}\right)^2}{4\rho^2} \\
-\frac{\left(\sigma^2_{t|t-1}\right)^2}{\sigma^2_{t|t-1}+\sigma^2_t}+\frac{\rho^2\left(\sigma^2_{t|t-1}\right)^2}{\left(\sigma^2_{t|t-1}+\sigma^2_t\right)^2} &< -\frac{\left(\sigma^2_{t|t-1}\right)^2}{4\rho^2} \\
-\frac{1}{\sigma^2_{t|t-1}+\sigma^2_t}+\frac{\rho^2}{\left(\sigma^2_{t|t-1}+\sigma^2_t\right)^2} &< -\frac{1}{4\rho^2} \\
\frac{\rho^2-\left(\sigma^2_{t|t-1}+\sigma^2_t\right)}{\left(\sigma^2_{t|t-1}+\sigma^2_t\right)^2} &< -\frac{1}{4\rho^2} \\
\rho^2-\sigma^2_{t|t-1}-\sigma^2_t &< -\frac{\left(\sigma^2_{t|t-1}+\sigma^2_t\right)^2}{4\rho^2}
\end{aligned}$$

$$-\sigma_t^2 < -\frac{(\sigma_{t|t-1}^2 + \sigma_t^2)^2}{4\rho^2} + \sigma_{t|t-1}^2 - \rho^2$$
$$\sigma_t^2 > \frac{(\sigma_{t|t-1}^2 + \sigma_t^2)^2 - 4\rho^2\sigma_{t|t-1}^2 + 4\rho^4}{4\rho^2}.$$

In other words, complete attention ($\lambda_t = 1$) is optimal only when uncertainty σ_t^2 is above these two thresholds. If that is not the case, incomplete attention ($\lambda_t < 1$) is optimal.

3. $\mu_1 = 0, \mu_2 \neq 0$: $\lambda_t = 0$ is binding
From Complementary slackness conditions:
$\lambda_t = 0$

From FOC:

$$\begin{aligned}
\rho^{-1}k_t\sigma_{t|t-1}^2 + \mu_2 = 0 \\
\mu_2 &= -\rho^{-1}k_t\sigma_{t|t-1}^2 < 0, \text{ since } 0 \leq k_t, \rho \leq 1 \\
&\to \mu_2 < 0 \\
&\to \text{ does not satisfy condition that } \mu_2 \geq 0, \\
&\text{so the solution is not feasible.}
\end{aligned}$$

4. $\mu_1 \neq 0, \mu_2 \neq 0$: both $(1 - \lambda_t) = 0$ and $\lambda_t = 0$ are binding
From Complementary slackness conditions:

$$\begin{aligned}
\lambda_t &= 1 \\
\lambda_t &= 0 \\
&\to \text{ this solution is not feasible.}
\end{aligned}$$

$\therefore$ The objective function is maximized under Case 1, where:

$$\lambda_t = \frac{\rho^{-1}\sigma_{t|t-1}^2}{2k_t},$$

unless uncertainty is above the two thresholds:

$$\sigma_t^2 \geq 2\rho^2 - \sigma_{t|t-1}^2$$
$$\sigma_t^2 > \frac{(\sigma_{t|t-1}^2 + \sigma_t^2)^2 - 4\rho^2\sigma_{t|t-1}^2 + 4\rho^4}{4\rho^2},$$

in which case complete attention ($\lambda_t = 1$) is optimal.

Checking SOC: $V''(\lambda_t) < 0$ needed to be local maximum:

$$\begin{aligned}
\frac{\partial^2 V}{\partial \lambda_t^2} &= \frac{\partial}{\partial \lambda_t}[\rho^{-1} k_t \sigma^2_{t|t-1} - 2\lambda_t k_t^2] \\
&= -2k_t^2 < 0, \text{ since } 0 \leq k_t \leq 1 \\
&\rightarrow \lambda_t = \frac{\rho^{-1}\sigma^2_{t|t-1}}{2k_t} \text{ is local maximum.}
\end{aligned}$$

$V(\lambda_t)$ is concave, so the local maximum is a global maximum.

D.2 Markov-switching Model Solution

Agents observe a common costless signal:

$$x_t = \mu_{\sigma_t} + \phi x_{t-1} + \epsilon_t, \tag{D.1}$$

with a state-dependent mean μ_{σ_t} and error term $\epsilon_t \sim i.i.d.\ \mathcal{N}(0,\ \sigma^2_{\sigma_t})$, about the hidden state of economic uncertainty σ_t^2 at time t. The hidden state σ_t^2 is assumed to follow an irreducible aperiodic two-state Markov chain, where by definition $P\{\sigma_t^2|\sigma_{t-1}^2, \sigma_{t-2}^2, ..., \sigma_1^2\} = P\{\sigma_t^2|\sigma_{t-1}^2\}$. The transition of the states is a stochastic process; however, the dynamics of the switching process are known and driven by a matrix of transitional probabilities:

$$P = \begin{bmatrix} p_{LL} & p_{LH} \\ p_{HL} & p_{HH} \end{bmatrix}, \tag{D.2}$$

where element in row i, column j (p_{ij}) is the probability of switching from state j to state i, $0 \leq p_{i,j} \leq 1$, and the p_{ij} sum to 1 in each column.

The Markov-switching model is estimated with maximum likelihood. The log likelihood of the model in equation (D.1), conditional on the state being known, is:

$$\ln L = \sum_{t=1}^{T} \ln \left[\frac{1}{\sqrt{2\pi\sigma^2_{\sigma_t}}} \exp\left(- \frac{x_t - \mu_{\sigma_t} - \phi x_{t-1}}{2\sigma^2_{\sigma_t}} \right) \right]. \tag{D.3}$$

Yet, since the states of uncertainty are unknown, the notation of the likelihood function becomes $f(x_t|\sigma_t^2 = j;\theta)$ for state j and conditional on a set of parameters $\theta = (\mu_L,\ \mu_H,\ \sigma_L^2,\ \sigma_H^2,\ p_{LL},\ p_{HH})$. The full log likelihood function of the model is given by:

$$\ln L = \sum_{t=1}^{T} \ln \sum_{j\in\{L,H\}} \Big(f(x_t|\sigma_t^2 = j;\theta) Pr(\sigma_t^2 = j) \Big), \tag{D.4}$$

which is a weighted sum of the likelihood in each state j with the weights equal to the state's probabilities. Since these probabilities are not observed, equation (D.4) cannot be applied directly; however, we can make inference about the probabilities based on the available information.

The probabilities of each state $Pr(\sigma_t^2 = j)$ are estimated based on the new incoming information using Hamilton's iterative algorithm, as follows:

1. Set a guess for the initial probabilities at time $t = 0$ for each state $Pr(\sigma_0^2 = j)$ for $j \in \{LU, HU\}$, conditional on the information set at time $t = 0$, I_0, i.e. $Pr(\sigma_0^2 = j) = 0.5$

or the steady-state unconditional probabilities of σ_t^2:

$$Pr(\sigma_0^2 = LU|I_0) = \frac{1 - p_{LL}}{2 - p_{LL} - p_{HH}}$$
$$Pr(\sigma_0^2 = HU|I_0) = \frac{1 - p_{HH}}{2 - p_{LL} - p_{HH}}.$$

2. For a following period t, calculate the probabilities of each state, given information up to time $t-1$, I_{t-1}, where p_{ji} are the transitional probabilities from the Markov chain in equation (D.2):

$$Pr(\sigma_t^2 = j|I_{t-1}) = \sum_{i \in \{L,H\}} p_{ji}\big(Pr(\sigma_{t-1}^2 = i|I_{t-1})\big). \tag{D.5}$$

3. Update the probability of each state with the new incoming information at time t using the set of parameters θ to calculate the log likelihood function for each state $f(x_t|\sigma_t^2 = j; I_{t-1})$ for time t and update the filtered or conditional probability of each state $\pi_{j|t}$, given the new information, as follows:

$$\pi_{j|t} = Pr(\sigma_t^2 = j|I_t) = \frac{f(x_t|\sigma_t^2 = j; I_{t-1})Pr(\sigma_t^2 = j|I_{t-1})}{\sum_{j \in \{L,H\}} f(x_t|\sigma_t^2 = j; I_{t-1})Pr(\sigma_t^2 = j|I_{t-1})}. \tag{D.6}$$

4. Set $t = t+1$ and repeat steps 2-3 until $t = T$ to obtain a set of filtered probabilities $\pi_{j|t}$ of each state for every period in the sample.

Using the set of filtered probabilities $\pi_{j|t}$, the log likelihood of the model can be calculated by maximum likelihood as a function of the parameters of the model that maximize:

$$\ln L = \sum_{t=1}^{T} \ln \sum_{j \in \{L,H\}} \Big(f(x_t|\sigma_t^2 = j; \theta)\pi_{j|t} \Big). \tag{D.7}$$

D.3 Kalman Filter Solution

D.3.1 Individual Forecast

The observation equation is:

$$z_{it} = y_t + \eta_{it}, \tag{D.8}$$

where $\eta_{it} \sim i.i.d.\ \mathcal{N}(0,\, \sigma_t^2)$. The variance of the error η_{it}, σ_t^2, is stochastic and thus, the model is characterized by stochastic volatility. Furthermore, σ_t^2 is assumed to follow an irreducible aperiodic two-state Markov chain, where $\sigma_t^2 \in \{\sigma_{LU}^2, \sigma_{HU}^2\}$, the two states of economic uncertainty from Section 4.2. The state or transition equation for the macroeconomic state variable takes an AR(1) form:

$$y_t = \rho y_{t-1} + v_t, \text{ where } v_t \sim i.i.d.\ \mathcal{N}(0,\, \sigma_v^2) \text{ and } 0 < \rho < 1. \tag{D.9}$$

Equations (D.9) and (D.8) together describe the information structure in state space form (ssf). Each agent i generates her nowcast for the state variable and its mean squared error (MSE) at time t for the period t, respectively, as:

$$\begin{aligned} \hat{y}_{t|t} &= \mathbb{E}[y_t | z_{i1}, ..., z_{it}] \\ \sigma_{t|t}^2 &= \mathbb{E}[(y_t - \hat{y}_{t|t})^2] \end{aligned}$$

Each agent forecasts the state variable at time t for the period $t+1$. The forecast and its MSE are thus:

$$\begin{aligned} \hat{y}_{t+1|t} &= \mathbb{E}[y_{t+1} | z_{i1}, ..., z_{it}] \\ \sigma_{t+1|t}^2 &= \mathbb{E}[(y_{t+1} - \hat{y}_{t+1|t})^2]. \end{aligned}$$

The recursion is initialized with $\hat{y}_{1|0} = \mathbb{E}(y_1|z_{i0}) = 0$ with associated MSE $\sigma_{1|0}^2 = Var(y_1|z_{i0}) = \frac{\sigma_v^2}{1-\rho^2}$, which are just the unconditional mean and variance of y_1 using information from time $t = 0$ (Hamilton, 1994, Chapter 13).

Agents need to estimate the error in the incoming noisy signal. Using the law of iterated expectations and equation (D.8), at time $t-1$ attentive agents forecast the value of z_{it}:

$$\begin{aligned} \hat{z}_{it|t-1} &= \mathbb{E}[z_{it} | z_{i1}, ..., z_{it-1}] \\ &= \mathbb{E}[y_t | z_{i1}, ..., z_{it-1}] + \mathbb{E}[\eta_{it} | z_{i1}, ..., z_{it-1}] \\ &= \hat{y}_{t|t-1}. \end{aligned} \tag{D.10}$$

For each agent i, the Kalman filter assigns different weights to the incoming noisy information and the previous estimate depending on the precision of the new observed data and the perceived accuracy of the existing estimate. Thus, in order to be able to estimate the forecasts for the following periods, agents need to calculate the error in the existing estimate, as well as the error in the incoming noisy information signal. The Kalman gain k_t is a measure of how much the forecaster can trust her information signal. The more credible the signal, the more weight the forecaster will optimally put on it in updating her expectation. The Kalman gain k_t is defined as:

$$k_t = \frac{\rho \sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_t}. \tag{D.11}$$

The Kalman gain is increasing in the MSE of the existing estimate ($\sigma^2_{t|t-1}$) and state process persistence (ρ) and decreasing in the variance of the error in the observation of the signal (σ^2_t). Thus, the less accurate the previous estimate and the more persistent the state process, the more weight is optimally placed on the new information signal. In contrast, the more uncertain the signal (higher σ^2_t), the less weight is optimally placed on it in updating the prediction.

The optimal nowcast of the attentive agents, who update their inference about the value of y_t based on the new observed idiosyncratic signal z_{it}, using the formula for updating a linear projection, is:

$$\begin{aligned}
\hat{y}_{t|t} &= \mathbb{E}[y_t | z_{i1}, ..., z_{it}] \\
&= \hat{y}_{t|t-1} + \mathbb{E}[(y_t - \hat{y}_{t|t-1})(z_{it} - \hat{z}_{it|t-1})] \times \mathbb{E}[(z_{it} - \hat{z}_{it|t-1})^2]^{-1} \times (z_{it} - \hat{z}_{it|t-1}) \\
&= \hat{y}_{t|t-1} + \mathbb{E}[(y_t - \hat{y}_{t|t-1})(y_t - \hat{y}_{t|t-1} + \eta_{it})] \times \mathbb{E}[(y_t - \hat{y}_{t|t-1} + \eta_{it})^2]^{-1} \\
&\qquad \times (z_{it} - \hat{y}_{t|t-1}) \\
&= \hat{y}_{t|t-1} + \sigma^2_{t|t-1} \times \{\mathbb{E}[(y_t - \hat{y}_{t|t-1})^2] + \mathbb{E}[(\eta_{it})^2]\}^{-1} \times (z_{it} - \hat{y}_{t|t-1}) \\
&= \hat{y}_{t|t-1} + \sigma^2_{t|t-1} \times [\sigma^2_{t|t-1} + \sigma^2_t]^{-1} \times (z_{it} - \hat{y}_{t|t-1}) \\
&= \hat{y}_{t|t-1} + \rho^{-1} k_t \times (z_{it} - \hat{y}_{t|t-1}) \\
&= \rho^{-1} k_t z_{it} + (1 - \rho^{-1} k_t) \hat{y}_{t|t-1},
\end{aligned} \tag{D.12}$$

since $\mathbb{E}[\eta_{it}(y_t - \hat{y}_{t|t-1})] = 0$. The MSE of the optimal nowcast is:

$$
\begin{aligned}
\sigma^2_{t|t} &= \mathbb{E}[(y_t - \hat{y}_{t|t})^2] \\
&= \mathbb{E}[(y_t - \rho^{-1}k_t z_{it} - (1-\rho^{-1}k_t)\hat{y}_{t|t-1})^2] \\
&= \mathbb{E}[(\rho^{-1}k_t(y_t - z_{it}) + (1-\rho^{-1}k_t)(y_t - \hat{y}_{t|t-1}))^2] \\
&= \mathbb{E}[(-\rho^{-1}k_t\eta_{it} + (1-\rho^{-1}k_t)(y_t - \hat{y}_{t|t-1}))^2] \\
&= \rho^{-2}k_t^2\sigma_t^2 + (1-\rho^{-1}k_t)^2\sigma^2_{t|t-1} \\
&= \sigma^2_{t|t-1} - 2\rho^{-1}k_t\sigma^2_{t|t-1} + \rho^{-2}k_t^2(\sigma^2_{t|t-1} + \sigma_t^2) \\
&= \sigma^2_{t|t-1} - 2\rho^{-1}k_t\sigma^2_{t|t-1} + \rho^{-2}\frac{\rho^2\sigma^4_{t|t-1}}{(\sigma^2_{t|t-1} + \sigma_t^2)^2}(\sigma^2_{t|t-1} + \sigma_t^2) \\
&= \sigma^2_{t|t-1} - 2\rho^{-1}k_t\sigma^2_{t|t-1} + \frac{\sigma^4_{t|t-1}}{\sigma^2_{t|t-1} + \sigma_t^2} \\
&= \sigma^2_{t|t-1} - 2\rho^{-1}k_t\sigma^2_{t|t-1} + \rho^{-1}k_t\sigma^2_{t|t-1} \\
&= (1-\rho^{-1}k_t)\sigma^2_{t|t-1}. \qquad \text{(D.13)}
\end{aligned}
$$

The optimal nowcast from equation (D.12) $\hat{y}_{t|t}$ is a weighted sum of the new information z_{it} and the old information $\hat{y}_{t|t-1}$, weighted by the respective weights $\rho^{-1}k_t$ and $(1-\rho^{-1}k_t)$. This is the equivalent of equation (8) in Coibion and Gorodnichenko (2015) for the homoscedastic case. However, here the Kalman gain is different due to heteroscedasticity.

Finally, the attentive agents produce their forecasts for $\hat{y}_{t+1|t}$:

$$
\begin{aligned}
\hat{y}_{t+1|t} &= \mathbb{E}[y_{t+1}|z_{i1}, ..., z_{it}] \\
&= \rho\,\mathbb{E}[y_t|z_{i1}, ..., z_{it}] + \mathbb{E}[v_{t+1}|z_{i1}, ..., z_{it}] \\
&= \rho\hat{y}_{t|t} + 0 \\
&= \rho[\rho^{-1}k_t z_{it} + (1-\rho^{-1}k_t)\hat{y}_{t|t-1}] \\
&= k_t z_{it} + (\rho - k_t)\hat{y}_{t|t-1} \qquad \text{(D.14)}
\end{aligned}
$$

with MSE of the forecast:

$$
\begin{aligned}
\sigma^2_{t+1|t} &= \mathbb{E}[(y_{t+1} - \hat{y}_{t+1|t})^2] \\
&= \mathbb{E}[(\rho y_t + v_{t+1} - \rho \hat{y}_{t|t})^2] \\
&= \rho^2\, \mathbb{E}[(y_t - \hat{y}_{t|t})^2] + \mathbb{E}[v^2_{t+1}] \\
&= \rho^2 \sigma^2_{t|t} + \sigma^2_v \\
&= \rho^2[(1 - \rho^{-1} k_t)\sigma^2_{t|t-1}] + \sigma^2_v \\
&= \rho(\rho - k_t)\sigma^2_{t|t-1} + \sigma^2_v. \qquad \text{(D.15)}
\end{aligned}
$$

D.3.2 Mean Forecast

By the law of large numbers, assuming the population of forecasters is large, a state-dependent fraction λ_t of the population is attentive each period and a proportion $(1 - \lambda_t)$ is inattentive. In other words, the mean forecast of the entire population is the sum of the mean forecasts within each group of forecasters j, who all updated their information sets j periods ago, weighted by the respective proportion of each group in the total population.

The mean forecast of the entire population is denoted as the expectation across individuals i within the same group j and *across* all groups j: $\mathbb{E}_j[\mathbb{E}_i(\hat{y}_{t+1|t-j})]$. The mean forecast across all agents i *within* the same group j is $\mathbb{E}_i[\hat{y}_{t+1|t}] = \bar{y}_{t+1|t}$, where $\hat{y}_{t+1|t}$ are the individual forecasts. Unlike the hybrid sticky-noisy information model, where λ is constant, here λ_t is state-dependent. The mean forecast of the entire population of forecasters is thus:

$$
\begin{aligned}
\mathbb{E}_j[\bar{y}_{t+1|t-j}] = \lambda_t \bar{y}_{t+1|t} & \\
+ \lambda_{t-1}(1 - \lambda_t)\bar{y}_{t+1|t-1} & \\
+ \lambda_{t-2}(1 - \lambda_{t-1})(1 - \lambda_t)\bar{y}_{t+1|t-2} & \\
+ \ldots & \\
+ \lambda_1(1 - \lambda_2)\ldots(1 - \lambda_{t-1})(1 - \lambda_t)\bar{y}_{t+1|1}. & \qquad \text{(D.16)}
\end{aligned}
$$

We can substitute for the mean forecast of the agents in each group. Taking expectation across agents i in group $j = 0$ w.l.o.g. using equation (D.14), the mean forecast of this group (all the forecasters who update their forecast for time $t+1$ at time t) becomes:

$$
\begin{aligned}
\mathbb{E}_j[\bar{y}_{t+1|t-j}|j=0] = \mathbb{E}_i[\hat{y}_{t+1|t}] &= \bar{y}^{attentive}_{t+1|t} \\
&= k_t\,\mathbb{E}_i[z_{it}] + (\rho - k_t)\,\mathbb{E}_i[\hat{y}_{t|t-1}] \\
&= k_t\,\mathbb{E}_i[y_t + \eta_{it}] + (\rho - k_t)\bar{y}_{t|t-1} \\
&= k_t y_t + (\rho - k_t)\bar{y}_{t|t-1}.
\end{aligned} \tag{D.17}
$$

Note that equation (D.17) yields the mean forecast of the noisy information model, according to which all agents are attentive, updating their forecasts every period.

Combining equations (D.16) and (D.17), the mean forecast of all agents, attentive and inattentive, at time t for time $t+1$, where the mean prediction of the inattentive forecasters is $\bar{y}^{inattentive}_{t+1|t} = \bar{y}_{t+1|t-1} = \rho\bar{y}_{t|t-1}$, is:

$$
\begin{aligned}
\mathbb{E}_j[\bar{y}_{t+1|t-j}|j=0] = \bar{y}^{all}_{t+1|t} &= \lambda_t\bar{y}^{attentive}_{t+1|t} + (1-\lambda_t)\bar{y}^{inattentive}_{t+1|t} \\
&= \lambda_t[k_t y_t + (\rho - k_t)\bar{y}_{t|t-1}] + (1-\lambda_t)\rho\bar{y}_{t|t-1} \\
&= \lambda_t k_t y_t + (\rho - \lambda_t k_t)\bar{y}_{t|t-1}.
\end{aligned} \tag{D.18}
$$

D.4 State-dependent Attention λ_t and Kalman Gain k_t

In Section 4.2, the individual agent calculates her estimated filtered probabilities of the two states of economic uncertainty. However, it is not realistic to assume that all agents will estimate the same filtered probabilities, even when exposed to the same common costless signal, due to the different models they employ to process the information. Thus, each agent calculates filtered probabilities $\pi_{j|t}$ polluted with error: $\hat{\pi}_{j|t} = \pi_{j|t} + \epsilon_{\pi,i}$. Focusing on the probability of a high-uncertainty (HU) state, the mean estimated probability becomes:

$$\bar{\pi}_{HU|t} = \mathbb{E}_i[\pi_{HU|t} + \epsilon_{\pi,i}] = \pi_{HU|t}. \tag{D.19}$$

Moreover, from equation (D.6), $\pi_{LU|t} + \pi_{HU|t} = 1$. Thus,

$$\pi_{LU|t} = 1 - \pi_{HU|t}. \tag{D.20}$$

The expected value of economic uncertainty σ_t^2 at time t is:

$$\mathbb{E}[\sigma_t^2 | x_1, ..., x_t] = \sigma_{t|t}^2 = \sigma_{LU}^2 \times \pi_{LU|t} + \sigma_{HU}^2 \times \pi_{HU|t}. \tag{D.21}$$

Finally, in order to be able to express attention λ_t and the Kalman gain k_t in terms of the filtered probabilities of HU state, I follow Bloom (2009) and set the following relationship for σ_{LU}^2 and σ_{HU}^2:[18]

$$\begin{aligned} \sigma_{HU} &= 2 \times \sigma_{LU} \\ \sigma_{HU}^2 &= 4 \times \sigma_{LU}^2. \end{aligned} \tag{D.22}$$

Substituting equations (D.20) and (D.22) in equation (D.21), the expected value of economic uncertainty $\sigma_{t|t}^2$ becomes:

$$\begin{aligned} \sigma_{t|t}^2 &= \sigma_{LU}^2(1 - \pi_{HU|t}) + 4\sigma_{LU}^2 \pi_{HU|t} \\ &= \sigma_{LU}^2 - \sigma_{LU}^2 \pi_{HU|t} + 4\sigma_{LU}^2 \pi_{HU|t} \\ &= \sigma_{LU}^2(1 + 3\pi_{HU|t}). \end{aligned} \tag{D.23}$$

[18] As in Bloom (2009), I confirm the results for $\sigma_{HU} = 2 \times \sigma_{LU}$ are also valid for $\sigma_{HU} = 1.5 \times \sigma_{LU}$ and $\sigma_{HU} = 3 \times \sigma_{LU}$.

D.4.1 State-dependent Attention λ_t

. Substituting for the Kalman gain from equation (D.11) in the expression for λ_t that maximizes the objective function in the agent's optimization problem from equation (4.5), yields:

$$\lambda_t = \frac{\rho^{-1}\sigma^2_{t|t-1}}{\frac{2\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1}+\sigma^2_t}} = \frac{\rho^{-1}\sigma^2_{t|t-1} \times (\sigma^2_{t|t-1}+\sigma^2_t)}{2\rho\sigma^2_{t|t-1}} = \frac{\sigma^2_{t|t-1}+\sigma^2_t}{2\rho^2}. \tag{D.24}$$

Applying the result for $\sigma^2_{t|t}$ from equation (D.23) to the result for λ_t above, expected attention λ_t at time t becomes:

$$\begin{aligned}\mathbb{E}[\lambda_t|x_1,...,x_t] = \lambda_{t|t} &= \mathbb{E}_t\left[\frac{\sigma^2_{t|t-1}+\sigma^2_t}{2\rho^2}\right] \\ &= \frac{\sigma^2_{t|t-1}+\sigma^2_{t|t}}{2\rho^2} \\ &= \frac{\sigma^2_{t|t-1}+\sigma^2_{LU}(1+3\pi_{HU|t})}{2\rho^2}\end{aligned} \tag{D.25}$$

To determine how the previous MSE $\sigma^2_{t|t-1}$, the LU regime uncertainty σ^2_{LU} (therefore, also the HU regime uncertainty σ^2_{HU}), the conditional probability of HU state $\pi_{HU|t}$, and the persistence of the state process ρ influence expected attention $\lambda_{t|t}$, I examine the partial derivatives of $\lambda_{t|t}$ with respect to these four variables.

$$\frac{\partial\lambda_{t|t}}{\partial\sigma^2_{t|t-1}} = \frac{1}{2\rho^2} > 0 \rightarrow \text{ as previous MSE } \sigma^2_{t|t-1}\uparrow \Longrightarrow \text{ expected } \lambda_{t|t}\uparrow$$

$$\frac{\partial\lambda_{t|t}}{\partial\sigma^2_{LU}} = \frac{1+3\pi_{HU|t}}{2\rho^2} > 0 \rightarrow \text{ as LU } \sigma^2_{LU}\uparrow\rightarrow \text{ HU } \sigma^2_{HU}\uparrow \Longrightarrow \text{ expected } \lambda_{t|t}\uparrow$$

$$\frac{\partial\lambda_{t|t}}{\partial\pi_{HU|t}} = \frac{3\sigma^2_{LU}}{2\rho^2} > 0 \rightarrow \text{ as probability of HU } \pi_{HU|t}\uparrow \Longrightarrow \text{ expected } \lambda_{t|t}\uparrow$$

$$\frac{\partial\lambda_{t|t}}{\partial\rho} = -\frac{\sigma^2_{t|t-1}+\sigma^2_{LU}(1+3\pi_{HU|t})}{\rho^3} < 0 \rightarrow \text{ as persistence } \rho\uparrow \Longrightarrow \text{ expected } \lambda_{t|t}\downarrow$$

Hence, attention $\lambda_{t|t}$ is an increasing function of the previous MSE, so economic agents are expected to pay more attention when their previous forecasts have been more inaccurate. Moreover, attention $\lambda_{t|t}$ is an increasing function of LU-regime uncertainty and by construction, also HU-regime uncertainty, confirming that agents are expected to become more attentive during periods of greater volatility. Attention is also an increasing function of the conditional probability of a HU state, so that $\lambda_{t|t} \propto \pi_{HU|t}$. In other words, the greater the probability of a perceived HU state, the greater the probability of updating one's forecast. Finally, attention $\lambda_{t|t}$ is a decreasing function of the persistence of the state process ρ, suggesting that more persistent processes cause agents to optimally become less attentive. Overall, all these relationships are of the expected signs.

D.4.2 State-dependent Kalman gain $\mathbf{k_t}$

. Substituting the result for $\sigma^2_{t|t}$ from equation (D.23) in the Kalman gain equation (D.11) yields:

$$\begin{aligned}\mathbb{E}[k_t|x_1, ..., x_t] = k_{t|t} &= \mathbb{E}_t\left[\frac{\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_t}\right] \\ &= \frac{\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_{t|t}} \\ &= \frac{\rho\sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_{LU}(1 + 3\pi_{HU|t})}\end{aligned} \tag{D.26}$$

To determine how the previous MSE $\sigma^2_{t|t-1}$, the LU regime uncertainty σ^2_{LU} (therefore, also the HU regime uncertainty σ^2_{HU}), the conditional probability of HU state $\pi_{HU|t}$, and the persistence of the state process ρ influence the expected Kalman gain $k_{t|t}$, I examine the partial derivatives of $k_{t|t}$ with respect to these four variables:

$$\frac{\partial k_{t|t}}{\partial \sigma^2_{t|t-1}} = \frac{\rho\sigma^2_{LU}(1+3\pi_{HU|t})}{\left(\sigma^2_{t|t-1}+\sigma^2_{LU}(1+3\pi_{HU|t})\right)^2} > 0 \rightarrow \text{ as previous MSE } \sigma^2_{t|t-1}\uparrow \implies \text{ expected } k_{t|t}\uparrow$$

$$\frac{\partial k_{t|t}}{\partial \sigma^2_{LU}} = -\frac{\rho\sigma^2_{t|t-1}(1+3\pi_{HU|t})}{\left(\sigma^2_{t|t-1}+\sigma^2_{LU}(1+3\pi_{HU|t})\right)^2} < 0 \rightarrow \text{ as LU } \sigma^2_{LU}\uparrow\rightarrow \text{HU } \sigma^2_{HU}\uparrow \implies \text{ expected } k_{t|t}\downarrow$$

$$\frac{\partial k_{t|t}}{\partial \pi_{HU|t}} = -\frac{3\rho\sigma^2_{t|t-1}}{\left(\sigma^2_{t|t-1}+\sigma^2_{LU}(1+3\pi_{HU|t})\right)^2} < 0 \rightarrow \text{ as probability of HU } \pi_{HU|t}\uparrow \implies \text{ expected } k_{t|t}\downarrow$$

$$\frac{\partial k_{t|t}}{\partial \rho} = \frac{\sigma^2_{t|t-1}}{\sigma^2_{t|t-1}+\sigma^2_{LU}(1+3\pi_{HU|t})} > 0 \rightarrow \text{ as persistence } \rho\uparrow \implies \text{ expected } k_{t|t}\uparrow$$

Therefore, the expected weight agents put on new information, Kalman gain $k_{t|t}$, is an increasing function of the previous MSE, so economic agents are expected to trust the new information more when their previous forecasts have been more inaccurate. Moreover, the Kalman gain $k_{t|t}$ is a decreasing function of LU-regime uncertainty and by construction, also HU-regime uncertainty, since agents optimally place less weight on new incoming information when it is more noisy. The Kalman gain is also a decreasing function of the conditional probability of a HU state, so that $k_{t|t} \not\propto \pi_{HU|t}$. In other words, the greater the probability of a perceived HU state, the lesser the weight agents optimally put on the new information when updating their forecasts. Finally, the Kalman gain $k_{t|t}$ is an increasing function of the persistence of the state process ρ, suggesting that more persistent processes cause agents to optimally weigh new information more heavily. Overall, all these relationships are also of the expected signs.

D.5 Mean Forecasts of Alternative Models

D.5.1 Sticky Information Model

$$
\begin{aligned}
\mathbb{E}_j[\bar{y}_{t+1|t-j}] &= \lambda \bar{y}_{t+1|t} \\
&+ \lambda(1-\lambda)\bar{y}_{t+1|t-1} \\
&+ \lambda(1-\lambda)(1-\lambda)\bar{y}_{t+1|t-2} \\
&+ \ldots \\
&+ \lambda(1-\lambda)\ldots(1-\lambda)(1-\lambda)\bar{y}_{t+1|1} \\
&= \lambda \sum_{j=0}^{\infty} (1-\lambda)^j \bar{y}_{t+1|t-j}. \qquad \text{(D.27)}
\end{aligned}
$$

D.5.2 Noisy Information Model

The individual signal that each agent i gleans is described in the observation equation:

$$z_{it} = y_t + \eta_{it}, \qquad \text{(D.28)}$$

where $\eta_{it} \sim i.i.d.\ \mathcal{N}(0,\, \sigma_t^2)$.

Each agent i updates her forecast according to the state equation for the macroeconomic state variable, which takes an AR(1) form:

$$y_t = \rho y_{t-1} + v_t, \text{ where } v_t \sim i.i.d.\ \mathcal{N}(0,\, \sigma_v^2) \text{ and } 0 < \rho < 1. \qquad \text{(D.29)}$$

As derived in Section 4.3, the optimal forecast of each agent i is:

$$\hat{y}_{t+1|t} = k z_{it} + (\rho - k)\hat{y}_{t|t-1}, \qquad \text{(D.30)}$$

where k is the Kalman gain. Averaging across individual agents yields the mean forecast:

$$
\begin{aligned}
\mathbb{E}_i[\hat{y}_{t+1|t}] &= \bar{y}_{t+1|t} \\
&= k\, \mathbb{E}_i[z_{it}] + (\rho - k)\, \mathbb{E}_i[\hat{y}_{t|t-1}] \\
&= k\, \mathbb{E}_i[y_t + \eta_{it}] + (\rho - k)\bar{y}_{t|t-1} \\
&= k y_t + (\rho - k)\bar{y}_{t|t-1}. \qquad \text{(D.31)}
\end{aligned}
$$

D.5.3 Hybrid Sticky-Noisy Information Model

$$\begin{aligned}\mathbb{E}_j[\bar{y}_{t+1|t-j}] &= \lambda \sum_{j=0}^{\infty} (1-\lambda)^j \, \mathbb{E}_i[\hat{y}_{t+1|t-j}] \\ &= \lambda \sum_{j=0}^{\infty} (1-\lambda)^j [k y_t + (\rho - k)\bar{y}_{t|t-j-1}] \end{aligned} \tag{D.32}$$

D.5.4 State-dependent Sticky-Noisy Information Model

$$\begin{aligned}\mathbb{E}_j[\bar{y}_{t+1|t-j}] &= \lambda_t \bar{y}_{t+1|t} + \sum_{j=1}^{\infty} \lambda_{t-j} \left[\prod_{k=0}^{j-1} (1-\lambda_{t-k}) \right] \bar{y}_{t+1|t-j} \\ &= \lambda_t [k_t y_t + (\rho - k_t)\bar{y}_{t|t-1}] + \sum_{j=1}^{\infty} \lambda_{t-j} \left[\prod_{k=0}^{j-1} (1-\lambda_{t-k}) \right] \rho^j [k_{t-j} y_{t-j} + (\rho - k_{t-j})\bar{y}_{t-j|t-j-1}]. \end{aligned} \tag{D.33}$$