H.O. Stekler Research Program on Forecasting

WORKING PAPER SERIES

Expectation Formation and the Persistence of Shocks

Constantin Bürgi
St. Mary's College of Maryland
crburgi@smcm.edu

Working Paper No. 2020-005

First Version: August, 2020

This Version: September, 2020

H. O. STEKLER RESEARCH PROGRAM ON FORECASTING
Department of Economics
Columbian College of Arts & Sciences
The George Washington University
Washington, DC 20052
https://www2.gwu.edu/~forcpgm

Expectation Formation and the Persistence of Shocks*

Constantin Bürgi†

September 17, 2020

Abstract

Persistence of economic shocks is commonly treated as deviations from rational expectations attributed to frictions like information rigidity or noisy information. This paper shows that there is persistence even without these information frictions. In the presence of uncertainty about the future, optimal forecasts place a positive weight on past predictions about the same event. The overall weight on the past prediction varies markedly over time and has an inverse relationship with the magnitude of shocks as the larger revisions after large shocks reduce the weight. Empirical estimates show that agents put a significant weight on previous prediction of inflation and output and a substantial part of the weight and hence persistence cannot be attributed to information frictions.

JEL: C53, D80, E37, E66

Keywords: Uncertainty, Information Rigidity, Information, Rational Expectations, Full Information

*I would like to thank Heiner Mikosch, Julio Ortiz, Jane Ryngaert, Xuguang (Simon) Sheng, Tara Sinclair and all participants in the H.O. Stekler Research Seminar for their valuable comments and support.

† *crburgi@smcm.edu, St. Mary's College of Maryland, 47645 College Drive, St. Mary's City, MD, 20686*

1 Introduction

There is ample evidence that economic shocks are persistent.[1] The literature on information frictions argues that under rational expectations with full information about the past and present (FIRE), agents immediately and fully adjust their expectations after shocks. Persistence is then interpreted as a deviation from rational expectations or full information, which has led to a wealth of models on the cause of these deviations.[2] This paper shows that uncertainty about the future leads to persistence even under FIRE. In the presence of uncertainty, optimal predictions place a positive weight on previous predictions. This implies that in addition to the well established persistence due to deviations from FIRE, there is also persistence without deviations from FIRE due to the optimal forecasting behavior in the presence of uncertainty about the future.[3] This result has important implication for the dynamic stochastic general equilibrium (DSGE) models. Specifically, it makes classical real business cycle models much more similar to new Keynesian models as the rational expectation forecasting behavior of agents can lead to a persistence of transitory shocks in both.

This paper introduces an estimation approach to measure the overall weight on the previous prediction which includes three components: the first component is due to optimal forecasting behavior under FIRE that causes more persistence of economic shocks, the second one is due to information frictions and also causes more persistence of economic shocks and the third component is due to the persistence of the underlying variable which does not cause more persistence of economic shocks. Utilizing the Survey of Professional Forecasters

[1] E.g. Brunner et al. (1980), Diebold and Rudebusch (1989), Clarida et al. (1999), Galı and Gertler (1999), Christiano et al. (2005)

[2] E.g. see Calvo (1983) or Mankiw and Reis (2002) for rigidity and Lucas Jr (1972), Kydland and Prescott (1982), Woodford (2001), Sims (2003) or Coibion and Gorodnichenko (2015) for noisy information type models.

[3] The empirical literature on deviations includes for example Coibion and Gorodnichenko (2015), Dräger and Lamla (2012), Andrade and Le Bihan (2013), An et al. (2017), Khan and Zhu (2006), Milani (2007), Döpke et al. (2008), Klenow and Malin (2010) or Nakamura and Steinsson (2013)

(SPF) from the Federal Reserve Bank of Philadelphia and the Federal Reserve's Greenbook forecasts, the weight on the past prediction is significant for GDP and inflation for both data sets.[4]

Having newly estimated the overall weight immediately raises the question about decomposing the overall weight. Under rational expectations, agents want to minimize the variance of the prediction error and they weight the old prediction and the prediction implied by new information according the inverse of their respective prediction error variances (e.g. see Bates and Granger (1969)). While the ex post variances are typically unknown to the agents when making their prediction, they can be readily calculated ex post and used to construct the ex post optimal weights. One can then compare the optimal weights to the actual weights and the share of the estimated weight explained by the ex post optimal weight then measures the share of the persistence which exists even without information frictions. Aside from the CPI predictions in the SPF, the information frictions appear to be minimal and in most cases small, relative to the overall weight on the past prediction.[5] This implies that substantial persistence also exists under FIRE and is due to optimal forecasting behavior under uncertainty about the future.

While frictions do not appear to be the main driver of the weight in most cases for the overall sample, they might still be important for certain sub-periods. The weights depend on the variances of the expected prediction errors, which in turn are known to vary over time (e.g. Zarnowitz and Lambros (1987), Lahiri and Sheng (2010) or Jurado et al. (2015)). Using rolling window regressions, it is shown that this variation over time leads to a distinct pattern in the overall weights as they tend to become smaller during recessions and larger outside recessions for GDP predictions. The typical swing in the weights is large at around 0.4 for GDP. Both the estimated weights and the ex post optimal weights follow this cyclical pattern closely, but there are also clear differences between the two. However, these differences

[4]The traditional Greenbook has been merged into the Tealbook.

[5]For the CPI predictions in the SPF, around 60% of the overall weight is explained by information frictions.

cannot automatically be attributed to information frictions. Agents do not have access to the ex post optimal weights when they make predictions and need to estimate them. The estimated weights in turn are typically not observed and so it is not possible to distinguish, whether the observed differences between the estimated and ex post optimal weights are due to information frictions or due to agents making prediction errors. The former would imply deviations from FIRE, while the latter is consistent with optimal behavior under uncertainty.

Given the cyclical patterns in the weights, one might wonder, how agents are able to accurately estimate the large shifts in the ex post optimal weights. Specifically, agents could actively change their forecasting behavior and update their predictions more often as suggested in Baker et al. (2018) for example, or putting a larger weight on new information leading to larger revisions. Alternatively, one could hypothesize that no active changes are necessary and the larger shocks automatically lead to the lower weight on the past prediction. In order to test this, it is checked whether the cycles are also apparent in a simple auto-regressive model and whether forecasters revise their predictions more often in periods where the weight on the past prediction is smaller. Based on the Wall Street Journal Survey, only around 10% of forecasters leave their prediction unchanged from month to month and while there might be a very weak cyclical pattern, the month to month variations make it difficult to identify. At the same time, the cyclical pattern is clearly visible in the simple auto-regressive model. Both findings together suggest that while agents might adjust their prediction behavior, a cyclical pattern arises even without changes to the forecasting behavior.

The remainder of the paper is structured as follows: The next section presents the econometric model followed by the data utilized in the paper. Next, the models are applied to the SPF forecasts and the forecasts in the Greenbook of the Federal Reserve for output and inflation over the entire sample. The next section looks at the time varying behavior of the weights and the last section concludes.

2 Econometric Model

Assume that agents have access to all past realizations of an underlying variable A_t, meaning there is full information about the past. Similar to the noisy information model, it is assumed that agents continuously update their predictions based on the information they receive. Further assume, that A_t is auto-correlated of the form

$$A_t = \psi + \phi A_{t-1} + \eta_t \tag{1}$$

where η_t is independent from η_{t-k} $\forall k > 0$ and $\phi \in [0, 1]$. Without any signal, the best possible prediction in period $t-1$ for period t is

$$F_{t,t-1} = \psi + \phi A_{t-1} \tag{2}$$

with the prediction error η_t. Now assume that agents receive a noisy signal about η_t every period $x_{t,t-h} = \eta_t + \nu_{t,t-h}$, where $\nu_{t,t-h}$ is independent across both t and h.[6] An intuitive example of how these expectations are formed is as follows: Assume a forecaster predicts inflation. Further assume that the auto-correlation implies that inflation for the next period is 3%. In addition, the forecaster has learned that there are currently tensions in the middle east which are likely to increase inflation next period. This causes the forecast to become 3.5% for example, as this information is not included in the forecast based exclusively on the auto-correlation. This is also consistent with the well known fact that survey expectations outperform auto-regressive models at least at the short horizon (e.g. see Ang et al. (2007), or survey predictions around the 2020 pandemic).

Based on the series of independent signals, agents create a prediction for η_t denoted $\hat{\eta}_{t,t-h}$ made in period $t-h$. Denote $\varepsilon_{t,t-h}$ the error of this prediction, meaning that $\hat{\eta}_{t,t-h} = \eta_t + \varepsilon_{t,t-h}$, which is based on the optimally weighted average between the past prediction with error $\varepsilon_{t,t-h-1}$ and the new signal $x_{t,t-h}$. The forecast made in period $t-1$ for period t then

[6] Note that this setup encompasses the case where the signal is about A_t instead of just η_t, as any such signal could be decomposed into the already known auto-correlation part and the new information about η_t.

becomes

$$F_{t,t-1} = \psi + \phi A_{t-1} + \eta_t + \varepsilon_{t,t-1} = A_t + \varepsilon_{t,t-1} \tag{3}$$

More generally, the prediction made in period $t-1-h$ for $h \geq 0$ is

$$F_{t,t-1-h} = \psi + \phi F_{t-1,t-1-h} + \eta_t + \varepsilon_{t,t-1-h} = A_t + \sum_{k=0}^{h} \phi^k \varepsilon_{t-k,t-1-h} = A_t + \varepsilon^*_{t,t-1-h} \tag{4}$$

Note that $F_{t-1,t-1} = A_{t-1}$ in the case of $h = 0$. Denote $\gamma \in [0,1]$ the optimal weight on the past prediction of η_t.[7] Hence

$$\hat{\eta}_{t,t-h} = \gamma \hat{\eta}_{t,t-h-1} + (1-\gamma) x_{t,t-h} \tag{5}$$

where $\hat{\eta}_{t,t-h}$ is the prediction of η_t made in period $t-h$. Subtracting the actual η_t, one obtains

$$\varepsilon_{t,t-h} = \gamma \varepsilon_{t,t-h-1} + (1-\gamma) \nu_{t,t-h} \tag{6}$$

One important property of this setup is that it generates a channel separate from the auto-correlation through which the prediction errors are correlated across time for the same event if $\gamma > 0$. One way to see this is to remove the auto-correlation by setting $\phi = 0$. Then the forecast error for the prediction made in period $t-1-h$ simplifies to $\varepsilon^*_{t,t-1-h} = \varepsilon_{t,t-1-h}$, which depends on earlier forecast errors made for the same period t through equation 6.

In the typical rational expectations setting (e.g. a dynamic stochastic general equilibrium (DSGE) model), it is assumed that $\gamma = 0$ due to FIRE and there is no persistence, while frictions cause $\gamma > 0$ and persistence (e.g. Mankiw and Reis (2002), Sims (2003) or Coibion and Gorodnichenko (2015)). The assumption follows from the argument that rational agents update their expectations after a shock and should not put any weight on the predictions that did not include the shock. However, this assumption is not generally consistent with rational expectations, as setting $\gamma = 0$ does not always minimize the prediction error. Specifically, $\gamma = 0$ only minimizes the prediction error provided that the expected prediction error variance of

[7] While γ is assumed to be time $t-h$ and horizon h specific, the subscripts are dropped here for readability.

new information $E\sigma^2_{\nu_{t,t-h}} = 0$, meaning that if agents anticipate that the new information they receive is noisy, they should not base their new prediction exclusively on the new information.

Theorem 2.1. *Provided $\sigma^2_{\nu_{t,t-h}} > 0$, predictions of rational agents satisfy $\gamma > 0$.*

Proof. If agents are rational, they minimize the (expected) prediction error and predictions are unbiased. Now suppose $\gamma = 0$ was optimal. In that case, $\hat{\eta}_{t,t-h} = x_{t,t-h}$ and the prediction error $\varepsilon_{t,t-h}$ simplifies to $\nu_{t,t-h}$ with variance $\sigma^2_{\nu_{t,t-h}} > 0$. While agents do not know the prediction error or variance ex ante, they form unbiased expectations for both, meaning that $E\sigma^2_{\nu_{t,t-h}} > 0$ as well. Given the independence assumption between $\nu_{t,t-h}$ and $\varepsilon_{t,t-h-1}$ and $\gamma = 0$, agents have access to $\hat{\eta}_{t,t-h-l}$, a prediction for η_t that is independent from $\hat{\eta}_{t,t-h}$. The second forecast $\hat{\eta}_{t,t-h-l}$ has the prediction error $\varepsilon_{t,t-h-1}$ and error variance $\sigma^2_{\varepsilon_{t,t-h-1}} > 0$ with the unbiased expected variance $E\sigma^2_{\varepsilon_{t,t-h-1}} > 0$. Now consider the alternative prediction

$$\hat{\eta}^*_{t,t-h} = \frac{E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\nu_{t,t-h}} + E\sigma^2_{\varepsilon_{t,t-h-1}}} x_{t,t-h} + \frac{E\sigma^2_{\nu_{t,t-h}}}{E\sigma^2_{\nu_{t,t-h}} + E\sigma^2_{\varepsilon_{t,t-h-1}}} \hat{\eta}_{t,t-h-l}$$

which uses the Bates and Granger (1969) inverse variance weights that are optimal for two independent predictions. If $\gamma = 0$ was optimal, $\hat{\eta}_{t,t-h}$ should have a lower expected error variance than $\hat{\eta}^*_{t,t-h}$. The expected prediction error variance of $\hat{\eta}^*_{t,t-h}$ is

$$E\sigma^2_{\varepsilon_{t,t-h}} = \frac{E\sigma^2_{\varepsilon_{t,t-h-1}} E\sigma^2_{\nu_{t,t-h}}}{E\sigma^2_{\nu_{t,t-h}} + E\sigma^2_{\varepsilon_{t,t-h-1}}} < \frac{E\sigma^2_{\varepsilon_{t,t-h-1}} E\sigma^2_{\nu_{t,t-h}} + (E\sigma^2_{\nu_{t,t-h}})^2}{E\sigma^2_{\nu_{t,t-h}} + E\sigma^2_{\varepsilon_{t,t-h-1}}} = E\sigma^2_{\nu_{t,t-h}}$$

Hence there is a contradiction and rational agents satisfy $\gamma > 0$ as the optimal inverse variance weights are always positive. □

As shown in Theorem 2.1, rational agents generally put a positive weight on their past prediction. This result implies that the requirement for transitory shocks to become persistent ($\gamma > 0$) is satisfied even under FIRE. This effect has previously been ignored in typical DSGE models and as a result, there is the distinction between classical real business cycle models without frictions and no persistence of transitory shocks and the new Keynesian models that include frictions and have persistent transitory shocks. Theorem 2.1 can help bridge this gap as it shows that there is some persistence even without deviations from FIRE. Indeed, as the estimates below show (e.g. Table 4), most of the persistence is driven by optimal forecasting behavior and not frictions.

Based on Theorem 2.1, the rational expectations (optimal Bates and Granger (1969)) weight on the past prediction in period $t-h$ for horizon h is

$$\gamma = \frac{E\sigma^2_{\nu_{t,t-h}}}{E\sigma^2_{\nu_{t,t-h}} + E\sigma^2_{\varepsilon_{t,t-h-1}}} \tag{7}$$

Intuitively, the weight on the past prediction is larger, the higher the ex ante uncertainty of new information obtained since the previous prediction. Similarly, the weight becomes smaller, the lower the ex ante uncertainty of the previous prediction.

As shown in Coibion and Gorodnichenko (2015) for example, there is evidence of deviations from FIRE in the form of information frictions that cause agents to put a higher than optimal weight on the past prediction. Due to the auto-correlation, there are three separate ways how the information friction can enter this model: First, agents could put an additional weight on the past prediction. This could be due to agents infrequently updating their prediction consistent with the Mankiw and Reis (2002) or Calvo (1983) setting. Second, agents could put a higher than optimal weight on the auto-regressive term in their prediction and third, agents could put a higher than optimal weight on the past signals. There are two main reasons, why the third case is utilized here: First, this leads to a tractable solution which can extract both γ and the friction coefficient. Second, in the data sets used here, forecasters change their predictions essentially every period making infrequent updating not very likely.[8] Empirical evidence in support of this choice is provided in Section 4.2.

Assume that the agent puts an additional weight of λ on the old prediction of $\hat{\eta}_{t,t-h-1}$

[8]The only exception are the CPI prediction in the SPF, where 20% of the forecasters keep the prediction unchanged.

and $1-\lambda$ on the rational expectations forecast.[9] The γ from equation 7 then changes to

$$\gamma = \frac{E\sigma^2_{\nu_{t,t-h}} + \lambda E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\varepsilon_{t,t-h-1}} + E\sigma^2_{\nu_{t,t-h}}} \tag{8}$$

A derivation of this expression can be found in the appendix. Note that λ can be estimated separately as in Coibion and Gorodnichenko (2015) for example. However, while estimating λ can determine whether there are deviations from FIRE, λ by itself does not generally provide information about what drives the persistence. Unless $E\sigma^2_{\nu_{t,t-h}} = 0$, the persistence due to λ depends on $\frac{E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\varepsilon_{t,t-h-1}}+E\sigma^2_{\nu_{t,t-h}}}$. In order to determine whether the weights and hence the persistence is mainly driven by the deviations of FIRE or are affected at least economically significanty, it is thus necessary to measure $\frac{\lambda E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\varepsilon_{t,t-h-1}}+E\sigma^2_{\nu_{t,t-h}}}$ and not just λ.

Having shown that there is persistence even under FIRE begs the question as to what data generating process (DGP) is consistent with this model. For example, an AR(1) model with unpredictable innovations would not be consistent with this model as agents do not receive independent signals. One DGP that is consistent with the model could similar to the setup by Kydland and Prescott (1982), who assume that agents cannot distinguish transitory and permanent shocks, but modified such that agents know everything about the present and only have uncertainty about the future. Specifically, one could assume that future shocks can either be correlated with all previous shocks or not. One can further assume that if the shocks were correlated often in the recent past, the future shock is less likely to be correlated with past shocks and conversely if the shocks were correlated with past shocks less often, the future shock is more likely to be correlated with all past shocks. Under rational expectation, agents predict that future shocks are always somewhat correlated with previous shocks, meaning they put a non-zero weight on all previous shocks. As a result, a transitory shock will have a persistent effect in the predictions of future periods because there is a chance that the future shock is correlated with the transitory shock.[10] The well documented information frictions

[9] As in the case of γ, the λ can vary across time and horizon but the subscripts are omitted to improve readability.

[10] In this setup, there are no strictly transitory shocks because there must be a correlated shock at some

can further increase the persistence of this simple model.

3 Data

Two data sets are utilized to measure the weight on old information an to see how it evolves over time. Since the model is defined for a single agent, the baseline results use the Greenbook (now Tealbook) forecasts from the Federal Reserve are prepared by the staff for every FOMC meeting. These forecasts are made available with a five year delay but are available for a number of variables for every meeting. Because the forecasts are for variables at a quarterly frequency, the weight cannot easily be estimated at a higher frequency. As a result, only the forecasts of the first meeting in a quarter are utilized.

The weight based on the mean prediction in the Survey of Professional Forecasters (SPF) is presented as a comparison. This survey does not have any delayed publication and allows to obtain a more real-time weight. However, as the predictions are the averages of a set of forecasters, this average is treated as if it was the forecast of a single forecaster. As shown in Mankiw and Reis (2002), the average forecast of a survey might be affected by (information) rigidity. Specifically, agents might not update their forecasts every period leading to a higher than optimal weight on the old prediction. As the Greenbook forecasts are newly prepared for every FOMC meeting, they should not be affected by this issue (e.g. see Messina et al. (2015)). This allows to compare the excess weights on old information of the two datasets. Both datasets start in the late 1960s and the Greenbook forecasts sample finishes in Q4 2013, while the SPF sample goes until Q4 2019.

For the realized value, the first release is used. The results presented here are robust to using any of the first three releases.

---

point in the future for the positive weight to be consistent with rational expectations. This can be remedied by assuming that the correlation is not with all previous periods but a finite subset of them.

4 Empirical Estimation

Data on the new information $x_{t,t-h}$ is not directly available even ex post and hence the weight γ cannot be extracted directly from equation 5. It is also not possible to simply regress the old prediction on the new prediction because the error term would be $x_{t,t-h}$ and hence the past prediction would be correlated with the error term as both include A_t. However, with the independence assumption between the noise of new information $\nu_{t,t-h}$ and the old prediction $\hat{\eta}_{t,t-h-1}$ mentioned above, one could estimate

$$\varepsilon_{t,t-h} = (1-\gamma)\nu_{t,t-h} + \gamma\varepsilon_{t,t-h-1} \tag{9}$$

where ε_j is the prediction error of the corresponding forecast $\hat{\eta}_j$. Because $\nu_{t,t-h}$ is independent from $\varepsilon_{t,t-h-1}$, one can then estimate the equation

$$\varepsilon_{t,t-h} = \gamma\varepsilon_{t,t-h-1} + \mu_t \tag{10}$$

where μ_t is the error term. Note that when estimating γ, one needs to take into account that μ_t follows an MA(h) as the observations are overlapping. Estimating this equation requires to obtain $\varepsilon_{t,t-h}$ and $\varepsilon_{t,t-h-1}$ in the first place. Rearranging equation 4, one can obtain

$$\varepsilon_{t,t-h} = \varepsilon^*_{t,t-h} - \phi\varepsilon^*_{t-1,t-h} \tag{11}$$

meaning that $\varepsilon_{t,t-h}$ is the corresponding overall prediction error minus ϕ times the prediction error of the prediction made in period $t-h$ for period $t-1$. Note that estimating this equation also tests whether there even is a signal. Without any signal, this regression would return a coefficient equal to one.[11]

In a first step, equation 9 is estimated for real GDP over the entire sample for both the Greenbook (GB) and the SPF in order to obtain an estimate for the weight agents put on the past prediction and hence the persistence of economic shocks. For the more recent forecast,

[11] This might allow to test, whether forecasters received any signal, which could be compared to other measures of forecasting ability like Bürgi and Boumans (2020).

the current quarter forecast is used and for the older forecast, two separate specifications are estimated. The first one uses the one quarter ahead forecast and the second uses the four quarter ahead forecasts. The results are presented in Table 1 with Newey-West standard errors.

Table 1: Overall Weights GDP

	Dependent variable:			
	Error Current Quarter			
	(1)	(2)	(3)	(4)
	GB	SPF	GB	SPF
Error 1Q	0.668***	0.637***		
	(0.050)	(0.032)		
Error 4Q			0.624***	0.565***
			(0.059)	(0.035)
Constant	0.179	0.225***	0.343**	0.301***
	(0.131)	(0.084)	(0.161)	(0.114)
Observations	179	202	154	194
R^2	0.657	0.725	0.582	0.623

This table reports the coefficient and Newey-West standard errors. *, ** and *** imply significantly different from 0 at 10%, 5% and 1% level, respectively.

The forecasts put less weight on the old forecast relative to the new information for the four quarter ahead forecast on average and hence there is less persistence. The one quarter ahead forecast includes three more quarters of information than the four quarter ahead forecast. As a result, the new information received until the current quarter forecast is less informative for the one quarter ahead forecast and the old forecast has a higher weight.

Furthermore, there is no statistically significant difference in the weight on old information between the two surveys. The model predicts a zero constant, which is rejected. These deviations from rational expectations affect the prediction, but they do not affect the weight on the past prediction and the persistence. Indeed, omitting the constant leaves the results and significance essentially unchanged. Additional regressions showing the robustness of these estimates across types of agents and how the coefficients vary by horizon are shown in the appendix.

As with output, the regression equation 9 is estimated for both CPI inflation and the GDP deflator. Table 2 shows the results. The GDP deflator results are very similar to the GDP results; the SPF and Greenbook forecasts put similar weights on the past prediction and the weight is very close to the ones for GDP. For the CPI, the results are different. For the Greenbook forecast, the weight on old information for the CPI is about half of the one for GDP or the deflator.

4.1 Optimal Weight On Past Predictions

Having established that there is a substantial weight on the past prediction immediately raises the question, what share of this weight is due to information frictions and what share would remain even without frictions. While the expected prediction error variances cannot be directly observed, one can observe the actual prediction error variances ex post. Specifically, $\sigma^2_{\varepsilon_{t,t-h-1}}$ is the variance of the past prediction error and $\sigma^2_{\nu_{t,t-h}}$, the variance of the independent new information, can be calculated based on equations 9 and 10 from the error term. Specifically, it is

$$\sigma^2_{\nu_{t,t-h}} = var(\nu_{t,t-h}) = var(\mu_t/(1-\gamma)) \tag{12}$$

This variance of the new information is also a new way to estimate the ex post uncertainty of new information and thus the information flow. The (ex post) prediction uncertainty (e.g. $\sigma^2_{\varepsilon_{t,t-h-1}}$) is well established in the literature (e.g. Zarnowitz and Lambros (1987), Lahiri and

Table 2: Overall Weights Inflation

	Dependent variable:							
	Error Current Quarter							
	CPI				Deflator			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	GB	SPF	GB	SPF	GB	SPF	GB	SPF
Error 1Q	0.395***	0.608***			0.552***	0.636***		
	(0.048)	(0.028)			(0.047)	(0.029)		
Error 4Q			0.340***	0.555***			0.533***	0.617***
			(0.040)	(0.025)			(0.063)	(0.031)
Constant	−0.055	0.016	−0.035	0.115***	−0.066	−0.080	−0.089	−0.103
	(0.074)	(0.027)	(0.082)	(0.035)	(0.100)	(0.064)	(0.079)	(0.066)
Observations	135	153	132	150	179	202	154	194
R^2	0.490	0.872	0.377	0.851	0.452	0.687	0.424	0.662

This table reports the coefficient and Newey-West standard errors. *, ** and *** imply significantly different from 0 at 10%, 5% and 1% level, respectively.

Sheng (2010) or Jurado et al. (2015)) and measures the uncertainty of a given event. The uncertainty of new information measures by how accurate the information was over a given time interval. This has the benefit that unlike the prediction uncertainty, is not influenced by events prior to the period. For example, the one quarter ahead prediction uncertainty might be low for a given period because it got reduced dramatically several quarters earlier, which is not an issue with the uncertainty of new information. As a result, it might be easier to compare this uncertainty to other uncertainty indices like the VIX or the one proposed by Baker et al. (2016) or Bürgi and Sinclair (2020). While analyzing and comparing this uncertainty index is beyond the scope of the paper, a graph of the ex post uncertainty of new information is included in the appendix. It is also important to note, that this uncertainty of new information based on $\sigma^2_{\nu_{t,t-h}}$ is model specific. This means that two forecasters receiving the same data might not have the same uncertainty of new information, as they weight the data they receive differently.

After calculating the two variances, the ex post optimal weight on the past prediction then becomes

$$\gamma_{opt} = \frac{\sigma^2_{\nu_{t,t-h}}}{\sigma^2_{\nu_{t,t-h}} + \sigma^2_{\varepsilon_{t,t-h-1}}} \tag{13}$$

In the case with an auto-correlated underlying variable, this expression includes the part of the weight due to said auto-correlation. One can now compare the estimated γ that potentially contains information frictions to the ex post optimal weight γ_{opt}. Before doing so, it is important to check whether there are statistically significant deviations from FIRE. Nordhaus (1987) proposed the following two regressions in order to test whether there are significant deviations from rational expectations:

$$A_t - F_{t,t-h} = \alpha + \beta(F_{t,t-h} - F_{t,t-h-1}) + \mu_t \tag{14}$$

$$F_{t,t-h} - F_{t,t-h-1} = \alpha + \beta(F_{t,t-h-1} - F_{t,t-h-2}) + \mu_t, \tag{15}$$

In both cases, $\alpha = 0$ implies that the forecasts are unbiased and $\beta = 0$ implies that the correct weight is put on the past prediction. Coibion and Gorodnichenko (2015) in turn showed that the β coefficient in Equation 14 is equal to $\frac{\lambda}{1-\lambda}$ and allows to easily measure λ.[12] Also, both tests for deviations from the optimal weights ignore periods where the forecasts for both horizons are identical. As a result, they cannot identify Mankiw and Reis (2002) type rigidity due to leaving forecasts unchanged at the individual level. In order to get around this issue, Coibion and Gorodnichenko (2015) proposed to average individual forecasts across agents, while Bürgi (2016) proposed to instrument the individual forecasts

[12]Note that the λ obtained through estimating Equation 14 is numerically identical to the λ obtained by estimating γ and rearranging Equation 8 when using the ex post variances. A derivation of this is provided in the appendix.

utilizing the average forecast. As the mean SPF prediction is utilized here, this should not be an issue and the Greenbook results do not appear to be much affected by this either. In order to estimate these equations based on the model in the present paper, it is necessary to replace F_j by $\hat{\eta}_j$ and A_t by η_t. Table 3 shows the regression coefficients. The regressions show that forecasters put significantly too high a weight on the past prediction for the one quarter GDP predictions in the SPF and the CPI predictions in the SPF. The λ implied by these estimates are $\lambda = \frac{\beta}{1+\beta}$ and they are all sizable at 0.271 for GDP, 0.885 for the one quarter CPI prediction and 0.765 for the four quarter CPI prediction. Conversely, a negative and significant coefficient is observed for the GDP deflator in the Greenbook.

Table 3: Estimate of $\frac{\lambda}{1-\lambda}$

One Quarter	α	β	Four Quarters	α	β
GDP GB	-0.034	-0.032	GDP GB	-0.132	-0.107
	(0.162)	(0.143)		(0.145)	(0.129)
GDP SPF	-0.16	0.271**	GDP SPF	-0.12	0.138
	(0.146)	(0.128)		(0.133)	(0.114)
CPI GB	0.082	0.145	CPI GB	0.054	0.053
	(0.101)	(0.094)		(0.101)	(0.132)
CPI SPF	0.013	0.885***	CPI SPF	-0.133*	0.765***
	(0.062)	(0.165)		(0.077)	(0.139)
DEF GB	-0.001	-0.214**	DEF GB	0.066	-0.228***
	(0.105)	(0.089)		(0.08)	(0.068)
DEF SPF	0.102	0.147	DEF SPF	0.13	0.122
	(0.086)	(0.103)		(0.084)	(0.096)

This table reports the coefficient and Newey-West standard errors in brackets for the regression $-\varepsilon_{t,t-h} = \alpha + \beta(\varepsilon_{t,t-h} - \varepsilon^*_{t,t-h-1}) + \mu_t$. *, ** and *** imply significantly different from 0 at the 10%, 5% and 1% level, respectively.

In order to see, how these significant deviations relate to the overall weight on the past prediction, it is necessary to calculate what share of the overall weight these λs imply. One way to do so is to compare the estimated γ that potentially contains information frictions to the ex post optimal weight γ_{opt}. Note that because the estimated γ uses the expected variances and γ_{opt} uses the ex post realized variances, one would expect that there are small differences between the two.

Table 4: Share of Weight on Old Prediction (γ) Explained by Optimal Weight (γ_{opt})

	1 Quarter		4 Quarters	
	(1)	(2)	(1)	(2)
	GB	SPF	GB	SPF
GDP	1.016	0.846	1.064	0.894
CPI	0.777	0.429	0.897	0.386
Deflator	1.174	0.916	1.199	0.924

As Table 4 shows, γ_{opt} makes up more than three quarters of γ with the exception of the CPI prediction in the SPF. For CPI inflation, λ accounts for around 60% of the weight, implying that the weight without friction is still substantial. As a result, information frictions are not generally the main driver of the weight on past information. Overall, the results presented so far have three important implications: First, there are deviations from rational expectations and they cause a positive weight on the past prediction which is economically meaningful. Second, even without these deviations, there is a substantial weight on the past prediction. Third, the part of the weight on the past prediction due to rational expectations is in most cases larger than the part due to deviations from rational expectations.

4.2 What Causes The Excess Weight

In section 2, it is assumed that if agents deviate from rational expectations, they neither put an additional overall weight on previous predictions, nor put a higher than optimal weight on the auto-regressive term. Instead, it is assumed that agents put a higher than optimal weight on previous signals. This section aims to provide some empirical evidence in support of this assumption. Consider the following modified equation 11

$$\varepsilon^*_{t,t-h} = \varepsilon_{t,t-h} + \theta\varepsilon^*_{t-1,t-h} + \lambda_{all}\varepsilon^*_{t,t-h-1} \tag{16}$$

Where $\varepsilon_{t,t-h}$ is the weighted average of the signal noise with a potentially larger than optimal weight on past signals, $\theta\varepsilon^*_{t-1,t-h}$ is the auto-correlation term where $\theta > \phi$ if agents put a higher than optimal weight on this term, and $\lambda_{all}\varepsilon^*_{t,t-h-1}$ captures any additional overall weight placed on the previous prediction. This equation hence includes all three types of deviations discussed in section 2. While all the variables denoted with a * are known, $\varepsilon_{t,t-h}$ is not and is correlated with $\varepsilon^*_{t,t-h-1}$. As a result, estimating this equation with all known variables leads to biased estimates.[13]

Instead, one can estimate

$$\varepsilon^*_{t,t-h} = \alpha + \theta\varepsilon^*_{t-1,t-h} + \mu_t \tag{17}$$

where μ_t is the error term that includes $\varepsilon_{t,t-h} + \lambda_{all}\varepsilon^*_{t,t-h-1}$ and α should be equal to 0. Under the hypothesis that there are either no deviations from rational expectations or that the deviations only come from a higher than optimal weight on the previous signals, $\lambda_{all} = 0$ and $\theta = \phi$. If instead, agents put too high a weight on the auto-correlation term, $\theta > \phi$. Similarly, if agents put too high a weight on the overall past prediction and $\lambda_{all} > 0$, then $\varepsilon^*_{t-1,t-h}$ is correlated with the error term leading to $\theta > \phi$.[14] One can thus utilize this regression to test whether deviations from rational expectations are due to a higher overall weight on past

[13] As $cov(\varepsilon_{t,t-h}, \varepsilon^*_{t,t-h-1}) > 0$ and $cov(\varepsilon^*_{t-1,t-h}, \varepsilon^*_{t,t-h-1}) > 0$, one would expect that $\hat{\lambda}_{all}$ would be biased upward and $\hat{\theta}$ in turn would be biased downward if the regression was run. Furthermore, $cov(\varepsilon^*_{t-1,t-h}, \varepsilon_{t,t-h}) = 0$ as the two errors are for different time periods.

[14] There is also an additional inefficiency case where $\theta > \phi$. Specifically, the prediction error across time periods could be correlated (i.e. $cov(\varepsilon^*_{t-1,t-h}, \varepsilon_{t,t-h}) > 0$), which would also constitute a deviation from rational expectations.

predictions or due to overestimating the noise in the new information resulting in a higher than optimal weight on past signals. If the estimated θ is larger than the auto-correlation of the underlying variable, the assumption made in section 2 about the expectation formation for the data sets needs to be revised.

Table 5 shows the regression results for the one and four quarter ahead predictions for GDP. The auto-correlation coefficient for GDP is 0.49, meaning that θ is not larger than ϕ. This result is also in line with finding limited evidence of deviations from rational GDP expectations in Table 3. Comparing the results to the ones obtained there, one notices that the coefficients for the one quarter ahead predictions in the SPF and the Greenbook are very similar despite finding deviations from full information rational expectations for the SPF. This suggests as well, that the deviations are not due to a higher overall weight or over-weighting the auto-correlation term.

The regressions are repeated for CPI inflation and the deflator and the results reported in Table 6. For the CPI, there auto-correlation coefficients are 0.602 for the Greenbook and 0.371 for the SPF. The difference stems from the Greenbook sample starting earlier.[15] The auto-correlation coefficient for the deflator is 0.835. As with the GDP results above, comparing the results here with the ones in Table 3 also suggests that the deviations from full information rational expectations are not due to over-weighting the auto-correlation term or the overall past prediction.

The estimated coefficients in Table 6 suggest that there is little evidence that $\theta > \phi$ or $\lambda_{all} > 0$ as the estimated coefficients are not significantly larger than the auto-correlation of the underlying series. These estimates are thus broadly in line with the assumption on deviations from rational expectations made in Section 2. In addition, the results presented in this section provide some evidence against the sticky information model where agents infrequently update their predictions. This is because the deviations from FIRE identified

[15] When matching samples are used, both auto-correlation coefficients are around 0.37 and the one quarter ahead coefficient for the Greenbook becomes significantly larger than ϕ.

Table 5: Test For Alternative Deviations From Rational Expectations (GDP)

	Dependent variable:			
	Error 1Q		Error 4Q	
	(1)	(2)	(3)	(4)
	GB	SPF	GB	SPF
$\varepsilon_{t-1,t-1}$	0.289***	0.273***		
	(0.108)	(0.102)		
$\varepsilon_{t-1,t-4}$			0.363***	0.498***
			(0.058)	(0.078)
Constant	−0.196	−0.217	−0.325	−0.436**
	(0.208)	(0.176)	(0.199)	(0.195)
Observations	179	202	154	194
R^2	0.055	0.041	0.161	0.252

This table reports the coefficient and Newey-West standard errors in brackets for the regression $\varepsilon_{t,t-h} = \alpha + \beta\varepsilon_{t-1,t-h} + \mu_t$. *, ** and *** imply significantly different from 0 at the 10%, 5% and 1% level, respectively. The AR(1)-coefficient for GDP is 0.490 with a standard error of 0.091.

Table 6: Test For Alternative Deviations From Rational Expectations (Inflation)

	Dependent variable:							
	CPI				Deflator			
	Error 1Q		Error 4Q		Error 1Q		Error 4Q	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GB	SPF	GB	SPF	GB	SPF	GB	SPF
$\varepsilon_{t-1,t-1}$	0.724***	0.439***			0.470***	0.396***		
	(0.159)	(0.099)			(0.151)	(0.148)		
$\varepsilon_{t-1,t-4}$			0.263***	0.283***			0.678***	0.712***
			(0.085)	(0.086)			(0.178)	(0.130)
Constant	−0.053	−0.162	−0.070	−0.365**	0.077	−0.053	0.011	−0.040
	(0.159)	(0.124)	(0.183)	(0.158)	(0.120)	(0.128)	(0.089)	(0.120)
Observations	135	153	132	150	179	202	154	194
R^2	0.144	0.076	0.071	0.077	0.138	0.093	0.348	0.416

This table reports the coefficient and Newey-West standard errors in brackets for the regression $\varepsilon_{t,t-h} = \alpha + \beta\varepsilon_{t-1,t-h} + \mu_t$. *, ** and *** imply significantly different from 0 at the 10%, 5% and 1% level, respectively. The AR(1)-coefficient for the deflator is 0.8347 with a standard error of 0.0513. The AR(1)-coefficients for the CPI are 0.602 and 0.3712 for the Greenbook and the SPF, respectively, each with a standard error of 0.102 and 0.0725.

in Table 3 do not correspond to a higher than optimal weight on the auto-correlation term. Also, the weight on the auto-correlation term does not appear greater for predictions with deviations from FIRE relative to predictions that are in line with FIRE.

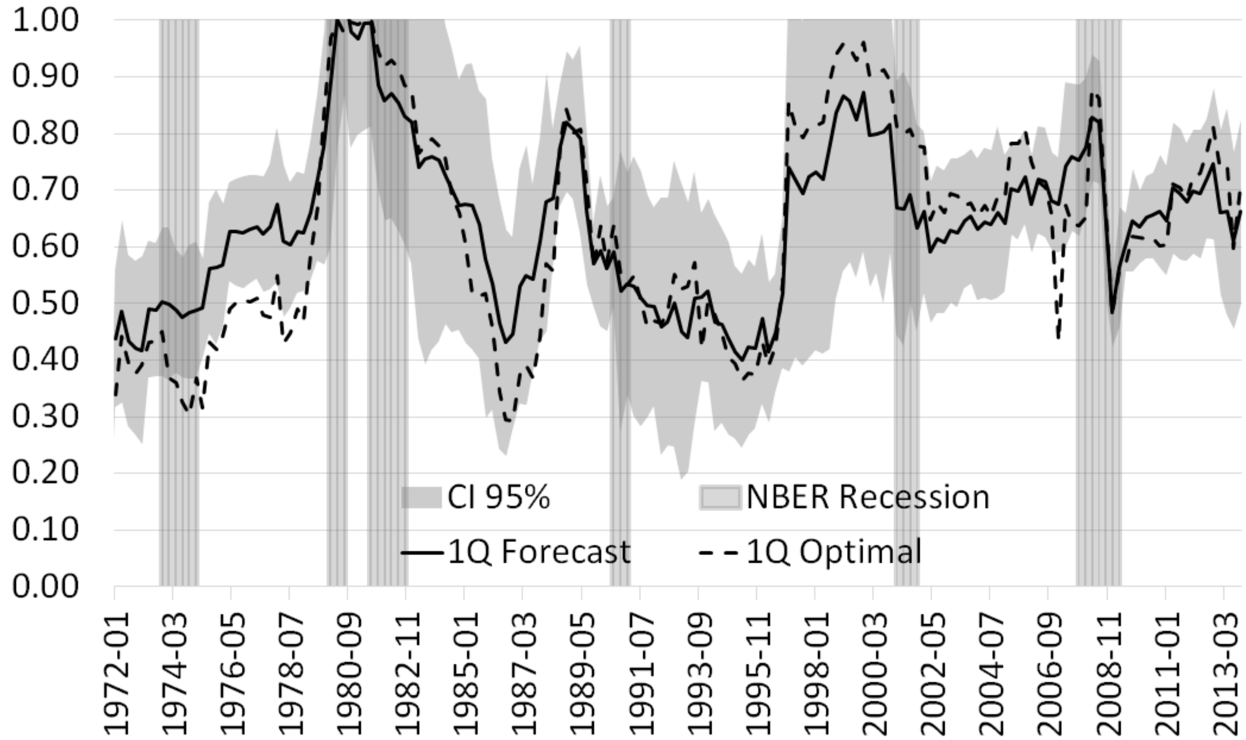
5 Variation over time

The weight on the old prediction is directly linked to the (expected) variances of the prediction error (uncertainty). The uncertainty is well known to vary over time, which results in the weights varying over time as well.[16] As a result one might be interested to know how well forecasters are tracking this natural fluctuation in the optimal weights. Also, tracking the weights over time allows to determine how the overall weights and the part of the weight due to information frictions evolve over time. For example, it could be the information frictions are only affecting a small part of the sample. In order to estimate time varying weights, a 20 quarter rolling window regression is estimated with OLS and the 95% confidence intervals in the graphs use Newey-West standard errors. Aside from the rolling window, the setup is the same as for Table 1. The regression coefficient γ for the one quarter ahead Greenbook forecasts is then plotted in Figure 1 together with shadings for the National Bureau of Economic Research (NBER) recessions.

The graph shows large cyclical swings in the weights over time. These swings are largely driven by the uncertainty of the previous prediction and new information (see Equations 7 and 13 above or the appendix on the uncertainty of new information). When the weight is high, the uncertainty of new information is relatively high and the uncertainty of old information relatively low and the opposite is the case when the weight is low. One exception to this is around the early 80s, when the weights essentially reach one, which is caused by the uncertainty of new information increasing to extraordinarily high levels. As high ex post uncertainty also implies larger shocks on average, this suggests that the weight on old

[16] See for example Zarnowitz and Lambros (1987), Lahiri and Sheng (2010) or Jurado et al. (2015)

Figure 1: Real GDP Weight on the Old Forecast (γ) vs Optimal Greenbook

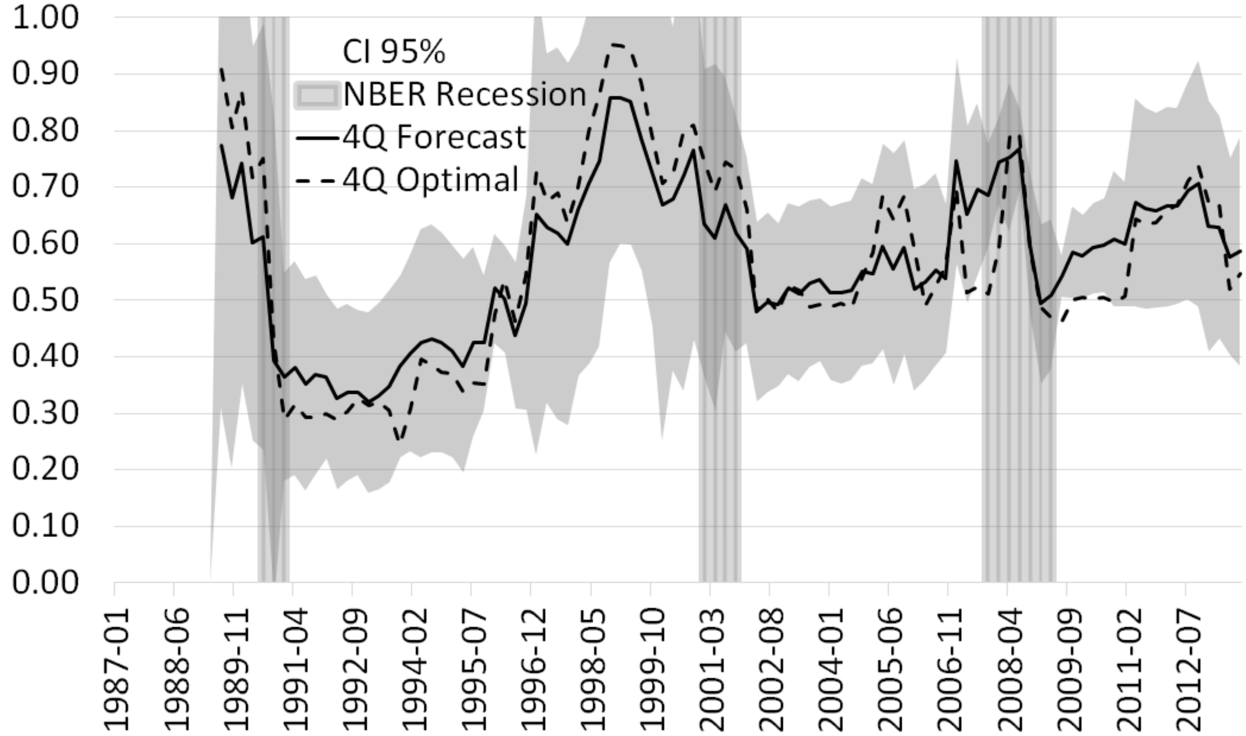


information is inversely related to the magnitude of the shocks. As a result, the weight after smaller shocks is larger than the estimates in Table 1 and above 0.7, but the weight after larger shocks might be closer to 0.4-0.5.

The figure also shows the ex post optimal weight γ_{opt}. A comparison of the optimal weight to the actual weight γ shows that the two are relatively similar. Specifically, the actual weight is able to capture the large swings in the optimal weight quite well. While the Greenbook forecasts sometimes put too much and sometimes too little weight on the past forecasts for even extended periods, these deviations are small relative to the large swings of the optimal weights and typically within the 95% confidence intervals. The only exception is the period at the beginning of the sample, where the difference is statistically significant.

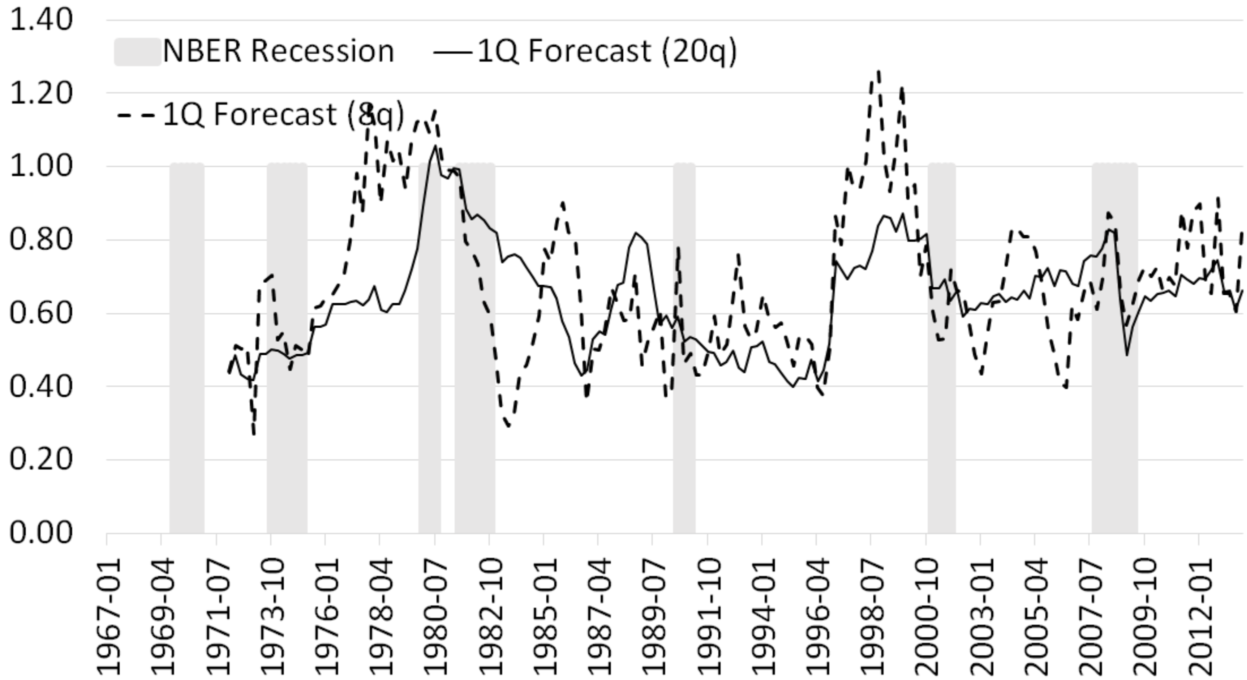
Figure 2 uses the same setup as Figure 1 but uses the four quarter ahead prediction instead. One can notice that the cycles at the different horizons are very similar. Specifically, in periods when the weight is high, there is little difference between the weights for the one or four quarter ahead prediction and both are around 0.8. However, in the periods when the weight is low, there is less weight placed on the four quarter ahead forecast (0.3-0.4) than the one quarter ahead forecast (0.4-0.6). For the shorter four quarter ahead sample, the ex post optimal weight is also almost always within two standard deviations of the estimated γ.

Figure 2: Real GDP Weight on the Old Forecast (γ) vs Optimal Greenbook



Because there is an increase in the weight before recessions, one might think that it could be used as a predictor as well. However, there are two drawbacks to this. First, the weight on old information can only be estimated ex post, once the realized GDP growth is known. As realized GDP data is released with a significant lag, this is a limiting factor on how timely this indicator would be. Second, the 20 quarter window length might influence this finding. While the former issue is not easily resolved, it is possible to alter the window length to check whether the pattern is a statistical artifact or not. Figure 3 presents a comparison of the 20 quarter rolling window and a much shorter eight quarter rolling window estimation of the actual weight on old information.

Figure 3: Real GDP Weight on the Old Forecast (γ) Greenbook 20 Quarters vs. 8 Quarters

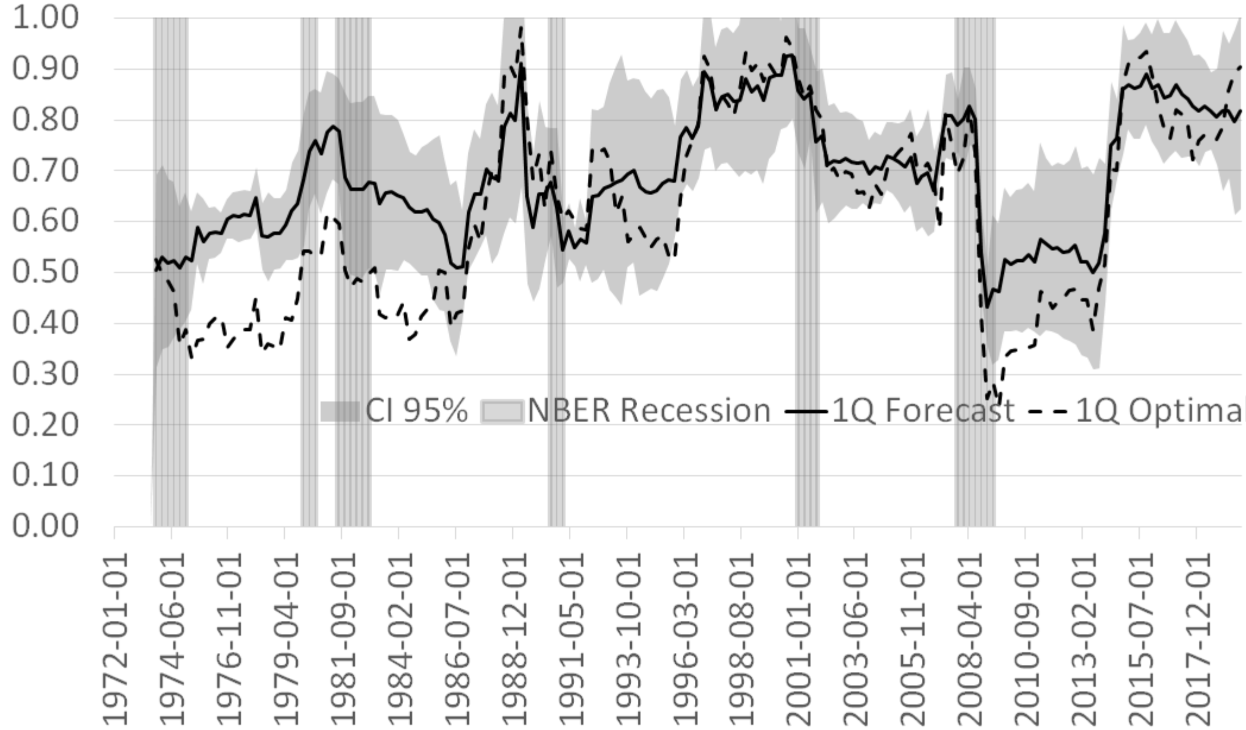


The comparison of the shorter and longer estimation windows shows that there is not always a dramatic increase of the weight on old information before recessions for both window lengths. In particular, there was no dramatic increase before the 2008 or the 90s recessions for the eight quarter window but there were clear increases before the 80s and 2000 recessions. There is still a dramatic decrease in the weight once a recession starts which is included for the next few quarters in the rolling regression.

Next, the weight on old information for the one quarter ahead SPF real GDP forecasts are presented in Figure 4.

The pattern for SPF is very similar to the one for the Greenbook in Figure 1. The cycles are almost identical and the general pattern is similar with the main difference before the

Figure 4: Real GDP Weight on the Old Forecast (γ) vs Optimal SPF

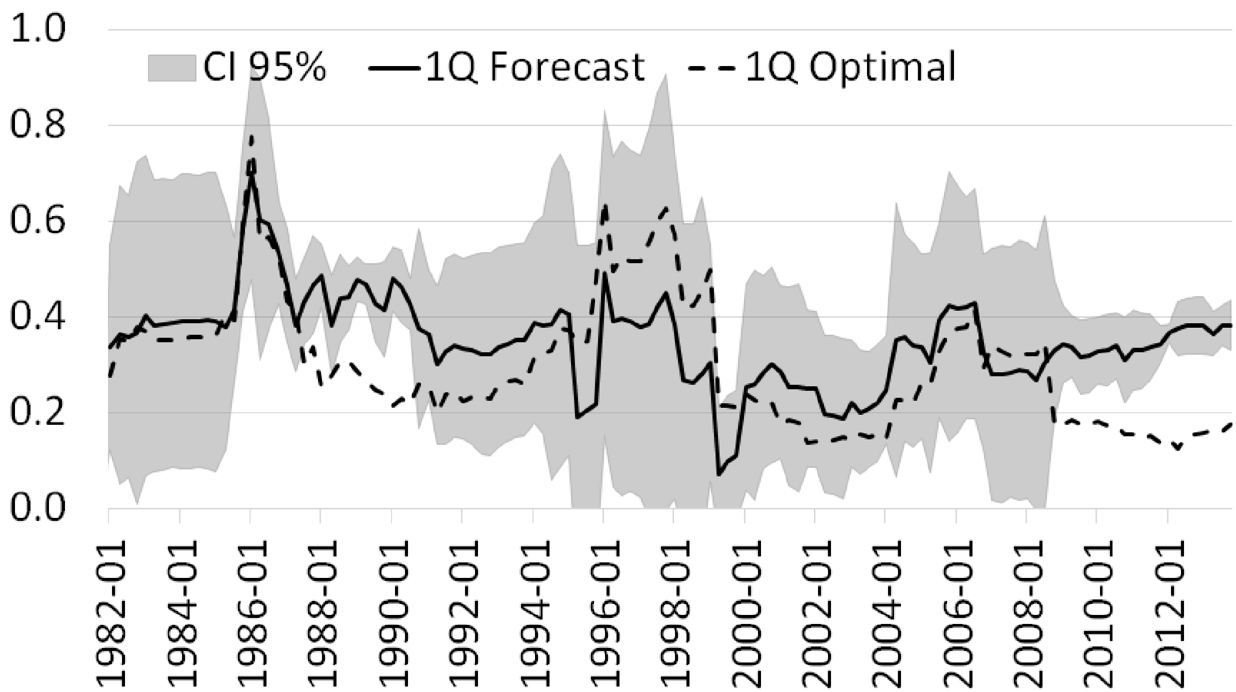


mid 80s. There is also some difference in levels for the two series. In the early 90s, the SPF had a much higher weight than the Greenbook. Also, since the SPF does not have the five year lag, it is possible to cover more recent periods as well. In the most recent years, the weight on the old prediction picked up significantly again. Comparing the optimal and actual weight on the old forecast for the SPF, the two are again fairly close to each other (except for the very beginning of the sample). The large swings in the optimal weights are again captured by the actual weights.

5.1 Inflation

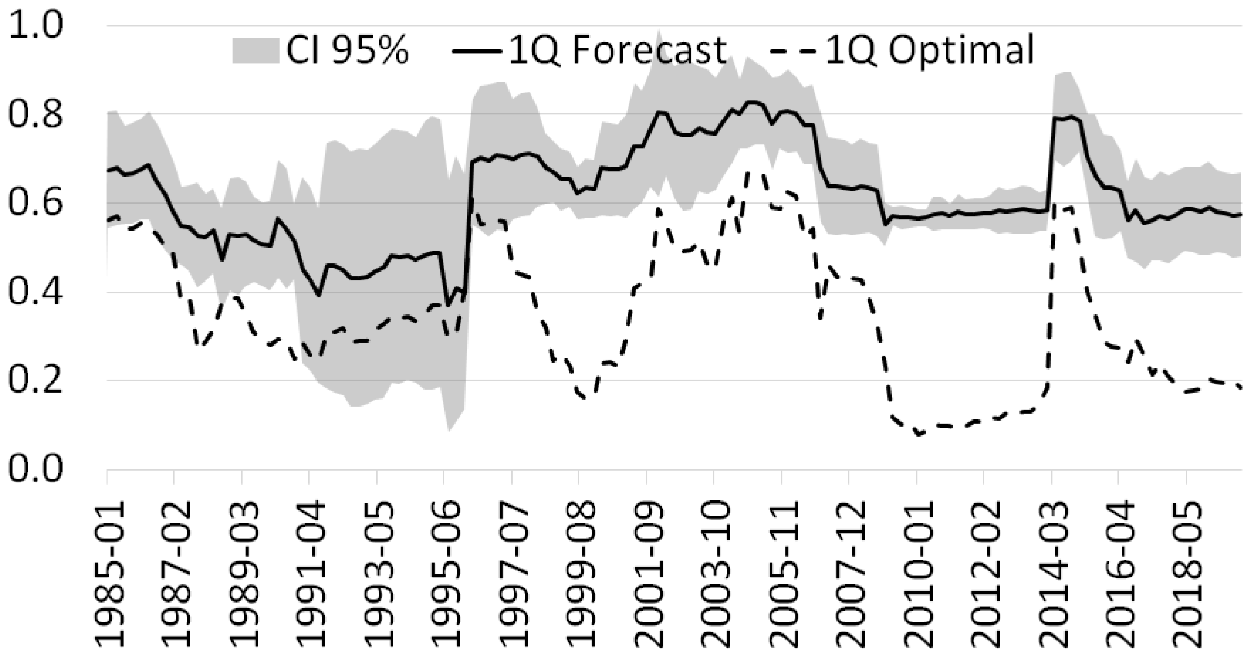
As the above results show large deviations for the CPI, this is the focus of this section and the graphs for the deflator are omitted. The pattern of the CPI weights is analyzed graphically across time. As with GDP, a rolling 20 quarter regression is run and Newey-West confidence intervals are shown. Unlike recessions for GDP, there are no cyclically higher and lower periods of inflation. As shown in Figure 5, the weight on old information for Greenbook forecasts also show less substantial swings, relative to output. While forecasters are able to track some of the larger movements in the optimal weight, forecasters struggle to get the exact magnitude right. In the most recent period, forecasters put too high a weight on the past prediction relative to the ex post optimal weight.

Figure 5: Weight on Old Information CPI Inflation Greenbook



A comparison with the SPF in Figure 6 shows a break in 1996 for the SPF. While the Greenbook forecasts remain within two standard deviations of the optimal weight, the SPF puts too much weight on old information and does not reduce it back down. As a result, the SPF puts a significantly larger than optimal weight on the past prediction since 1996. Note that already before this break, the SPF puts a higher than optimal weight on the past prediction, but to a much lesser extent than after. There are several potential explanation for the additional persistence observed in inflation forecasts. For example the averaging the forecast (see Mankiw and Reis (2002), Coibion and Gorodnichenko (2015) or Bürgi (2016) for a detailed mechanism), or a change in how CPI inflation is predicted after the Boskin et al. (1996) report. However, determining the exact cause is beyond the scope of this analysis. The smaller than optimal movements in the weights after 1996 appear to be mainly driven by the higher than optimal weights. Assuming a constant λ of 0.5 after 1996 matches the the actual closely.

Figure 6: Weight on Old Information CPI Inflation SPF



Another difference between the two data sources is that while the cycles appear to be synchronized, their magnitudes are not. For the Greenbook forecasts, the increase in the optimal weight in the late 90s is much larger than in the mid 2000s. In contrast, the increase in the mid 2000s for the SPF is the same or higher. The cycles in the weights for the CPI forecasts appear to coincide with the cycles in the wights for output. Specifically, there are periods of higher optimal weights on old information in the late 90s, mid 2000s and mid

2010s that are visible in both variables. However, these cycles are a bit less pronounced for CPI inflation. For both output and inflation, there is also jump in the weights in 2014, just after the last available data point for the Greenbook.

5.2 Time Varying Information Frictions

The GDP graphs above showed that particularly the beginning of the sample saw some periods with significant deviations from the ex post optimal weight γ_{opt} while the remainder of the sample saw little deviations. As a result, one might conclude that the deviations in the beginning are due to information frictions, which disappear after that period. One problem with this approach is that the ex post optimal weights might not be an appropriate approximation of the optimal weights in small samples. As γ is based on the expected variances, it is conceivable that the period was just a cluster of random prediction errors. One indicator of this period was special could be that the ex post optimal weights are the lowest during that period as well. Due to this issue, it is not directly possible to distinguish whether temporary deviations from the ex post optimal weights are due to information frictions or due to prediction errors. The exception to this are the CPI graphs from the SPF above as the differences are apparent throughout almost the entire sample. At least for the CPI, it can thus be concluded that the differences are likely due to deviations from rational expectations in the form of information frictions.

This problem persists for the above approaches on determining significant deviations from rational expectations. For example, the β coefficient in equation 14 without the assumption that expected and actual variances are equal is

$$\hat{\beta} = \frac{\frac{\gamma}{1-\gamma}\sigma^2_{\varepsilon_{t,t-h-1}} - \sigma^2_{\nu_{t,t-h}}}{\sigma^2_{\nu_{t,t-h}} + \sigma^2_{\varepsilon_{t,t-h-1}}} \tag{18}$$

and in general does not identify λ separately. A derivation of this result is presented in the appendix. Instead of the deviation from rational expectations, this expression is linked to the overall weights and estimating this expression in small samples might capture the cycles in the overall weights instead. Due to this, it is not possible to definitively determine whether the differences between γ and γ_{opt} found at the beginning of the GDP sample are due to information frictions.

6 Conclusion

This paper shows that agents put a positive weight on their past predictions for a fixed event even under rational expectations. This implies that information frictions are not necessary for transitory or permanent economic shocks to have persistent effects. Similarly, this also implies that the frictions are not necessary for monetary policy to have lasting effects as there can be persistence even without frictions. A new approach to estimate the combined weight with and without information frictions is presented. For GDP and inflation based on the GDP deflator, it is estimated that the overall weight is at least 0.55, while it is lower with 0.33-0.55 for CPI inflation on average based on the SPF and Greenbook predictions. It is further shown that a substantial part of this weight on the past prediction is present even without information frictions. Indeed, in most cases, while information frictions were found to increase the weight agents put on their past predictions, it is not the main determinant of that weight. Instead, the majority of the weight on the past prediction is due to optimal forecasting behavior under rational expectations. This has important implications for DSGE models, as the persistence of transitory shocks in new Keynesian models can mainly be attributed to rational expectations instead of frictions, and there is also persistence in classical real business cycle models without frictions.

The new estimation approach is also able to identify the ex post uncertainty of new information, a measure of the uncertainty of the information flow. Unlike the prediction uncertainty of a specific event in the future, this measure captures the reduction of the uncertainty over a given period. Further research might be able to determine how his measure and its properties compare to the many other uncertainty measures in the literature.

References

An, Z., Sheng, X., and Wallen, J. (2017). Forecaster inattention: Measurement, determinants and policy implications. *Mimeo*.

Andrade, P. and Le Bihan, H. (2013). Inattentive professional forecasters. *Journal of Monetary Economics*, 60(8):967 – 982.

Ang, A., Bekaert, G., and Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of monetary Economics*, 54(4):1163–1212.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.

Baker, S. R., McElroy, T. S., and Sheng, X. S. (2018). Expectation formation following large unexpected shocks. *Review of Economics and Statistics*, pages 1–45.

Bates, J. M. and Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.

Boskin, M., Dulberger, E., Gordon, R., Griliches, Z., and Jorgenson, D. (1996). The boskin commission report: Toward a more accurate measure of the cost of living. *Washington DC: Senate Finance Committee*.

Brunner, K., Cukierman, A., Meltzer, A. H., et al. (1980). Stagflation, persistent unemployment and the permanence of economic shocks. *Journal of Monetary Economics*, 6(4):467–92.

Bürgi, C. (2016). What Do We Lose When We Average Expectations? Working Papers 2016-013, The George Washington University, Department of Economics, Research Program on Forecasting.

Bürgi, C. and Boumans, D. (2020). Categorical Forecasts and Non-Categorical Loss Functions. CESifo Working Paper Series 8266, CESifo.

Bürgi, C. and Sinclair, T. M. (2020). What does forecaster disagreement tell us about the state of the economy? *Applied Economics Letters*, 0(0):1–5.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, 12(3):383–398.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45.

Clarida, R., Gali, J., and Gertler, M. (1999). The science of monetary policy: a new keynesian perspective. *Journal of economic literature*, 37(4):1661–1707.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78.

Diebold, F. X. and Rudebusch, G. D. (1989). Long memory and persistence in aggregate output. *Journal of monetary economics*, 24(2):189–209.

Döpke, J., Dovern, J., Fritsche, U., and Slacalek, J. (2008). Sticky information phillips curves: European evidence. *Journal of Money, Credit and Banking*, 40(7):1513–1520.

Dräger, L. and Lamla, M. J. (2012). Updating inflation expectations: Evidence from micro-data. *Economics Letters*, 117(3):807–810.

Galı, J. and Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of monetary Economics*, 44(2):195–222.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

Khan, H. and Zhu, Z. (2006). Estimates of the Sticky-Information Phillips Curve for the United States. *Journal of Money, Credit and Banking*, 38(1):195–207.

Klenow, P. J. and Malin, B. A. (2010). Microeconomic evidence on price-setting. *Handbook of Monetary Economics*, 3:231 – 284.

Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.

Lahiri, K. and Sheng, X. (2010). Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics*, 25(4):514–538.

Lucas Jr, R. E. (1972). Expectations and the neutrality of money. *Journal of economic theory*, 4(2):103–124.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4):1295–1328.

Messina, J. D., Sinclair, T. M., and Stekler, H. (2015). What can we learn from revisions to the greenbook forecasts? *Journal of Macroeconomics*, 45:54–62.

Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of monetary Economics*, 54(7):2065–2082.

Nakamura, E. and Steinsson, J. (2013). Price rigidity: Microeconomic evidence and macroeconomic implications. *Annual Review of Economics*, 5(1):133–163.

Nordhaus, W. D. (1987). Forecasting efficiency: concepts and applications. *The Review of Economics and Statistics*, pages 667–674.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy. Technical report, National Bureau of Economic Research.

Zarnowitz, V. and Lambros, L. A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political economy*, 95(3):591–621.

A Derivations

Assume the standard Nordhaus/CG setup. That is

$$F_{t,t-h} = \gamma F_{t,t-h-1} + (1-\gamma)(A_t + \nu_{t,t-h}), \tag{19}$$

where γ is the weight put on old information. It is assumed, that $\nu_{t,t-h}$ and the forecast error of the old prediction ($\varepsilon_{t,t-h-1}$) are independent from each other and that forecasters set γ as specified in equation 8. This allows the regression specification

$$A_t - F_{t,t-h} = \alpha + \beta(F_{t,t-h} - F_{t,t-h-1}) + \varepsilon_t, \tag{20}$$

where under weak efficiency $\alpha = 0$ and $\beta = 0$. The estimator for the coefficient of interest then becomes

$$\hat{\beta} = \frac{cov[A_t - F_{t,t-h}, F_{t,t-h} - F_{t,t-h-1}]}{var[F_{t,t-h} - F_{t,t-h-1}]} \tag{21}$$

$$= \frac{cov[-\gamma\varepsilon_{t,t-h-1} - (1-\gamma)\nu_{t,t-h}, (\gamma-1)\varepsilon_{t,t-h-1} + (1-\gamma)\nu_{t,t-h}]}{var[(\gamma-1)\varepsilon_{t,t-h-1} + (1-\gamma)\nu_{t,t-h}]} \tag{22}$$

$$= \frac{cov[-\gamma\varepsilon_{t,t-h-1} - (1-\gamma)\nu_{t,t-h}, (1-\gamma)(\nu_{t,t-h} - \varepsilon_{t,t-h-1})]}{(1-\gamma)^2 var[(\nu_{t,t-h} - \varepsilon_{t,t-h-1})]} \tag{23}$$

$$= \frac{\gamma(1-\gamma)\sigma^2_{\varepsilon_{t,t-h-1}} - (1-\gamma)^2\sigma^2_{\nu_{t,t-h}}}{(1-\gamma)^2(\sigma^2_{\nu_{t,t-h}} + \sigma^2_{\varepsilon_{t,t-h-1}})} \tag{24}$$

$$= \frac{\frac{\gamma}{1-\gamma}\sigma^2_{\varepsilon_{t,t-h-1}} - \sigma^2_{\nu_{t,t-h}}}{\sigma^2_{\nu_{t,t-h}} + \sigma^2_{\varepsilon_{t,t-h-1}}} \tag{25}$$

$$=^* \frac{\lambda}{1-\lambda} \tag{26}$$

where $\sigma^2_{\varepsilon_{t,t-h-1}}$ is the variance of $\varepsilon_{t,t-h-1}$ and $\sigma^2_{\nu_{t,t-h}}$ is the variance of $\nu_{t,t-h}$.[17] $=^*$ only holds if the expected uncertainty is equal to the ex post uncertainty ($E\sigma^2_{\nu_{t,t-h}} = \sigma^2_{\nu_{t,t-h}}$ and $E\sigma^2_{\varepsilon_{t,t-h-1}} = \sigma^2_{\varepsilon_{t,t-h-1}}$) which might only hold as the number of observations goes to infinity. Alternatively, one can estimate

$$F_{t,t-h} - F_{t,t-h-1} = \alpha + \beta(F_{t,t-h-1} - F_{t,t-h-2}) + \varepsilon_t \tag{27}$$

[17]Due to the independence assumption between $\varepsilon_{t,t-h-1}$ and $\nu_{t,t-h}$, all the covariance terms disappear.

Again, the null assumes that $\alpha = 0$ and $\beta = 0$. In that case, the coefficient is more complicated, as it does not simplify to the ratio of the excess weights when the expected ex ante uncertainty is assumed equal to the ex post uncertainty. In addition, it depends on past excess weights λ_{t-1} as well and the subscripts are hence not omitted:

$$\hat{\beta}_t = \frac{cov[F_{t,t-h} - F_{t,t-h-1}, F_{t,t-h-1} - F_{t,t-h-2}]}{var[F_{t,t-h-1} - F_{t,t-h-2}]} \tag{28}$$

$$= \frac{cov[(\gamma_t - 1)\varepsilon_{t,t-h-1} + (1-\gamma_t)\nu_{t,t-h}, (\gamma_{t-1} - 1)\varepsilon_{t,t-h-2} + (1-\gamma_{t-1})\nu_{t,t-h-1}]}{var[(\gamma_{t-1} - 1)\varepsilon_{t,t-h-2} + (1-\gamma_{t-1})\nu_{t,t-h-1}]} \tag{29}$$

$$= \frac{cov[(\gamma_t - 1)(\gamma_{t-1}\varepsilon_{t,t-h-2} + (1-\gamma_{t-1})\nu_{t,t-h-1}), (1-\gamma_{t-1})(\nu_{t,t-h-1} - \varepsilon_{t,t-h-2})]}{(1-\gamma_{t-1})^2 var[(\nu_{t,t-h-1} - \varepsilon_{t,t-h-2})]} \tag{30}$$

$$= \frac{\gamma_{t-1}(1-\gamma_t)(1-\gamma_{t-1})\sigma^2_{\varepsilon_{t,t-h-2}} - (1-\gamma_t)(1-\gamma_{t-1})^2\sigma^2_{\nu_{t,t-h-1}}}{(1-\gamma_{t-1})^2(\sigma^2_{\nu_{t,t-h-1}} + \sigma^2_{\varepsilon_{t,t-h-2}})} \tag{31}$$

$$= (1-\gamma_t)\frac{\frac{\gamma_{t-1}}{1-\gamma_{t-1}}\sigma^2_{\varepsilon_{t,t-h-2}} - \sigma^2_{\nu_{t,t-h-1}}}{\sigma^2_{\nu_{t,t-h-1}} + \sigma^2_{\varepsilon_{t,t-h-2}}} \tag{32}$$

$$=^* (1-\gamma_t)\frac{\lambda_{t-1}}{1-\lambda_{t-1}} = \frac{(1-\lambda_t)\sigma^2_{\varepsilon_{t,t-h-1}}}{\sigma^2_{\nu_{t,t-h}} + \sigma^2_{\varepsilon_{t,t-h-1}}}\frac{\lambda_{t-1}}{1-\lambda_{t-1}} \tag{33}$$

Again, $=^*$ assumes that the expected uncertainty is equal to the ex post uncertainty. If $\lambda = 0$, both equation 25 or 32 will render the numerator equal to 0 and the estimate should be equal to 0 as well. If γ is larger than the optimal weight, the coefficient will become positive. This implies that agents put more than efficient weight on old information and under adjust to new information. If γ is smaller than the optimal weight, the coefficient will become negative. This implies that agents put less-than-efficient weight on old information and over-adjust to new information. Both tests will have the same sign, but the interpretation of the coefficient is different. Also, neither version allows allows the estimation of time varying coefficients without further assumptions (e.g. ex ante and ex post uncertainty is the same).

If it is assumed that forecasters put an additional weight of λ on the old prediction and 1-λ on the rational prediction and plugging in the optimal weights, equation 5 can be restated as

$$\hat{\eta}_{t,t-h} = (1-\lambda)\left[\frac{E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\varepsilon_{t,t-h-1}} + E\sigma^2_{\nu_{t,t-h}}}(\eta_t + \nu_{t,t-h}) \right. \\ \left. + \frac{E\sigma^2_{\nu_{t,t-h}}}{E\sigma^2_{\varepsilon_{t,t-h-1}} + E\sigma^2_{\nu_{t,t-h}}}\hat{\eta}_{t,t-h-1}\right] + \lambda\hat{\eta}_{t,t-h-1} \tag{34}$$

$$\hat{\eta}_{t,t-h} = \frac{(1-\lambda)E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\varepsilon_{t,t-h-1}} + E\sigma^2_{\nu_{t,t-h}}}(\eta_t + \nu_{t,t-h}) \\ + \frac{E\sigma^2_{\nu_{t,t-h}} + \lambda E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\varepsilon_{t,t-h-1}} + E\sigma^2_{\nu_{t,t-h}}}\hat{\eta}_{t,t-h-1} \tag{35}$$

Hence the weight on the old prediction becomes

$$\gamma = \frac{E\sigma^2_{\nu_{t,t-h}} + \lambda E\sigma^2_{\varepsilon_{t,t-h-1}}}{E\sigma^2_{\varepsilon_{t,t-h-1}} + E\sigma^2_{\nu_{t,t-h}}} \tag{36}$$

B Coefficients for Different Types of Agents

It was already shown above that the Greenbook forecasts and the SPF have a similar weight on their past predictions for output growth and inflation based on the deflator and different weights for the CPI inflation. This appendix tests, whether this is also the case for the three categories within the SPF. The three categories are financial services, non-financial services and unknown. As the Tables 7, 8 and 9 show, there is some variation across the three forecast categories financial industry, non-financial industry and unknown industry. With very few exceptions (e.g. unknown industry for the four quarter ahead GDP forecasts), the weight is similar to the overall coefficient.

C Coefficients for Different Horizons

From the model, the weight on old information depends on the noise in the new information. The further out into the future predictions are made, the less informative is the new information and the coefficient should get closer to unity.

Table 7: GDP Weight on Old Information by Industry

	Dependent variable:							
	Error Current Quarter							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Fin	Non-Fin	Unknown	All	Fin	Non-Fin	Unknown
Error 1Q	0.637***	0.665***	0.705***	0.530***				
	(0.032)	(0.062)	(0.045)	(0.071)				
Error 4Q					0.560***	0.617***	0.681***	0.435***
					(0.036)	(0.086)	(0.073)	(0.076)
Constant	0.216**	0.180*	0.243***	0.255**	0.287**	0.154	0.260**	0.351**
	(0.084)	(0.099)	(0.084)	(0.130)	(0.113)	(0.103)	(0.103)	(0.140)
Observations	202	116	116	107	194	113	113	104
R^2	0.726	0.641	0.740	0.400	0.623	0.603	0.707	0.277
Adjusted R^2	0.725	0.638	0.738	0.395	0.621	0.599	0.705	0.270

This table reports the coefficient and Newey-West standard errors. *, ** and *** imply significantly different from 0 at 10%, 5% and 1% level, respectively.

Table 8: CPI Weight on Old Information by Industry

	Dependent variable:							
	Error Current Quarter							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Fin	Non-Fin	Unknown	All	Fin	Non-Fin	Unknown
Error 1Q	0.608***	0.498***	0.655***	0.665***				
	(0.028)	(0.034)	(0.035)	(0.031)				
Error 4Q					0.535***	0.487***	0.675***	0.628***
					(0.028)	(0.037)	(0.033)	(0.050)
Constant	0.016	0.042	−0.010	0.051	0.134***	0.149***	0.135***	0.184**
	(0.027)	(0.043)	(0.036)	(0.060)	(0.039)	(0.042)	(0.029)	(0.073)
Observations	153	118	118	109	150	115	115	106
R^2	0.872	0.754	0.897	0.802	0.826	0.771	0.925	0.754
Adjusted R^2	0.871	0.752	0.896	0.801	0.825	0.769	0.924	0.752

This table reports the coefficient and Newey-West standard errors. *, ** and *** imply significantly different from 0 at 10%, 5% and 1% level, respectively.

Table 9: Deflator Weight on Old Information by Industry

	Dependent variable:							
	Error Current Quarter							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Fin	Non-Fin	Unknown	All	Fin	Non-Fin	Unknown
Error 1Q	0.633***	0.807***	0.919***	0.780***				
	(0.030)	(0.049)	(0.033)	(0.082)				
Error 4Q					0.617***	0.695***	0.829***	0.547***
					(0.032)	(0.041)	(0.038)	(0.137)
Constant	−0.090	−0.024	0.011	0.182*	−0.118*	0.063	0.079**	0.178*
	(0.064)	(0.052)	(0.028)	(0.094)	(0.067)	(0.040)	(0.037)	(0.098)
Observations	202	116	116	107	194	113	113	104
R^2	0.686	0.761	0.905	0.564	0.657	0.710	0.832	0.352
Adjusted R^2	0.684	0.759	0.904	0.560	0.655	0.707	0.831	0.346

This table reports the coefficient and Newey-West standard errors. *, ** and *** imply significantly different from 0 at 10%, 5% and 1% level, respectively.

Table 10: GDP Regressions for Different Horizons

	Dependent variable:					
	Error 1Q		Error 2Q		Error 3Q	
	(1)	(2)	(3)	(4)	(5)	(6)
	GB	SPF	GB	SPF	GB	SPF
Error 2Q	0.893***	0.666***				
	(0.025)	(0.035)				
Error 3Q			0.915***	0.677***		
			(0.019)	(0.059)		
Error 4Q					1.028***	1.062***
					(0.023)	(0.046)
Constant	0.086	−0.365***	−0.095	0.617**	0.003	−0.207*
	(0.076)	(0.106)	(0.101)	(0.279)	(0.027)	(0.109)
Observations	174	200	165	199	154	194
R^2	0.877	0.627	0.899	0.460	0.970	0.859
Adjusted R^2	0.876	0.626	0.898	0.457	0.970	0.858

This table reports the coefficient and Newey-West standard errors. *, ** and *** imply significantly different from 0 at 10%, 5% and 1% level, respectively.

Table 11: Inflation Regressions for Different Horizons

	Dependent variable:											
	CPI						GDP Deflator					
	Error 1Q		Error 2Q		Error 3Q		Error 1Q		Error 2Q		Error 3Q	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	GB	SPF	GB	SPF	GB	SPF	GB	SPF	GB	SPF	GB	SPF
Error 2Q	0.847***	0.967***					0.712***	0.749***				
	(0.035)	(0.007)					(0.018)	(0.014)				
Error 3Q			0.955***	0.978***					0.879***	0.969***		
			(0.011)	(0.010)					(0.081)	(0.065)		
Error 4Q					0.975***	0.970***					0.993***	0.975***
					(0.015)	(0.017)					(0.017)	(0.027)
Constant	−0.118	0.086***	0.021	0.040**	0.075**	0.050***	0.021	0.041	−0.008	−0.061	0.010	−0.027
	(0.108)	(0.020)	(0.061)	(0.017)	(0.032)	(0.014)	(0.063)	(0.045)	(0.074)	(0.067)	(0.029)	(0.030)
Observations	134	152	133	151	132	150	174	200	165	199	154	194
R^2	0.777	0.981	0.851	0.983	0.978	0.988	0.626	0.727	0.682	0.730	0.947	0.933
Adjusted R^2	0.776	0.981	0.850	0.983	0.978	0.988	0.624	0.726	0.680	0.729	0.947	0.932

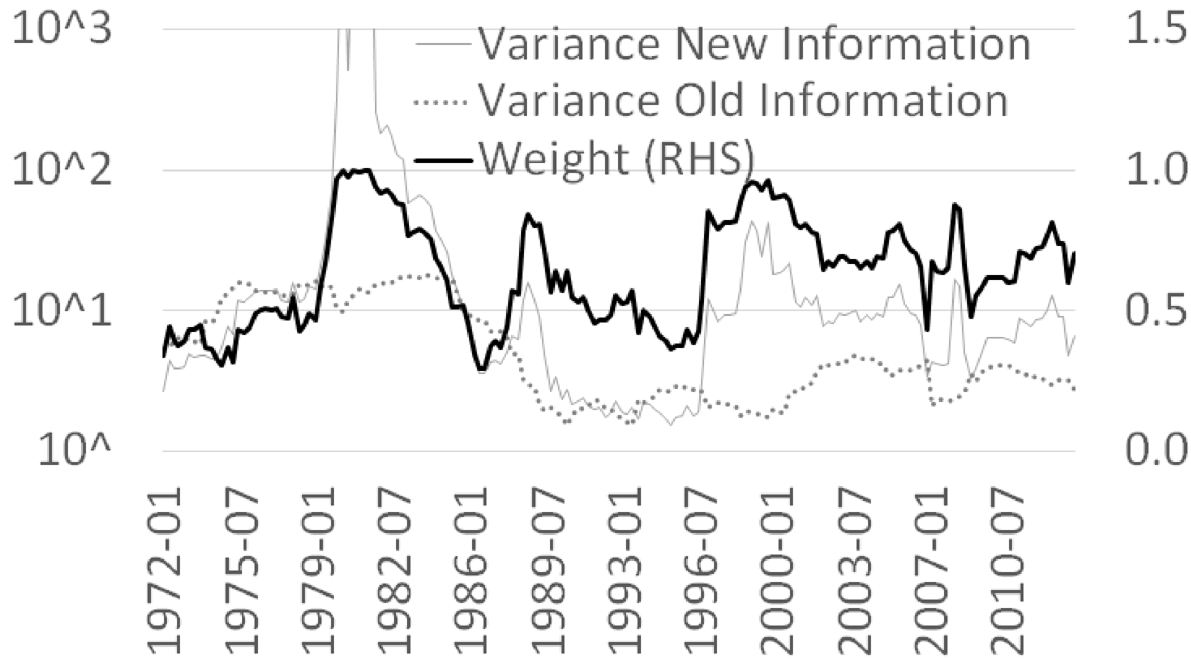
This table reports the coefficient and Newey-West standard errors. *, ** and *** imply significantly different from 0 at 10%, 5% and 1% level, respectively.

Indeed, both Tables 10 and 11 show that as the horizon increase, the coefficient increases as well, meaning that the weight on old information increases as well. Indeed, for longer horizons, the coefficient is not significantly different from unity anymore and hence the case of no signal cannot be rejected anymore.

D Uncertainty of New Information

As mentioned in the section on the ex post optimal γ (4.1), the estimation approach allows to extract the variance of the new information $\sigma^2_{\nu_{t,t-h}}$. In order to obtain a time series of this variance, the same 20 quarter rolling window regressions as in the section on the variation over time (5) is utilized. Figure 7 shows the uncertainty of new information against the uncertainty of the old prediction on a logarithmic scale for the one quarter ahead GDP predictions in the Greenbook. In addition, the graph has the weight on the previous prediction on the right scale for comparison.

Figure 7: Uncertainty vs. Weight for One Quarter GDP Predictions (GB)



The graph shows that the uncertainty of new information fluctuates differently from the uncertainty of the past prediction. Specifically, the two appear to move somewhat in opposite directions (which can also be seen when comparing them to the weight places on the past prediction). In addition, the variance of new information show an exceptionally uncertain period in the late 70s and early 80s.