Accepted: 12 December 2025
DOI: 10.1111/ecca.70026

ORIGINAL ARTICLE

Economica LSE

Adaptive expectations and reaction to information

Junyi Liao

University of Essex

Correspondence
Dr Junyi Liao, Essex Business School, University of Essex, Wivenhoe Park, Colchester CO4 3SQ
Email: junyi.liao.sam@gmail.com

Abstract
This paper develops a model combining adaptive expectations with noisy signals, and derives three coefficients and one impulse response function (IRF): the Coibion–Gorodnichenko (CG) coefficient capturing consensus under-reaction to information, the Bordalo–Gennaioli–Ma–Shleifer coefficient capturing individual over-reaction, the Kohlhas–Walther coefficient capturing extrapolation, and the Angeletos–Huo–Sastry IRF capturing delayed overshooting. There exists a parameter region in which the model reconciles all four moments with the data simultaneously. The model also delivers a testable prediction linking the CG coefficient to variable persistence, distinguishing adaptive expectations from Kalman-filter updating, and I present supporting evidence for adaptive expectations. The model's fit to survey data is evaluated.

1 | INTRODUCTION

The rational expectations hypothesis has long dominated the modelling of macroeconomic expectations. Yet recent survey evidence consistently rejects its validity: forecast errors are predictable across a wide range of surveys (Mankiw *et al.* 2003; Coibion and Gorodnichenko 2012, 2015; Adam *et al.* 2017; Fuhrer 2018; Bordalo *et al.* 2020; Kohlhas and Walther 2021; Angeletos *et al.* 2021; Wang 2021; d'Arienzo 2020; de Silva and Thesmar 2023; Kucinskas and Peters 2024). This empirical challenge has spurred a wave of alternative models designed to account for systematic deviations from rational expectations, including diagnostic expectations (Bordalo *et al.* 2018, 2019, 2020), extrapolative or behavioural models (Gennaioli *et al.* 2016; Angeletos *et al.* 2021; Reis 2020), cognitive discounting, and level-k thinking (Farhi and Werning 2019; Gabaix 2020).

Before the rational expectations revolution, adaptive expectations provided the standard framework for modelling expectations (Cagan 1956; Friedman 1957). In this framework, forecasts are formed as a weighted average of past forecasts and current observations, with the weight on new information termed the adjustment parameter. While widely used in the 1950s–1970s, adaptive expectations was criticized for its inability to adapt to regime shifts—Lucas (1976) emphasized that expectations should adjust immediately to policy changes. As a result, adaptive expectations was displaced by rational expectations during the rational expectations revolution.

This paper revisits adaptive expectations. I show that once combined with noisy signals, this simple framework can account for four influential empirical findings: (a) consensus forecasts under-react to new information (the Coibion–Gorodnichenko (CG) coefficient is positive); (b) individual forecasts over-react (the Bordalo–Gennaioli–Ma–Shleifer (BGMS) coefficient is negative); (c) forecasts extrapolate from recent observations (the Kohlhas–Walther (KW) coefficient is negative); and (d) forecasts display delayed overshooting relative to outcomes (the Angeletos–Huo–Sastry (AHS) impulse response function (IRF)). The model yields closed-form expressions for these moments and simple parameter conditions under which all four can be reconciled simultaneously. Moreover, adaptive expectations and the canonical noisy information model make opposite predictions about how the CG coefficient relates to persistence. Using US survey data, I find evidence in favour of adaptive expectations.

The model features a continuum of forecasters who receive noisy signals about a fundamental variable, assumed to follow an AR(1) process. Rather than applying a Kalman filter, forecasters update adaptively, weighting their signal and past forecast. Since the benchmark adaptive expectations does not generate a term structure of forecasts, I impose an iterated law of forecasts assumption—analogous to the iterated law of expectations under rational expectations—which yields a flat term structure. At the consensus level, idiosyncratic noise cancels out, so the consensus forecast is equivalent to that of a representative agent who observes the true variable with adaptive expectations updating.

As a first step, the model implies a division between moments driven by adaptive updating and those shaped by noise. Specifically, the consensus-level moments—the CG and KW coefficients, and the AHS IRF—depend only on the adaptive expectations mechanism, since individual noise cancels out in the aggregate. By contrast, the BGMS coefficient is jointly determined by adaptive expectations and the noisy information component.

Fixing persistence, there exists a threshold for the adjustment parameter below which forecasts exhibit under-reaction, yielding a positive CG coefficient: the less forecasts adjust to current observations, the stronger the under-reaction. Similarly, there exists a threshold above which forecasts extrapolate from recent events, producing a negative KW coefficient: the more forecasts tilt towards current observations, the greater the extrapolation. Crucially, the extrapolation threshold is lower than the under-reaction threshold, implying that under adaptive expectations there exists a parameter region where under-reaction (positive CG coefficient) and extrapolation (negative KW coefficient) coexist—a result that the literature had considered unattainable in simple models (Kohlhas and Walther 2021).

Adaptive expectations yield a simple condition for the AHS pattern of delayed overshooting: the adjustment parameter must be smaller than the persistence of the underlying variable. When this condition holds, forecasts initially lag behind actual outcomes, then overshoot as they continue adjusting. If instead the adjustment parameter exceeds persistence, then forecasts overshoot immediately after a shock, with no initial under-reaction. In the long run, forecasts always overshoot because they are backward-looking on the initial spike of the outcome.

Combining this condition with the thresholds derived above shows that the observed signs of the CG and KW coefficients and the AHS IRF can be jointly reconciled under adaptive expectations by a single inequality linking the adjustment parameter to persistence. By contrast, the BGMS coefficient arises from the noisy information element. In the BGMS regression, the current forecast enters both sides with opposite signs, and noise in individual forecasts pushes the coefficient negative. Since this noise cancels out in the aggregate, it affects only BGMS and not CG. If noise were absent, then the CG and BGMS coefficients would coincide.

I then conduct a quantitative estimation exercise to assess the model's fit to survey data, using Consumer Price Index (CPI) and gross domestic product (GDP) price index forecasts from the Philadelphia Fed's Survey of Professional Forecasters. First, I estimate the AR(1) process for inflation. Second, given the persistence estimates, I back out the adjustment parameter by matching either the CG or KW coefficient between model and data. Third, I estimate the variance of noisy signals by matching the variance of individual forecast revisions. A good fit requires the adjustment parameters obtained from CG and KW coefficients to overlap, and the model-implied BGMS coefficient and AHS IRF to align with their empirical counterparts. The results show that the model performs reasonably well, though not perfectly: for the CPI, it nearly always reconciles all four moments simultaneously; for the GDP Price Index, it fits CG and BGMS coefficients, and AHS IRF, but not when including the KW coefficient.

In further analysis, I first follow Kohlhas and Walther (2021) by decomposing inflation into different components, each assigned distinct adjustment parameters (i.e. different degrees of attention). Although reconciling the CG and KW coefficients under adaptive expectations does not require such decomposition, this exercise reinforces the intuition in Kohlhas and Walther (2021): forecasters should assign higher adjustment parameters to more procyclical components. Second, to address the concern that my adjustment parameter estimates may be driven by the specific targeted moments, I re-estimate the parameter using a model-implied regression that does not target any moment. The resulting estimates are somewhat smaller but remain of the same order of magnitude as the benchmark estimates.

Finally, I test the model prediction on which adaptive expectations and rational noisy-information models diverge: the relationship between the CG coefficient and persistence. Under adaptive expectations, for a given adjustment parameter, greater persistence implies stronger under-reaction and thus a larger CG coefficient. In contrast, under the canonical noisy-information model, greater persistence leads agents to place more weight on current signals, implying a smaller CG coefficient. This difference arises because adaptive expectations takes the adjustment parameter as exogenous, whereas in the noisy information model, the Kalman gain depends endogenously on persistence. Using rolling window estimates of the CG coefficient and persistence, I find time series evidence in support of adaptive expectations, in contrast to the cross-sectional evidence favouring noisy information in Coibion and Gorodnichenko (2015).

The contribution of this paper is both theoretical and empirical.

Theoretically, the paper derives closed-form solutions for four moments of central interest, and establishes parameter conditions under which all moments can be reconciled simultaneously. It demonstrates that adaptive expectations can jointly match the CG and KW coefficients—previously thought unattainable by simple extrapolation models (Kohlhas and Walther 2021)—and shows that adaptive expectations alone can reconcile the three consensus-level moments (CG, KW and AHS IRF), which under Kalman-filter updating require additional modelling elements (Angeletos *et al.* 2021). Finally, it delivers a testable prediction linking the CG coefficient to persistence, on which adaptive expectations and Kalman-filter updating diverge.

Empirically, the paper provides new estimates of the adjustment parameter, directly related to the learning rate in the constant gain learning literature (Marcet and Nicolini 2003; Orphanides and Williams 2005, 2006; Branch and Evans 2006; Milani 2007, 2008; Eusepi and Preston 2011). It also presents time series evidence favouring adaptive expectations over rational Kalman-filter updating, complementing the cross-sectional evidence in support of noisy information in Coibion and Gorodnichenko (2015).

The structure of the paper is as follows. Section 2 provides a brief introduction to adaptive expectations. Section 3 presents the four moments of interest from the literature, develops the model framework, and derives the analytical results. Section 4 reports the estimation exercises, and evaluates the model's fit to survey data. Section 5 discusses three additional exercises, with particular emphasis on testing the model's key prediction. Section 6 concludes.

2 | AN INTRODUCTION TO ADAPTIVE EXPECTATIONS

The adaptive expectations hypothesis dates back to the 1930s and was formally introduced in the 1950s (see Cagan 1956; Friedman 1957). It became widely used in macroeconomics during the 1960s and 1970s. A common application was the modelling of inflation expectations as adaptive expectations. Let w_t denote the variable of interest, and let $\mathcal{F}_t w_{t+1}$ denote the period-t forecast of w_{t+1}. Under adaptive expectations, the forecast is given by

$$\mathcal{F}_t w_{t+1} = \gamma w_t + (1-\gamma)\mathcal{F}_{t-1} w_t, \tag{1}$$

where the forecast of w_{t+1} formed in period t, $\mathcal{F}_t w_{t+1}$, is a weighted average of the current realization w_t and the previous forecast $\mathcal{F}_{t-1} w_t$. The adjustment parameter γ determines the weight placed on the most recent observation. Under adaptive expectations, agents form forecasts using current information and past expectations. Because past expectations themselves are based on earlier realizations, current expectations ultimately reflect the entire history of observations. This relationship can be expressed as an iterated sum:

$$\mathcal{F}_t w_{t+1} = \gamma \sum_{i=0}^{\infty} (1-\gamma)^i w_{t-1-i}. \tag{2}$$

Mathematically, the adjustment parameter γ must lie in $(0, 2)$ to ensure that $\mathcal{F}_t w_{t+1}$ does not diverge. When γ is between 1 and 2, the model generates a form of 'extrapolation'. According to equation (2), the current expectation is a weighted average of past observations, with progressively smaller weights placed on more distant observations.

The noisy information model is an important framework in the study of expectations formation. In these models, agents update forecasts using the Kalman filter, which resembles equation (1). However, two key differences arise. First, the Kalman filter implies rational updating: agents recognize how forecasts relate across horizons. By contrast, adaptive expectations make no assumption about the term structure, updating only forecasts of the same horizon across periods. Second, in the Kalman filter, the update rate is determined endogenously by the noise-to-signal ratio and the persistence of the underlying variable. In adaptive expectations, the adjustment parameter γ is specified exogenously. In Subsection 5.3, I show that the rational noisy information model and the adaptive expectations model generate opposite predictions, which can be tested in the data.

The backward-looking feature of adaptive expectations may partially capture how people form expectations, but it is subject to the famous Lucas critique (Lucas 1976). People's expectations should adapt to policy changes. When considering the future, humans are far more sophisticated than the simple rule described above. This critique led to the rational expectations revolution, initiated by Muth (1961). Under rational expectations, agents fully comprehend the environment in which they operate, effectively acting as the smartest economists.

Over the past two decades, rational expectations has been extensively challenged in the literature, from various perspectives. Numerous biases in people's expectations formation have been documented (Mankiw *et al.* 2003; Coibion and Gorodnichenko 2015; Bordalo *et al.* 2020; Kohlhas and Walther 2021; Giglio *et al.* 2021; Angeletos *et al.* 2021). In response, many alternative models of expectations formation have emerged, deviating from rational expectations to better align with empirical findings from forecast surveys. The literature also shows that these models help to explain other empirical puzzles beyond forecast surveys.

3 | THEORETICAL INSIGHTS FROM ADAPTIVE EXPECTATIONS

In this section, I show how adaptive expectations help to explain a set of recent empirical findings from forecast surveys. Subsection 3.1 introduces three regression coefficients and an IRF that capture how forecasters react to new information. Subsection 3.2 then develops a framework that combines noisy information with adaptive expectations, derives closed-form expressions for the coefficients and the IRF, and identifies the parameter conditions under which the model can match the data. Finally, I compare this framework with the existing literature.

3.1 | Three regression coefficients and one IRF

In a seminal contribution, Coibion and Gorodnichenko (2015) propose a regression of consensus forecast errors on consensus forecast revisions to test for aggregate over-reaction or under-reaction to new information in survey forecasts. Building on this approach, Bordalo *et al.* (2020) evaluate the same regression at the level of individual forecasts. Let w_t denote a macroeconomic variable, such as inflation. The following two equations present the 'forecast errors on forecast revisions' regressions at the consensus and individual levels, henceforth referred to as the CG and BGMS regressions:

$$w_{t+h} - \overline{\mathcal{F}}_t w_{t+h} = \beta_0 + \beta_{CG}(\overline{\mathcal{F}}_t w_{t+h} - \overline{\mathcal{F}}_{t-1} w_{t+h}) + u_{t,t+h}, \tag{3}$$

$$w_{t+h} - \mathcal{F}_{i,t} w_{t+h} = \beta_0 + \beta_{BGMS}(\mathcal{F}_{i,t} w_{t+h} - \mathcal{F}_{i,t-1} w_{t+h}) + u_{i,t,t+h}. \tag{4}$$

A key distinction between the two regressions is that in equation (3), $\overline{\mathcal{F}}_t w_{t+h}$ denotes the mean forecast across all forecasters, while in equation (4), $\mathcal{F}_{i,t} w_{t+h}$ denotes an individual forecast. In the literature, a positive β_{CG} is interpreted as under-reaction: when revisions occur, forecast revisions are too small relative to actual realizations. Conversely, a negative β_{CG} indicates over-reaction: revisions are excessively large relative to realizations. The same interpretation applies at the individual level, with $\beta_{BGMS} > 0$ corresponding to under-reaction, and $\beta_{BGMS} < 0$ to over-reaction. Empirical studies typically find $\beta_{CG} > 0$ and $\beta_{BGMS} < 0$ (Coibion and Gorodnichenko 2015; Bordalo *et al.* 2020; Fuhrer 2018).

In a closely related article, Kohlhas and Walther (2021) show that macroeconomic forecasts both extrapolate from recent events and under-react to new information. The evidence for under-reaction comes from the same regression in Coibion and Gorodnichenko (2015), namely $\beta_{CG} > 0$. Extrapolation is identified through the regression

$$w_{t+h} - \overline{\mathcal{F}}_t w_{t+h} = \beta_0 + \beta_{KW} w_t + u_{t,t+h}.$$

On the left-hand side, the dependent variable is the consensus forecast error, as in the CG regression. On the right-hand side, the explanatory variable is the current realization w_t. Kohlhas and Walther (2021) find $\beta_{KW} < 0$. In this case, an increase in w_t is associated with negative forecast errors, indicating that forecasts place too much weight on the rise in w_t—that is, they extrapolate.

Angeletos *et al.* (2021) move beyond linear regressions to characterize forecast errors in a more dynamic setting. They study responses to two key shocks: one that drives most business-cycle variation in unemployment and other macroeconomic variables, and another that primarily explains business-cycle variation in inflation. They find that the IRFs of average unemployment and inflation forecasts initially display under-reaction, but subsequently overshoot relative to actual realizations. They formally define the IRF of forecast errors as $\{\zeta_k\}_{k=1}^{+\infty}$ in Definition 1.

Definition 1. Let $\{\zeta_k\}_{k=1}^{+\infty}$ be the IRF of the average, one-step-ahead forecast error. For all $k \geq 1$,

$$\zeta_k \equiv \frac{\partial(w_{t+k} - \overline{\mathcal{F}}_{t+k-1} w_{t+k})}{\partial e_t}$$

is the kth coefficient in the moving-average representation of the average forecast error, where e_t is the shock at period t.

Forecasts exhibit initial under-reaction and subsequent overshoot if and only if for some $k_{IRF} \in (1, +\infty)$, $\zeta_k > 0$ for $1 \leq k < k_{IRF}$, and $\zeta_k < 0$ for $k_{IRF} < k$.

As shown in Definition 1, the initial under-reaction and subsequent overshoot of forecasts relative to actual realizations correspond to an IRF $\{\zeta_k\}_{k=1}^{+\infty}$ of forecast errors that starts positive and later turns negative.

In the next subsection, I develop a framework that combines noisy information with adaptive expectations, and derive analytical expressions for three coefficients and one IRF. I then introduce the parameter conditions under which the signs of these coefficients can be reconciled simultaneously.

3.2 | A model of adaptive expectations and noisy signals

3.2.1 | The model environment

To compute forecast revisions in the CG regressions, forecasts at different horizons, $\overline{\mathcal{F}}_t w_{t+h}$ and $\overline{\mathcal{F}}_{t-1} w_{t+h}$, are needed. Adaptive expectations, however, do not impose an explicit structure across horizons. This limitation has kept the literature from assessing the ability of adaptive expectations to account for the empirical patterns documented in the CG and BGMS regressions (see Afrouzi *et al.* 2023).

With a lenient assumption—the 'law of iterated forecasts' in Assumption 1 below—it is possible to derive a term structure for adaptive expectations. The law of iterated forecasts mirrors the form of the law of iterated expectations, which holds under rational expectations. In the adaptive expectations framework, agents update beliefs using past data. The law of iterated forecasts captures the idea that if forecasts are formed consistently under the same rule (e.g. a weighted average of past outcomes), then forecasts at different horizons should be internally consistent.

Assumption 1. Assume the following form of the law of iterated forecasts:

$$\mathcal{F}_t w_{t+h} = \mathcal{F}_t(\mathcal{F}_{t+s} w_{t+h}),$$

for integers $0 \leq s \leq h$.

With Assumption 1 and when forecasts follow adaptive expectations, the term structure of forecasts is as follows.

Lemma 1. *The term structure of forecasts under adaptive expectations is flat:*

$$\mathcal{F}_t w_{t+1} = \mathcal{F}_t w_{t+i}, \quad i \geq 2.$$

The proof of Lemma 1 is provided in the Appendix. Lemma 1 shows that adaptive expectations yield identical forecasts across horizons. This result is unsurprising: if forecasts are always formed as weighted averages of past observations, and different horizons do not assign different weights to these observations, then forecasts must coincide across horizons.

After deriving the term structure of forecasts, assumptions are needed about the process governing the underlying variable. Following Coibion and Gorodnichenko (2015), Bordalo *et al.* (2020), Angeletos *et al.* (2021) and Kohlhas and Walther (2021), I assume that the variable follows an AR(1) process.

Assumption 2. The underlying variable follows an AR(1) process:

$$w_t = \rho w_{t-1} + e_t,$$

where e_t is independent and identically distributed, and follows a normal distribution, $e_t \sim N(0, \sigma_e^2)$, with $0 < \rho < 1$.

The AR(1) assumption is standard in the literature. Nonetheless, Subsection A.0.2 of the Appendix shows with a numerical example that the ability of the model to explain the empirical findings does not rely on this assumption. The result also holds under an AR(2) process.

To generate dispersed individual forecasts, I assume that forecasters disagree because they receive heterogeneous signals (Sims 2003; Woodford 2001). Each forecaster then forms expectations adaptively, weighting both observed signals and past forecasts.

Assumption 3. There is a continuum of forecasters. In each period, each forecaster receives a noisy signal about the underlying variable:

$$w_{i,t} = w_t + \varepsilon_{i,t},$$

where $\varepsilon_{i,t}$ follows a normal distribution $\varepsilon_{i,t} \sim N(0, \sigma_\varepsilon^2)$, and is independent and identically distributed across forecasters and time. Forecasters form forecasts using adaptive expectations:

$$\mathcal{F}_{i,t} w_{t+h} = \gamma w_{i,t} + (1-\gamma)\mathcal{F}_{i,t-1} w_{t+h-1}, \tag{5}$$

where $0 < \gamma < 2$.

Allowing γ to lie in $(1, 2)$—beyond the usual $(0, 1)$ range—introduces a form of over-extrapolation. However, as shown analytically below, reconciling the empirical patterns requires γ to fall within $(0, 1)$.

Under the benchmark noisy information framework (Sims 2003; Woodford 2001), forecasters update rationally using the Kalman filter, where the update speed depends on the noise-to-signal ratio and the persistence of the variable ρ. By contrast, in the naive adaptive expectations model, the adjustment parameter γ is specified exogenously. Subsection 5.3 shows that the two frameworks yield opposite predictions for the relationship between β_{CG} and persistence ρ, which can be tested in the data.

By definition, the consensus forecast is given by

$$\overline{\mathcal{F}}_t w_{t+h} = \lim_{N\to+\infty} \frac{\sum_{i=1}^{N} \mathcal{F}_{i,t} w_{t+h}}{N}.$$

The number of forecasters is large enough that noisy individual signals get averaged out for each t, that is,

$$\lim_{N\to+\infty} \frac{\sum_{i=1}^{N} \varepsilon_{i,t}}{N} = 0.$$

As a result, when N is large enough,

$$\overline{\mathcal{F}}_t w_{t+h} = \gamma w_t + (1-\gamma)\overline{\mathcal{F}}_{t-1} w_{t+h-1}. \tag{6}$$

Individual noisy signals affect individual forecasts. In contrast, consensus forecasts are unaffected because averaging over forecasters eliminates the noise.

3.2.2 | Model-implied coefficients and IRF

In the environment described above, the three regression coefficients can be derived analytically, as stated in Proposition 1. The proof is provided in the Appendix.

Proposition 1. *Under Assumptions 1, 2 and 3, coefficients β_{CG}, β_{BGMS} and β_{KW} are given as follows:*

$$\beta_{CG} = \rho^h \left[\frac{1}{\gamma} - \frac{1}{2}\right] - \frac{1}{2}, \tag{7}$$

$$\beta_{BGMS} = \frac{\rho^h\left((2/\gamma) - 1\right) - 1 - \kappa\sigma_\varepsilon^2/\sigma_e^2}{2 + 2\kappa\sigma_\varepsilon^2/\sigma_e^2}, \tag{8}$$

$$\beta_{KW} = \rho^h - \frac{\gamma}{1-(1-\gamma)\rho}, \tag{9}$$

where $\kappa = [1-\rho(1-\gamma)](1+\rho) > 0$.

Fixing the forecast horizon h, β_{CG} and β_{KW} are jointly determined by persistence ρ and the adjustment parameter γ. Importantly, because individual signals average out at the consensus level, the noisy information component does not affect β_{CG} or β_{KW}. These coefficients are determined solely by the benchmark adaptive expectations model.

The noisy information component plays an important role in determining β_{BGMS}. The noise-to-signal ratio $\sigma_\varepsilon^2/\sigma_e^2$ governs the gap between β_{CG} and β_{BGMS}. Before turning to discussing each coefficient in detail, Corollary 1 summarizes several key relationships between coefficients and parameters. The proof is provided in the Appendix.

Corollary 1. *For* $\rho \in (0,1)$ *and* $\gamma \in (0,2)$, *the following inequalities hold:*

$$\frac{\partial\beta_{CG}}{\partial\gamma} < 0, \quad \frac{\partial\beta_{KW}}{\partial\gamma} < 0,$$

$$\frac{\partial\beta_{BGMS}}{\partial(\sigma_\varepsilon^2/\sigma_e^2)} < 0,$$

$$\frac{\partial\beta_{CG}}{\partial\rho} > 0.$$

Here, β_{CG} increases with ρ, and decreases with γ. Holding γ fixed, a higher ρ implies greater under-reaction: shocks to the variable persist further into the future than adaptive expectations incorporate. Holding ρ fixed, a higher γ implies greater over-reaction: adaptive expectations update forecasts too strongly relative to the true persistence of the variable. The model-implied relationship between β_{CG} and ρ is the opposite of the prediction from the rational noisy information model in Coibion and Gorodnichenko (2015), a point developed further in Subsection 5.3 and tested formally.

Also, β_{KW} decreases with γ. A larger adjustment parameter implies stronger extrapolation. As shown in the corollary below, there exists a range of γ for which under-reaction to new information ($\beta_{CG} > 0$) and extrapolation from recent events ($\beta_{KW} < 0$) hold simultaneously.

The ratio $\sigma_\varepsilon^2/\sigma_e^2$ enters the numerator of β_{BGMS} with a negative sign, pushing β_{BGMS} downwards. When the noisy signal vanishes, β_{CG} and β_{BGMS} coincide. This highlights the mechanical source of their difference: in the CG and BGMS regressions, current forecasts appear on both

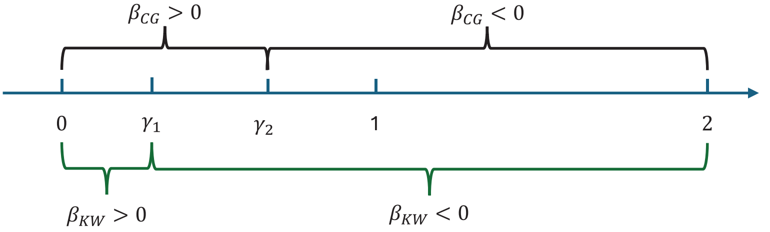


FIGURE 1 β_{CG} and β_{KW} over the domain of γ. *Notes*: This figure shows how β_{CG} and β_{KW} change sign over the domain of γ. There exists an interval $[\gamma_1, \gamma_2]$ in which $\beta_{CG} > 0$ and $\beta_{KW} < 0$ hold simultaneously, where $\gamma_1 = \rho^h(1-\rho)/(1-\rho^{h+1})$ and $\gamma_2 = 2\rho^h/(1+\rho^h)$.

sides of the equation but with opposite signs. At the individual level, idiosyncratic signals push the coefficient downwards, while at the consensus level, these signals average out and have no effect. This point has been noted in the literature (de Silva and Thesmar 2023; Liao 2024). Liao (2024) shows that in a parsimonious model of expectations formation, the gap between β_{CG} and β_{BGMS} can be reconciled quantitatively when forecast disagreement is modelled as individual heterogeneity and general measurement error.

For a given h, there exists a range of values for γ, ρ and $\sigma_\varepsilon^2/\sigma_e^2$ such that $\beta_{CG} > 0$, $\beta_{BGMS} < 0$ and $\beta_{KW} < 0$ hold simultaneously, as stated in Corollary 2. The proof is provided in the Appendix.

Corollary 2. *$\beta_{CG} > 0$, $\beta_{BGMS} < 0$ and $\beta_{KW} < 0$ hold simultaneously if and only if*

$$\frac{\rho^h(1-\rho)}{1-\rho^{h+1}} < \gamma < \frac{2\rho^h}{1+\rho^h} < 1,$$

$$\frac{\rho^h\left((2/\gamma)-1\right)-1}{\kappa} < \frac{\sigma_\varepsilon^2}{\sigma_e^2}.$$

According to the corollary, the noise-to-signal ratio must be sufficiently large for $\beta_{BGMS} < 0$. The adjustment parameter γ must also lie within a range where under-reaction to new information, in the sense of Coibion and Gorodnichenko (2015), overlaps with extrapolation from recent events, in the sense of Kohlhas and Walther (2021), as illustrated in Figure 1.

After analysing the three regression coefficients under the model framework, I now turn to the IRF. The proposition below states the parameter conditions under which forecasts initially under-react and later overshoot.

Proposition 2. *Under Assumptions 2 and 3, ζ_k, as defined in Definition 1, is given by*

$$\zeta_k = \frac{(\rho-1)\rho^k + \gamma(1-\gamma)^k}{\rho+\gamma-1}.$$

The necessary and sufficient condition for forecasts to display initial under-reaction followed by overshooting is

$$\gamma \in (0, \rho).$$

The condition for initial under-reaction followed by overshooting is intuitive: the adjustment parameter must be smaller than ρ. When $\gamma > \rho$, forecasters revise too aggressively relative to the decay of the initial shock, so no initial under-reaction occurs. Under adaptive expectations, however, forecasts always overshoot in the long run. This arises from the backward-looking nature of the model: forecasts are based on the earlier path of higher realizations. Moreover, the smaller γ, the longer it takes for overshooting to appear.

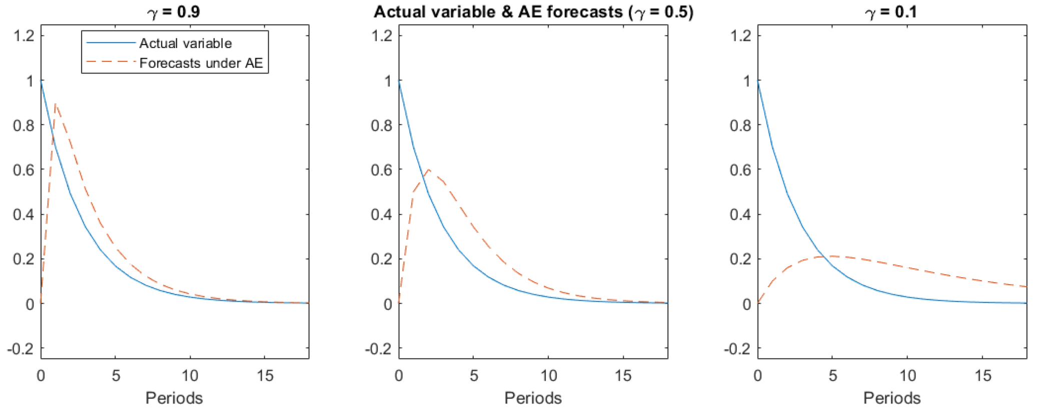


FIGURE 2 IRF of the variable and forecasts under adaptive expectations. *Notes*: This figure shows the IRFs of the variable and its forecasts under adaptive expectations to an unexpected shock at period 0. I set $\rho = 0.7$. From left to right, γ is set to 0.9, 0.5 and 0.1.

Because the IRF is defined on the consensus forecast, the model's ability to reconcile the documented delayed overshooting of forecasts stems from the adaptive expectations component, unrelated to the noisy information part.

Figure 2 presents a numerical example of the IRFs of both the variable and the forecasts in response to an unexpected shock at period 0. I set $\rho = 0.7$ and illustrate three cases with $\gamma = 0.9, 0.5, 0.1$. Consistent with Proposition 2, when $\gamma = 0.9 > \rho$, forecasts overshoot immediately from $t = 1$, with no initial under-reaction. When $\rho > \gamma = 0.5$ or 0.1, forecasts initially under-react and later overshoot. The smaller γ, the longer the period of under-reaction before overshooting occurs.

For the three sets of empirical findings on consensus forecasts to hold, the restrictions on ρ and γ are summarized in the following corollary.

Corollary 3. $\beta_{CG} > 0$, $\beta_{KW} < 0$, *and the IRF of forecast errors exhibits initial under-reaction and subsequent overshoot if and only if*

$$\frac{\rho^h(1-\rho)}{1-\rho^{h+1}} < \gamma < \min\left\{\frac{2\rho^h}{1+\rho^h}, \rho\right\}. \tag{10}$$

The lower bound in inequality (10) ensures extrapolation from recent events. The upper bound ensures under-reaction to new information and initial under-reaction in the forecast IRF.

We summarize how the model can theoretically reconcile the four sets of empirical findings.

- Findings on consensus forecasts: $\beta_{CG} > 0$, $\beta_{KW} < 0$, and an IRF of forecast errors that displays initial under-reaction followed by overshooting are dictated by the parameters of the adaptive expectations model—specifically, the relative magnitudes of γ and ρ—as stated in Corollary 3. The noisy information component does not affect these three moments.
- Findings on individual forecasts: $\beta_{BGMS} < 0$ arises from the large magnitude of idiosyncratic noise, as shown in equation (8). The noisy information component is what drives the difference between β_{CG} and β_{BGMS}.

3.2.3 | Relation to existing literature

Before relating the model to the existing literature, it is important to emphasize that $\beta_{CG} > 0$, $\beta_{KW} < 0$ and the sign of the IRF are determined solely by the adaptive expectations component, and are unaffected by the noisy information component. The noisy information component matters only for the distinction between β_{CG} and β_{BGMS}. For this reason, I sometimes refer below only to the adaptive expectations model when discussing the first three moments.

Coibion and Gorodnichenko (2015) explain β_{CG} using both the sticky-information and noisy-information models. In both frameworks, under-reaction to new information arises from the slow diffusion of information. Kohlhas and Walther (2021), Bordalo *et al.* (2020) and Angeletos *et al.* (2021), discussed below, also incorporate the noisy-information model, in which forecasters update rationally using the Kalman filter. In my framework, by contrast, the noisy-information component serves only to generate dispersed individual forecasts; there is no rational updating with the Kalman filter. More importantly, the noisy-information model and adaptive expectations yield opposite predictions for the relationship between β_{CG} and ρ, a point examined in Subsection 5.3.

Building on the noisy-information framework, Kohlhas and Walther (2021) propose a rational model in which asymmetric attention to different components of the variable can simultaneously account for $\beta_{CG} > 0$ and $\beta_{KW} < 0$. I show that the adaptive expectations model can reconcile both findings within a certain range of the adjustment parameter γ, without decomposing the variable. This reflects an implicit overlap between under-reaction to new information and extrapolation from recent events. In Subsection 5.1, I extend the benchmark adaptive expectations model to allow for a decomposition of the variable into multiple components, following Kohlhas and Walther (2021), and derive β_{KW}. I then discuss the role of variable decomposition under adaptive expectations.

Bordalo *et al.* (2020) reconcile $\beta_{CG} > 0$ and $\beta_{BGMS} < 0$ by combining the noisy-information model with diagnostic expectations. In their framework, diagnostic forecasters over-react to individual signals, generating $\beta_{BGMS} < 0$, while $\beta_{CG} > 0$ arises from the slow diffusion of information at the aggregate level due to noisy information. In my framework, by contrast, agents' reactions to information are governed solely by the adjustment parameter γ, and the gap between $\beta_{CG} > 0$ and $\beta_{BGMS} < 0$ is largely mechanical, stemming from noisy signals at the individual level.

Angeletos *et al.* (2021) propose a model that reconciles the signs of the three coefficients as well as the initial under-reaction and subsequent overshooting of the forecast IRF. Their framework relies on three assumptions: (a) a noisy-information environment; (b) perceived noise precision that differs from true precision; and (c) perceived persistence of the variable that differs from true persistence. They show that (b) and (c) are necessary to reconcile $\beta_{CG} > 0$ and $\beta_{KW} < 0$. Together with (a), these assumptions also reconcile $\beta_{BGMS} < 0$. Explaining the forecast IRF requires the full set of (a), (b) and (c). By contrast, adaptive expectations provide an alternative and to some extent simpler explanation of $\beta_{CG} > 0$, $\beta_{KW} < 0$ and the IRF, through the single inequality (10).

Given the proliferation of models that depart from full information rational expectations, it is important to ask whether existing frameworks can already reconcile the empirical puzzles before proposing new ones. The adaptive expectations model, although no longer at the centre of the literature, remains a useful benchmark that can deliver sharp insights.

In next section, I fit the model to the data to assess the extent to which it matches the empirical moments quantitatively.

4 | ASSESSING THE MODEL'S FIT TO THE DATA

4.1 | Data

The dataset used in this article is the Survey of Professional Forecasters (SPF) conducted by the Federal Reserve Bank of Philadelphia. Initiated in 1968, the SPF collects quarterly forecasts

TABLE 1 Summary statistics.

		Count	Mean	S.D.	Min	25%	50%	75%	Max
Individual level									
CPI	FE	5834	−0.02	1.56	−10.86	−0.82	−0.18	0.60	9.05
	FR	5438	−0.09	0.82	−10.90	−0.38	−0.02	0.25	8.00
PGDP	FE	8279	0.19	1.66	−22.43	−0.73	−0.13	0.74	11.42
	FR	7527	0.04	1.15	−22.57	−0.31	0.00	0.31	24.72
Consensus level									
CPI	FE	168	−0.01	1.35	−4.07	−0.62	−0.19	0.47	5.43
	FR	168	−0.09	0.49	−1.94	−0.26	−0.04	0.12	1.64
PGDP	FE	219	0.07	1.28	−1.98	−0.66	−0.25	0.47	5.30
	FR	219	0.05	0.51	−1.30	−0.17	−0.03	0.20	3.12

Notes: This table reports the summary statistics of forecast errors (FE) and forecast revisions (FR) at the individual level and consensus level. S.D. is the standard deviation when pooling all observations. All forecast errors and revisions are calculated at horizon $t + 3$. CPI forecast data span from 1981Q3 to 2024Q3, while PGDP data span from 1968Q4 to 2024Q3.

of US macroeconomic indicators from professional forecasters. Historical realizations are taken from vintage data, the real-time dataset provided by the Federal Reserve Bank of Philadelphia, which ensures that realizations align with the information available to forecasters. To illustrate how adaptive expectations can quantitatively account for the empirical patterns, I focus on inflation forecasts from the SPF.

The SPF reports the CPI as a growth rate—the CPI inflation rate—but reports the GDP Price Index (PGDP) in levels. This study uses a one-year forecast horizon. The forecasted one-year PGDP inflation rate is computed as the annual growth rate between the realization in quarter $t - 1$ and the forecasted PGDP level in quarter $t + 3$. The forecasted one-year CPI inflation rate is computed as the average of the CPI inflation rates in quarters t, $t + 1$, $t + 2$ and $t + 3$.

I use CPI forecast data from 1981Q3 to 2024Q3, and PGDP forecast data from 1968Q4 to 2024Q3. On average, each survey wave includes 35 forecasters for the PGDP, and 32 for the CPI. Summary statistics for forecast errors and forecast revisions at horizon $t + 3$ are reported in Table 1.

In Table 1, the standard deviations are computed by pooling across quarters and forecasters. The number of observations is larger for forecast errors than for forecast revisions because calculating a revision requires both the $t + 3$ forecast and the lagged $t + 4$ forecast, and some $t + 4$ forecasts are missing.

Forecast errors and revisions are consistently indistinguishable from zero, indicating no systematic bias in forecasts or asymmetry in revisions. Extreme observations are not winsorized, since forecasters may have received genuinely extreme private signals.

The three regression results are reported in Table 2. The forecast horizon is one year ahead. For consensus time series regressions, standard errors are computed using the Newey–West estimator with a lag of four quarters. For individual-level panel regressions, standard errors are clustered by time.

Here, β_{BGMS} is consistently negative, although not statistically significant; β_{CG} is significantly positive. Including forecaster fixed effects does not materially change the individual coefficients. The two coefficients reported here are broadly consistent with Bordalo *et al.* (2020), with slight differences in magnitude. By contrast, $\beta_{KW} < 0$ does not hold consistently in this empirical application, with one estimate negative, and another positive.

TABLE 2 Three regression coefficients.

	Regression BGMS				Regression CG		Regression KW	
	CPI		PGDP		CPI	PGDP	CPI	PGDP
FR	−0.13	−0.02	−0.14	−0.09	0.72	1.07		
	(0.10)	(0.10)	(0.08)	(0.09)	(0.36)	(0.34)		
Actual observations							−0.05	0.12
							(0.12)	(0.07)
Constant	−0.01	0.004	0.15	0.14	0.06	0.02	0.12	−0.36
	(0.11)	(0.12)	(0.09)	(0.1)	(0.19)	(0.13)	(0.31)	(0.21)
Individual fixed effect	Yes	No	Yes	No				
R^2	0.01	0.00	0.01	0.02	0.07	0.19	0.004	0.06
Observations	5312	5312	7436	7436	168	220	169	219

Notes: This table reports β_{CG}, β_{BGMS} and β_{KW}. For consensus time series regressions, standard errors (in parentheses) are calculated using the Newey–West method with bandwidth 4 quarters. For individual-level panel regressions, standard errors (in parentheses) are clustered by time. CPI forecast data span from 1981Q3 to 2024Q3, while PGDP data span from 1968Q4 to 2024Q3.

4.2 | Estimation

The goal of this estimation exercise is to recover the parameters γ, ρ and $\sigma_\varepsilon^2/\sigma_e^2$ for both measures of inflation, and assess the extent to which the adaptive expectations model can reconcile the four sets of empirical findings simultaneously.

4.2.1 | Estimation strategy

The estimation strategies follow three steps, as below.

1. Fit an AR(1) process to the historical time series to estimate ρ and σ_e for the two measures of inflation.
2. After obtaining the estimates for ρ and σ_e, coefficients β_{CG} and β_{KW} are matched between the data and the model (equations (7) and (9)) to estimate γ. Denote the estimates as $\widehat{\gamma}_{CG}$ and $\widehat{\gamma}_{KW}$, respectively. Standard errors are computed using the delta method. If the model is the true data-generating process, then $\widehat{\gamma}_{CG}$ and $\widehat{\gamma}_{KW}$ should be similar. Comparing $\widehat{\gamma}_{CG}$ and $\widehat{\gamma}_{KW}$ therefore provides one dimension of the model's ability to fit the data.
3. Finally, the average cross-sectional forecast variance is matched between the data and the model to estimate σ_ε. In the data, forecast dispersion is computed in two steps: first, calculate the variance of forecasts within each quarter; second, average these variances across quarters. Denote the estimate as $\widehat{\mathrm{Var}}_i(\mathcal{F}_{i,t})$. I now derive the corresponding cross-sectional forecast variance in the model:

$$\mathrm{Var}_i(\mathcal{F}_{i,t}) = \mathrm{Var}_i\left(\sum_{s=0}^{+\infty}(1-\gamma)^s\gamma(w_{t-s}+\varepsilon_{i,t-s})\right)$$

$$= \mathrm{Var}_i\left(\sum_{s=0}^{+\infty}(1-\gamma)^s\gamma\varepsilon_{i,t-s}\right)$$

TABLE 3 Estimation results from matching β_{CG}.

	CPI			PGDP		
	1981Q3–2024Q2	1981Q3–2019Q4	1981Q3–2007Q4	1968Q4–2024Q3	1968Q4–2019Q4	1968Q4–2007Q4
$\hat{\gamma}_{CG}$	0.52	0.68	0.58	0.47	0.5	0.49
	(0.11)	(0.11)	(0.07)	(0.08)	(0.09)	(0.09)
$\hat{\sigma}_{\varepsilon}^{CG}/\hat{\sigma}_e$	1.61	1.40	1.96	3.25	3.35	3.41

Notes: The first row reports γ estimates from matching β_{CG} between the model and data. Standard errors are reported in parentheses. The second row reports the noise-to-signal ratio $\sigma_{\varepsilon}^{CG}/\sigma_e$ estimates using σ_e estimates from Table A2.

TABLE 4 Estimation results from matching β_{KW}.

	CPI			PGDP		
	1981Q3–2024Q2	1981Q3–2019Q4	1981Q3–2007Q4	1968Q4–2024Q3	1968Q4–2019Q4	1968Q4–2007Q4
$\hat{\gamma}_{KW}$	0.33	1.01	—	0.05	0.05	0.05
	(0.32)	(1.19)		(0.03)	(0.03)	(0.03)
$\hat{\sigma}_{\varepsilon}^{KW}/\hat{\sigma}_e$	2.14	0.99	—	11.27	12.07	12.13

Notes: The first row reports γ estimates from matching β_{KW} between the model and data. Standard errors are reported in parentheses. The second row reports the noise-to-signal ratio $\sigma_{\varepsilon}^{KW}/\sigma_e$ estimates using σ_e estimates from Table A2.

$$= \gamma^2 \sum_{s=0}^{+\infty} (1-\gamma)^{2s} \sigma_{\varepsilon}^2$$

$$= \frac{\gamma}{2-\gamma} \sigma_{\varepsilon}^2. \tag{11}$$

The second equality follows from the fact that w_t is common to all forecasters. Since there are two estimates of γ, $\hat{\gamma}_{CG}$ and $\hat{\gamma}_{KW}$, there are correspondingly two variance estimates, $\hat{\sigma}_{\varepsilon}^{CG}$ and $\hat{\sigma}_{\varepsilon}^{KW}$.

4.2.2 | Estimation results

To assess the robustness of parameter estimates across time, I repeat the estimation procedure for three subsamples: from the beginning of the sample to the end of 2007, to the end of 2019, and to early 2024. This allows me to test whether the 2008 financial crisis and the 2020 pandemic had significant effects on the estimates. The three empirical coefficients across these periods are reported in Appendix Table A3.

Tables 3 and 4 report the estimation results for the adjustment parameter γ and the noise-to-signal ratio $\sigma_{\varepsilon}/\sigma_e$ across the three sample periods. Table 3 is based on targeting β_{CG}, while Table 4 is based on targeting β_{KW}. Parameter estimates obtained from targeting different coefficients are denoted with the subscripts CG or KW. Estimates of ρ and σ_e are reported separately in Appendix Table A2.

First, consider the γ estimates in Tables 3 and 4. In Table 3, $\hat{\gamma}_{CG}$ is stable across sample periods for both measures of inflation. For the PGDP, $\hat{\gamma}_{CG}$ is consistently slightly below 0.5, while for the CPI, it is somewhat larger but still below 0.7. In Table 4, by contrast, the $\hat{\gamma}_{KW}$ estimates are mixed. The empty cell for the CPI in the pre-2008 subsample arises because, given the ρ estimate and $h = 3$, equation (9) cannot match its empirical counterpart within the range $\gamma \in [0, 2]$ in this specific case. This failure to match indicates that the model is not always flexible enough to capture β_{KW} in the data for a fixed ρ. Appendix Figure A3 illustrates this case: the minimum value of equation (9) remains above the empirical β_{KW} for the CPI during 1981Q3–2007Q4. Where the model does succeed in matching the data, the magnitude of $\widehat{\gamma}_{KW}$ varies widely, from 0.05 to 1.01. The dispersion in $\widehat{\gamma}_{KW}$ reflects the dispersion in the empirical estimates of β_{KW}. As shown in Appendix Table A3, point estimates of β_{KW} are positive and insignificant for the PGDP, while for the CPI they are negative, with two cases significant. This evidence suggests that the empirical finding $\beta_{KW} < 0$ is not always robust.

There is a large literature on constant-gain learning (CGL), where the estimated gain parameter is analogous to the adjustment parameter γ in this paper (see Orphanides and Williams 2005; Branch and Evans 2006; Milani 2007, 2008; Marcet and Nicolini 2003). In these studies, the calibrated gain typically ranges from 0.01 to 0.05, an order of magnitude smaller than most γ estimates here. Only the γ_{KW} estimates for the PGDP are comparable. One reason is the difference in model specification. The simple benchmark adaptive expectations model used here differs from the CGL models in those studies. Many papers assume that agents estimate an econometric model to learn the coefficients of structural models (Orphanides and Williams 2005, 2006; Milani 2007, 2008; Eusepi and Preston 2011), which contrasts with the setup in this paper. Another reason is the choice of moments and estimation methods. This paper estimates γ by targeting two new moments, β_{CG} and β_{KW}, which the CGL literature has not considered. Previous studies have used different approaches, including Bayesian estimation within structural dynamic stochastic general equilibrium models (Milani 2007, 2008), targeting root mean square errors (Orphanides and Williams 2005), and targeting forecast errors (Eusepi and Preston 2011).

In Subsection 5.2, I estimate γ using a regression directly implied by equation (1). This approach yields a more natural estimate in the sense that it does not target specific moments. The estimates range from 0.15 to 0.24, smaller than but of the same order of magnitude as most results in Tables 3 and 4. They remain, however, an order of magnitude larger than the learning rates typically found in the CGL literature. In the next subsubsection, I examine model performance, and show that $\gamma = 0.05$ fits poorly, failing to match two other untargeted moments.

Second, consider the noise-to-signal ratio estimates. For the CPI, the ratio ranges from 0.99 to 2.14, and for the PGDP, it ranges from 3.25 to 12.07. In Table 4, the high ratios reflect the low values of $\widehat{\gamma}_{KW}$. From equation (11), holding constant the cross-sectional variance of forecasts in the data, a smaller γ requires a larger σ_ε^2 to match the observed dispersion. Estimates of the noise-to-signal ratio in the literature also span a wide range. In Bordalo *et al.* (2020), the ratio ranges from 0.06 to 1.57 for the CPI, and from 0.73 to 9.26 for the PGDP, across different specifications. The estimates reported here are broadly comparable, though somewhat higher. Bordalo *et al.* (2020) emphasize that in their noisy-information framework, individual noise exceeding fundamental innovations is consistent with rigidity in consensus forecasts.

4.2.3 | Model performance

By construction, the estimation procedure matches empirical β_{CG} (or β_{KW}) and forecast dispersion exactly, since these moments are directly targeted. This subsubsection evaluates model performance using moments that are not targeted.

First, comparing $\widehat{\gamma}_{CG}$ and $\widehat{\gamma}_{KW}$ provides a test of whether the model can simultaneously match empirical β_{CG} and β_{KW}. If the model succeeds, then the confidence intervals of $\widehat{\gamma}_{CG}$ and $\widehat{\gamma}_{KW}$ should align closely. In Tables 3 and 4, the 95% confidence intervals overlap for the CPI but not for the PGDP, indicating that model performance on this dimension varies across variables.

Second, I test whether the model can also match β_{BGMS} in the data, which is not directly targeted. The model-implied β_{BGMS} is computed using the estimated parameters and equation (9). Figure 3 compares the model-implied and empirical values. Figure 3(a) uses the parameter set obtained by targeting β_{CG}, while Figure 3(b) uses the set estimated through targeting β_{KW}. The model-implied β_{BGMS} is shown as blue dots; 95% confidence intervals of the empirical estimates of β_{BGMS} are in red. As Figure 3 shows, empirical β_{BGMS} is matched well when β_{CG} is targeted. When β_{KW} is targeted, the model matches β_{BGMS} only for the CPI, and performs poorly for the PGDP.

To assess quantitatively whether adaptive expectations can fit the dynamic overshooting pattern documented in Angeletos *et al.* (2021), Figure 4 shows the dynamic responses of the inflation outcome, actual forecasts, and forecasts based on adaptive expectations. The IRFs of the inflation outcome and actual aggregate forecasts are taken from Angeletos *et al.* (2021). They estimate the IRFs using two methods: instrumental variable ARMA and local projection. Shocks are taken from Angeletos *et al.* (2020). Standard errors are heteroscedasticity- and autocorrelation-robust, computed using a Bartlett kernel with four lags. Forecasts under adaptive expectations are calculated from equation (1), using the IRF path of the inflation outcome as inputs. Consistent with Angeletos *et al.* (2020), the plotted forecasts are those made three quarters earlier, $\mathcal{F}_{t-3}[\pi_{t,t-4}]$. Figure 4(a) uses $\gamma = 0.47$, based on the γ_{CG} estimates, while Figure 4(b) uses $\gamma = 0.05$, based on the γ_{KW} estimates.

Figure 4(a) shows that with $\gamma = 0.47$, forecasts under adaptive expectations fit the actual forecasts reasonably well. The adaptive expectations forecasts display a similar pattern: initial undershooting followed by overshooting. Quantitatively, the adaptive forecasts overshoot slightly more than the actual forecasts early on, though the difference is not significant. In Figure 4(b), with $\gamma = 0.05$, the model-implied forecasts fit much less well. In this case, forecasts under adaptive expectations respond too weakly to the shock.

The two exercises above shed light on the γ estimates in Tables 3 and 4. Although $\hat{\gamma}_{KW}$ values around 0.05 are closer to the learning rates found in the CGL literature, they perform poorly in matching β_{BGMS} and the delayed overshooting pattern. By contrast, $\hat{\gamma}_{CG} \in [0.47, 0.68]$ fits these two untargeted moments much better.

In summary, the model fits the data reasonably well, though not perfectly. Performance varies by variable. For the CPI, the model can nearly reconcile all four sets of empirical findings simultaneously. For the PGDP, it performs less well in reconciling the findings quantitatively. Nonetheless, the exercise highlights the considerable quantitative potential of the model.

5 | FURTHER DISCUSSION

In this section, I conduct three additional analyses. First, I extend the benchmark adaptive expectations model to allow for asymmetric adjustment parameters across components of the variable. Second, I run a model-implied regression to estimate the adjustment parameter γ. Third, I test a key prediction on which the canonical noisy-information model and adaptive expectations diverge.

5.1 | Asymmetric adjustment parameters to variable components

Kohlhas and Walther (2021) document the coexistence of extrapolation from recent events and under-reaction to new information in the SPF. To reconcile $\beta_{CG} > 0$ and $\beta_{KW} < 0$ simultaneously, they develop a noisy-information model in which agents update forecasts rationally. The key feature of their framework is that agents observe noisy signals from different components of a variable, with the precision of each signal determining the attention paid to that component. They show that as long as agents place more weight on procyclical components relative

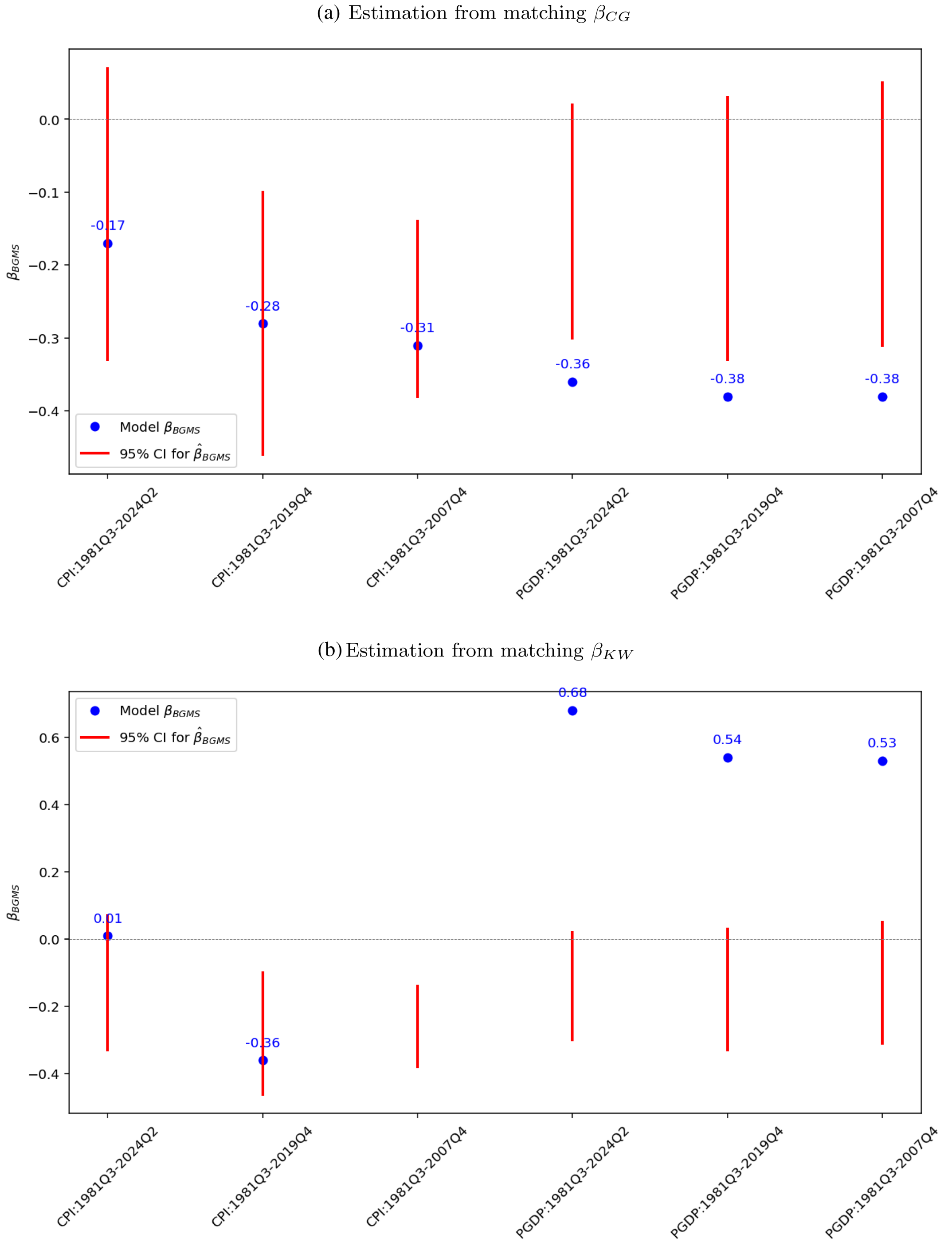
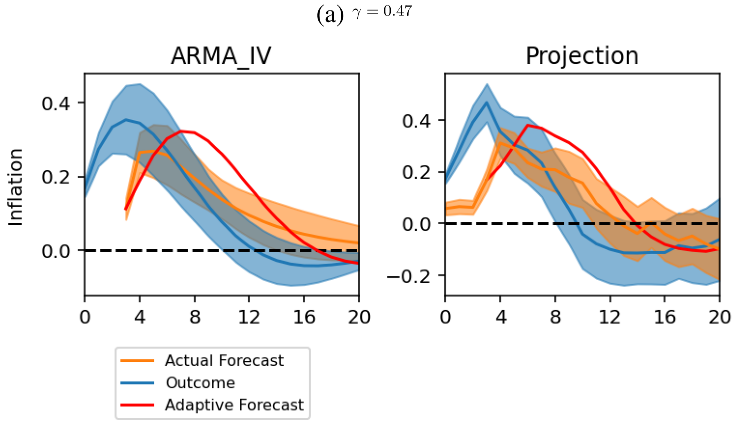


FIGURE 3 Model-implied and empirical β_{BGMS}. *Notes*: These graphs plot the model-implied β_{BGMS} as blue dots, and the 95% confidence intervals for the empirically estimated β_{BGMS} as red lines, for (a) parameters estimated by matching β_{CG} between the model and data, and (b) parameters estimated by matching β_{KW} between the model and data.



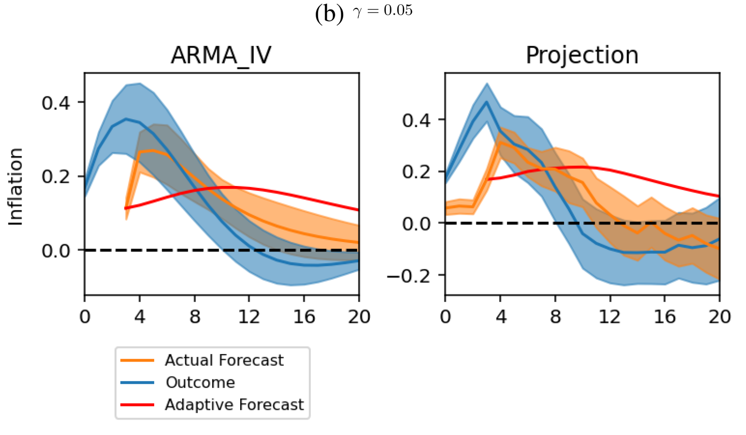


FIGURE 4 Dynamic responses of outcomes and forecasts. *Notes*: These graphs plot the IRFs of the inflation outcome and forecasts from Angeletos *et al.* (2021), shown as blue and orange lines, respectively. The shaded areas represent 68% confidence intervals. The sample period is 1968Q1–2017Q4. The red line represents the adaptive forecasts based on the IRF of the inflation outcome, calculated using (a) $\gamma = 0.47$, (b) $\gamma = 0.05$. The horizontal axis denotes the number of quarters since the shock.

to countercyclical components, the model can generate both $\beta_{CG} > 0$ and $\beta_{KW} < 0$. Intuitively, under-reaction to new information ($\beta_{CG} > 0$) arises from the noisy-information structure, as rational agents gradually update beliefs. Extrapolation from recent events, in turn, 'can be viewed as an outcome of underreactions to countercyclical components'.

I have already shown analytically that the benchmark adaptive expectations model can simultaneously account for $\beta_{CG} > 0$ and $\beta_{KW} < 0$. To shed further light on the role of decomposing the variable into different components, I extend the benchmark model in a direction similar to that of Kohlhas and Walther (2021): each component with its own adjustment parameter in the adaptive expectations framework. This extension provides a deeper interpretation of $\beta_{KW} < 0$.

In the following framework, I focus on a representative forecaster rather than distinguishing between individual and consensus forecasts. This assumption is innocuous for β_{KW}. Following Kohlhas and Walther (2021), suppose that the variable of interest is comprised of N structural components:

$$y_t = \sum_{j=1}^{N} x_{j,t}. \tag{12}$$

The forecaster directly observes the components $x_{j,t}$ for $j \in \{1, 2, \dots, N\}$, and attaches different adjustment parameters γ_j to different components in her expectations:

$$\mathcal{F}_t x_{j,t+h} = \gamma_j x_{j,t} + (1-\gamma_j)\mathcal{F}_{t-1} x_{j,t+h-1}. \tag{13}$$

Her forecast of y_t is the sum of forecasts of each component:

$$\mathcal{F}_t y_{t+h} = \sum_{j=1}^{N} \mathcal{F}_t x_{j,t+h}. \tag{14}$$

Each component is correlated with a latent factor θ_t, which follows an AR(1) process:

$$x_{j,t} = a_j \theta_t \quad \text{for each } i \in \{1, 2, \dots, N\}, \tag{15}$$

$$\theta_t = \rho\theta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \tag{16}$$

where a_j measures the correlation between component j and the latent factor θ_t. Without loss of generality, I assume $\sum_{j=1}^N a_j > 0$. In such an environment, I provide the following result. The proof is provided in the Appendix.

Proposition 3. *Under the framework setup from equations (12)–16, the expression for β_{KW} is given as*

$$\beta_{KW} = \frac{\sum_{j=1}^N a_j \left(\rho^h - \dfrac{\gamma_j}{1-(1-\gamma_j)\rho}\right)}{\sum_{j=1}^N a_j}. \tag{17}$$

Unsurprisingly, equation (9) is a special case of equation (17). When the adjustment parameters are homogeneous across components, $\gamma_j = \gamma$ for all j, the two equations coincide. More generally, equation (17) can be viewed as a weighted-average version of equation (9), with weights given by $a_j / \sum_{j=1}^N a_j$. When adjustment parameters γ_j are heterogeneous across components, $\beta_{KW} < 0$ requires assigning larger weights a_j to components with larger γ_j. Recall that the expression $\rho^h - \gamma/(1-(1-\gamma)\rho)$ is decreasing in γ. Interpreting a_j as a measure of component cyclicality, and γ_j as a measure of 'attention', the mechanism parallels the asymmetric attention highlighted in Kohlhas and Walther (2021): forecasters must place greater attention (larger γ_j) on procyclical components (larger a_j).

5.2 | Model-implied regression

If adaptive expectations is the true data-generating process, then γ can be estimated by fitting the adaptive expectations model in equation (6) to the data. Equation (6) can be rearranged as

$$\overline{\mathcal{F}}_t w_{t+h} - \overline{\mathcal{F}}_{t-1} w_{t+h-1} = \gamma(w_t - \overline{\mathcal{F}}_{t-1} w_{t+h-1})$$

for $h = 0, 1, 2, 3$. Now the regression of interest turns into

$$\overline{\mathcal{F}}_t w_{t+h} - \overline{\mathcal{F}}_{t-1} w_{t+h-1} = \gamma_h^0 + \gamma_h^1(w_t - \overline{\mathcal{F}}_{t-1} w_{t+h-1}) + \varepsilon_t. \tag{18}$$

The subscript h in the coefficients γ_h^1 and γ_h^0 denotes potential variation in γ across forecast horizons. I use the rearranged regression model (18) rather than the original specification (6), since forecast differences are far more stationary than the forecasts themselves, which may trend with the underlying variable. This choice mitigates concerns about spurious regression. I also estimate the regression at the consensus level rather than the individual level, since in the individual-level adaptive expectations model in equation (5), the individual signal $w_{i,t}$ is unobserved by the econometrician. As above, I focus on the one-year horizon $t+3$, or $h=3$.

The estimates of γ_1 range from 0.15 to 0.24 across different inflation measures and time periods, with larger values for earlier subsamples. This implies that forecasters place progressively less weight on current observations and more weight on past forecasts.

These estimates are generally smaller than the γ estimates in Subsubsection 4.2.2. The discrepancy between the γ_1 estimates here and those in Subsubsection 4.2.2 is not surprising. One reason is that the simple regression (18) cannot be the true data-generating process for forecasts, given how complicated humans are. The benchmark adaptive expectations model captures key features of expectations formation and can match certain empirical patterns, but it is not a complete representation of the process.

This exercise tests the external validity of the γ estimates in Tables 3 and 4, since those earlier estimates are an order of magnitude larger than the learning rates in the CGL literature. To address the concern that large γ estimates might be driven by the specific targeted moments, this section estimates γ without targeting any moments. The resulting estimates are of the same order of magnitude as those in Tables 3 and 4.

5.3 | A key model prediction

The model implies predictions about the relationship between coefficients and parameters, as summarized in Corollary 1. The relationship between β_{CG} and ρ is the key prediction on which adaptive expectations and the rational noisy-information model diverge. In adaptive expectations, for a given adjustment parameter γ, greater persistence implies more under-reaction, as reflected in a larger β_{CG}. By contrast, the canonical noisy-information model predicts the opposite relationship. In that framework, β_{CG} should decline with ρ, since more persistent processes lead agents to place greater weight on current signals, reducing information rigidity (Coibion and Gorodnichenko 2015). The root of this difference is that adaptive expectations treat the adjustment parameter as exogenous, whereas in the rational noisy-information model, the Kalman filter endogenously depends on the persistence of the variable.

The empirical relationship between β_{CG} and ρ thus provides a way to distinguish the model in this paper from the canonical noisy-information framework. Coibion and Gorodnichenko (2015) conduct a cross-country, cross-variable test that supports the noisy-information model. By contrast, I present evidence below in favour of adaptive expectations.

Empirically I study a regression

$$\beta_{CG,t} = \phi^0 + \phi^1 \rho_t + \phi^2 \gamma_t + \varepsilon_t. \tag{19}$$

I obtain $\beta_{CG,t}$, ρ_t and γ_t through running 20-year rolling window regressions. Corollary 1 predicts that $\phi^1 > 0$ and $\phi^2 < 0$. Table 6 reports the estimates. The time series graphs of rolling window estimates are shown in Appendix Figures A4 and A5.

According to Table 6, the data are consistent with both sign predictions of the adaptive expectations model. All coefficients of interest are significant except one: the coefficient on γ_t for the CPI.

A natural concern with such time series regressions is spurious correlation arising from non-stationary data. Appendix Figures A4 and A5 suggest this issue visually. I formally test the stationarity of each series using the augmented Dickey–Fuller test and the Kwiatkowski–Phillips–Schmidt–Shin test. The results, reported in Appendix Table A4, show that none of the series passes both tests, confirming the concern. I therefore redo the empirical analysis

TABLE 5 Model-implied regression results.

	CPI			PGDP		
	1981Q3–2024Q2	1981Q3–2019Q4	1981Q3–2007Q4	1968Q4–2024Q3	1968Q4–2019Q4	1968Q4–2007Q4
γ^0	−0.05	−0.03	−0.05	−0.00	0.01	0.00
	(0.04)	(0.03)	(0.05)	(0.03)	(0.03)	(0.04)
γ^1	0.15	0.18	0.24	0.17	0.20	0.20
	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)	(0.05)
R^2	0.11	0.10	0.15	0.20	0.22	0.23
Observations	171	153	105	222	203	155

Notes: This table reports the regression results from equation (18) at the consensus level. Standard errors (in parentheses) are Newey–West with a lag of 4 quarters.

TABLE 6 Relationship between β_{CG} and parameters ρ, γ.

	CPI $\beta_{CG,t}$		PGDP $\beta_{CG,t}$	
ρ_t	6.68	5.36	21.64	23.59
	(3.02)	(2.33)	(2.45)	(2.88)
γ_t	−1.39		−1.94	
	(1.30)		(0.68)	
Constant	−5.95	−4.91	−20.30	−22.59
	(2.72)	(2.21)	(2.37)	(2.82)
Observations	89	89	139	139
Adjusted R^2	0.24	0.20	0.55	0.48

Notes: This table reports the regression results from equation (19). Standard errors (in parentheses) are Newey–West with a lag of 4 quarters.

TABLE 7 Relationship between β_{CG} and parameters ρ, γ—first difference regression.

	CPI $\Delta\beta_{CG,t}$		PGDP $\Delta\beta_{CG,t}$	
$\Delta\rho_t$	2.98	2.04	10.66	10.51
	(3.52)	(4.13)	(3.92)	(4.09)
$\Delta\gamma_t$	1.14		−0.73	
	(0.86)		(0.48)	
Constant	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	88	88	138	138
Adjusted R^2	0.04	0.01	0.28	0.26

Notes: This table reports the regression results from equation (19), after taking the first difference of all variables. Standard errors (in parentheses) are Newey–West with a lag of 4 quarters.

after first-differencing all series. The results, reported in Table 7, continue to show strong support for $\partial\beta_{CG}/\partial\rho > 0$, contradicting the prediction of the noisy-information model.

A similar test is done in Bordalo *et al.* (2020), but they focus on the relationship between their individual coefficient and ρ. Their model provides a clear positive sign on this relationship. Also, they use the variation across different variables, instead of the time series rolling window variation used in this test.

6 | CONCLUSION

While the rational expectations hypothesis has long been the cornerstone of modelling macroeconomic expectations, contemporary data consistently reject its validity. A growing set of models has emerged to account for empirical puzzles in forecast data. This paper shows that the benchmark adaptive expectations model—once central but largely displaced by rational expectations in the 1970s—combined with a noisy-information component, can account for four influential empirical findings both quantitatively and qualitatively. With respect to the relationship between the Coibion–Gorodnichenko coefficient and persistence, on which adaptive expectations and the canonical noisy-information model yield opposite predictions, I find evidence at the time series level in favour of adaptive expectations. In sum, the adaptive expectations framework remains a valuable tool in the evolving study of macroeconomic expectations.

ACKNOWLEDGMENTS

This paper is developed from the second chapter of my PhD thesis. I am deeply indebted to my supervisor, Wouter den Haan, for his invaluable guidance. I am grateful to three anonymous referees for their constructive feedback. I also thank Ricardo Nunes and participants at the London Behavioral Finance Group, the EBS Research Conference, and the Money, Macro and Finance Conference.

REFERENCES

Adam, K., Marcet, A. and Beutel, J. (2017). Stock price booms and expected capital gains. *American Economic Review*, **107**(8), 2352–408.

Afrouzi, H., Kwon, S. Y., Landier, A., Ma, Y. and Thesmar, D. (2023). Overreaction in expectations: evidence and theory. *Quarterly Journal of Economics*, **138**(3), 1713–64.

Angeletos, G.-M., Collard, F. and Dellas, H. (2020). Business-cycle anatomy. *American Economic Review*, **110**(10), 3030–70.

———, Huo, Z. and Sastry, K. A. (2021). Imperfect macroeconomic expectations: evidence and theory. *NBER Macroeconomics Annual*, **35**, 1–86.

Bordalo, P., Gennaioli, N., La Porta, R. and Shleifer, A. (2019). Diagnostic expectations and stock returns. *Journal of Finance*, **74**(6), 2839–74.

———, ———, Ma, Y. and Shleifer, A. (2020). Overreaction in macroeconomic expectations. *American Economic Review*, **110**(9), 2748–82.

———, ——— and Shleifer, A. (2018). Diagnostic expectations and credit cycles. *Journal of Finance*, **73**(1), 199–227.

Branch, W. A. and Evans, G. W. (2006). Intrinsic heterogeneity in expectation formation. *Journal of Economic Theory*, **127**(1), 264–95.

Cagan, P. (1956). The monetary dynamics of hyperinflation. In M. Friedman (ed.), *Studies in the Quantity Theory of Money*. Chicago, IL: University of Chicago Press, pp. 25–117.

Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, **120**(1), 116–59.

——— and ——— (2015). Information rigidity and the expectations formation process: a simple framework and new facts. *American Economic Review*, **105**(8), 2644–78.

d'Arienzo, D. (2020). *Maturity increasing overreaction and bond market puzzles*. Nova School of Business and Economics Working Paper.

de Silva, T. and Thesmar, D. (2023). Noise in expectations: evidence from analyst forecasts. *Review of Financial Studies*, **37**(5), 1494–537.

Eusepi, S. and Preston, B. (2011). Expectations, learning, and business cycle fluctuations. *American Economic Review*, **101**(6), 2844–72.
Farhi, E. and Werning, I. (2019). Monetary policy, bounded rationality, and incomplete markets. *American Economic Review*, **109**(11), 3887–928.
Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press.
Fuhrer, J. C. (2018). Intrinsic expectations persistence: evidence from professional and household survey expectations. Working Paper no. 18-9, Federal Reserve Bank of Boston.
Gabaix, X. (2020). A behavioral New Keynesian model. *American Economic Review*, **110**(8), 2271–327.
Gennaioli, N., Ma, Y. and Shleifer, A. (2016). Expectations and investment. *NBER Macroeconomics Annual*, **30**(1), 379–431.
Giglio, S., Maggiori, M., Stroebel, J. and Utkus, S. (2021). Five facts about beliefs and portfolios. *American Economic Review*, **111**(5), 1481–522.
Kohlhas, A. N. and Walther, A. (2021). Asymmetric attention. *American Economic Review*, **111**(9), 2879–925.
Kucinskas, S. and Peters, F. S. (2024). Measuring under- and overreaction in expectation formation. *Review of Economics and Statistics*, **106**(6), 1620–37.
Liao, J. (2024). Over/underreaction to new information and noise in expectations formation; available online at https://ssrn.com/abstract=4673339 (accessed 5 December 2025).
Lucas, R. E. (1976). Econometric policy evaluation: a critique. *Carnegie–Rochester Conference Series on Public Policy*, **1**, 19–46.
Mankiw, N. G., Reis, R. and Wolfers, J. (2003). Disagreement about inflation expectations. *NBER Macroeconomics Annual*, **18**, 209–48.
Marcet, A. and Nicolini, J. P. (2003). Recurrent hyperinflations and learning. *American Economic Review*, **93**(5), 1476–98.
Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics*, **54**(7), 2065–82.
——— (2008). Learning, monetary policy rules, and macroeconomic stability. *Journal of Economic Dynamics and Control*, **32**(10), 3148–65.
Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, **29**(3), 315–35.
Orphanides, A. and Williams, J. C. (2005). The decline of activist stabilization policy: natural rate misperceptions, learning, and expectations. *Journal of Economic Dynamics and Control*, **29**(11), 1927–50.
——— and ——— (2006). Monetary policy with imperfect knowledge. *Journal of the European Economic Association*, **4**(2/3), 366–75.
Reis, R. (2020). Comment on 'Imperfect expectations: theory and evidence'. *NBER Macroeconomics Annual*, **35**, 99–111.
Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, **50**(3), 665–90.
Wang, C. (2021). Under- and overreaction in yield curve expectations. University of Notre Dame Working Paper.
Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy. NBER Working Paper no. 8673.

How to cite this article: Liao, J. (2025). Adaptive expectations and reaction to information. *Economica*, 1–36. https://doi.org/10.1111/ecca.70026

APPENDIX

A.1 Proofs

A.1.1 Proof of Lemma 1

Under the assumption of the law of iterated forecasts, for a positive integer $h \geq 2$, we have

$$\mathcal{F}_t w_{t+h} = \mathcal{F}_t \mathcal{F}_{t+h-1} w_{t+h} = \mathcal{F}_t[\gamma w_{t+h-1} + (1-\gamma)\mathcal{F}_{t+h-2} w_{t+h-1}] = \mathcal{F}_t w_{t+h-1}.$$

A.1.2 Proof of Proposition 1

Using Lemma 1, and after iterating backwards on equation (5), we get

$$\mathcal{F}_{i,t} w_{t+h} = \mathcal{F}_{i,t} w_{t+1} = \sum_{s=0}^{+\infty} (1-\gamma)^s \gamma (w_{t-s} + \varepsilon_{i,t-s}).$$

Given the large number of forecasters, idiosyncratic noise gets averaged out for each t. The consensus forecast is given by

$$\overline{\mathcal{F}}_t w_{t+h} = \overline{\mathcal{F}}_t w_{t+1} = \sum_{s=0}^{+\infty} (1-\gamma)^s \gamma w_{t-s}.$$

Since w_t follows an AR(1) process, $\mathrm{Cov}(w_t, w_{t+h}) = \rho^h \sigma_e^2/(1-\rho^2)$.

First, let us derive β_{KW}. The denominator is $A = \mathrm{Var}(w_t) = \sigma_e^2/(1-\rho^2)$. The numerator is given by

$$\begin{aligned}
\mathrm{Cov}(w_{t+h} - \overline{\mathcal{F}}_t w_{t+h}, w_t) &= \mathrm{Cov}(w_{t+h}, w_t) - \mathrm{Cov}\left(\sum_{s=0}^{+\infty} (1-\gamma)^s \gamma w_{t-s}, w_t\right) \\
&= \rho^h \frac{\sigma_e^2}{1-\rho^2} - \sum_{s=0}^{+\infty} (1-\gamma)^s \gamma \rho^s \frac{\sigma_e^2}{1-\rho^2} \\
&= \frac{\sigma_e^2}{1-\rho^2}\left[\rho^h - \frac{\gamma}{1-(1-\gamma)\rho}\right].
\end{aligned}$$

As a result,

$$\beta_{KW} = \frac{\mathrm{Cov}(w_{t+h} - \overline{\mathcal{F}}_t w_{t+h}, w_t)}{\mathrm{Var}(w_t)} = \rho^h - \frac{\gamma}{1-(1-\gamma)\rho}.$$

Second, let us derive β_{CG}. The numerator is given by

$$\begin{aligned}
&\mathrm{Cov}(w_{t+h} - \overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_t w_{t+h} - \overline{\mathcal{F}}_{t-1} w_{t+h}) \\
&= \mathrm{Cov}(w_{t+h}, \overline{\mathcal{F}}_t w_{t+h}) - \mathrm{Cov}(w_{t+h}, \overline{\mathcal{F}}_{t-1} w_{t+h}) \\
&\quad - \mathrm{Cov}(\overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_t w_{t+h}) + \mathrm{Cov}(\overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_{t-1} w_{t+h}).
\end{aligned}$$

Let us derive each component one by one:

$$\begin{aligned}
\mathrm{Cov}(w_{t+h}, \overline{\mathcal{F}}_t w_{t+h}) &= \sum_{s=0}^{+\infty} (1-\gamma)^s \gamma \ \mathrm{Cov}(w_{t+h}, w_{t-s}) \\
&= \sum_{s=0}^{+\infty} (1-\gamma)^s \gamma \rho^{h+s} \frac{\sigma_e^2}{1-\rho^2} \\
&= \rho^h \gamma \frac{\sigma_e^2}{1-\rho^2} \frac{1}{1-(1-\gamma)\rho}, \\
\mathrm{Cov}(w_{t+h}, \overline{\mathcal{F}}_{t-1} w_{t+h}) &= \sum_{s=0}^{+\infty} (1-\gamma)^s \gamma \ \mathrm{Cov}(w_{t+h}, w_{t-1-s}) \\
&= \sum_{s=0}^{+\infty} (1-\gamma)^s \gamma \rho^{h+s+1} \frac{\sigma_e^2}{1-\rho^2} \\
&= \rho^{h+1} \gamma \frac{\sigma_e^2}{1-\rho^2} \frac{1}{1-(1-\gamma)\rho},
\end{aligned}$$

$$\mathrm{Cov}(\overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_t w_{t+h}) = \sum_{s=0}^{+\infty}\sum_{k=0}^{+\infty}(1-\gamma)^s\gamma(1-\gamma)^k\gamma\ \mathrm{Cov}(w_{t-s}, w_{t-k})$$
$$= \frac{\gamma^2\sigma_e^2}{1-\rho^2}\sum_{s=0}^{+\infty}\sum_{k=0}^{+\infty}(1-\gamma)^{s+k}\rho^{|s-k|}.$$

Suppose that $h = s-k$. When $h \geq 1$, $s = k+h$. When $h \leq -1$, let $m = -h$, $k = s-h = s+m$. Then

$$\frac{\gamma^2\sigma_e^2}{1-\rho^2}\sum_{s=0}^{+\infty}\sum_{k=0}^{+\infty}(1-\gamma)^{s+k}\rho^{|s-k|}$$
$$= \frac{\gamma^2\sigma_e^2}{1-\rho^2}\left[\sum_{h=1}^{+\infty}\rho^h\sum_{k=0}^{+\infty}(1-\gamma)^{2k+h} + \sum_{m=1}^{+\infty}\rho^m\sum_{s=0}^{+\infty}(1-\gamma)^{2s+m} + \sum_{s=0}^{+\infty}(1-\gamma)^{2s}\right]$$
$$= \frac{\gamma^2\sigma_e^2}{1-\rho^2}\left[\sum_{h=1}^{+\infty}\rho^h(1-\gamma)^h\ \frac{2}{1-(1-\gamma)^2} + \frac{1}{1-(1-\gamma)^2}\right]$$
$$= \frac{\gamma^2\sigma_e^2(1+\rho(1-\gamma))}{(1-\rho^2)(1-(1-\gamma)^2)(1-\rho(1-\gamma))}$$

and

$$\mathrm{Cov}(\overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_{t-1} w_{t+h}) = \gamma^2\sum_{s=0}^{+\infty}\sum_{r=0}^{+\infty}(1-\gamma)^{r+s}\ \mathrm{Cov}(w_{t-r}, w_{t-1-s})$$
$$= \frac{\gamma^2\sigma_e^2}{1-\rho^2}\sum_{s=0}^{+\infty}\sum_{r=0}^{+\infty}(1-\gamma)^{r+s}\rho^{|s+1-r|}.$$

Let $h = s+1-r$. When $h \geq 1$, $s = h+r-1 \geq 0$; when $h \leq -1$, $m = -h$, $r = s+1+m \geq 2$; when $h = 0$, $s+1 = r$. So

$$\frac{\gamma^2\sigma_e^2}{1-\rho^2}\sum_{s=0}^{+\infty}\sum_{r=0}^{+\infty}(1-\gamma)^{r+s}\rho^{|s+1-r|}$$
$$= \frac{\gamma^2\sigma_e^2}{1-\rho^2}\left[\sum_{h=1}^{+\infty}\rho^h\sum_{r=0}^{+\infty}(1-\gamma)^{h+2r-1} + \sum_{m=1}^{+\infty}\rho^m\sum_{s=0}^{+\infty}(1-\gamma)^{2s+1+m} + \sum_{s=0}^{+\infty}(1-\gamma)^{2s+1}\right]$$
$$= \frac{\gamma^2\sigma_e^2(\rho+1-\gamma)}{(1-\rho^2)(1-\rho(1-\gamma))(1-(1-\gamma)^2)}$$

and

$$\beta_{CG} = \frac{\mathrm{Cov}(w_{t+h} - \overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_t w_{t+h} - \overline{\mathcal{F}}_{t-1} w_{t+h})}{\mathrm{Var}(\overline{\mathcal{F}}_t w_{t+h} - \overline{\mathcal{F}}_{t-1} w_{t+h})}$$
$$= \frac{\mathrm{Cov}(w_{t+h}, \overline{\mathcal{F}}_t w_{t+h}) - \mathrm{Cov}(w_{t+h}, \overline{\mathcal{F}}_{t-1} w_{t+h}) - \mathrm{Cov}(\overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_t w_{t+h}) + \mathrm{Cov}(\overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_{t-1} w_{t+h})}{\mathrm{Var}(\overline{\mathcal{F}}_t w_{t+h}) + \mathrm{Var}(\overline{\mathcal{F}}_{t-1} w_{t+h}) - 2\ \mathrm{Cov}(\overline{\mathcal{F}}_t w_{t+h}, \overline{\mathcal{F}}_{t-1} w_{t+h})}.$$

After plugging in components derived earlier, and some tedious derivation, we have

$$\beta_{CG} = \rho^h\left(\frac{1}{\gamma} - \frac{1}{2}\right) - \frac{1}{2}.$$

Third, let us derive β_{BGMS}:

$$\beta_{BGMS} = \frac{\text{Cov}(w_{t+h} - \mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t}w_{t+h} - \mathcal{F}_{i,t-1}w_{t+h})}{\text{Var}(\mathcal{F}_{i,t}w_{t+h} - \mathcal{F}_{i,t-1}w_{t+h})}$$

$$= \frac{\begin{array}{c}\text{Cov}(w_{t+h}, \mathcal{F}_{i,t}w_{t+h}) - \text{Cov}(w_{t+h}, \mathcal{F}_{i,t-1}w_{t+h}) \\ - \text{Cov}(\mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t}w_{t+h}) + \text{Cov}(\mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t-1}w_{t+h})\end{array}}{\text{Var}(\mathcal{F}_{i,t}w_{t+h}) + \text{Var}(\mathcal{F}_{i,t-1}w_{t+h}) - 2\ \text{Cov}(\mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t-1}w_{t+h})}.$$

Let us derive each component one by one:

$$\begin{aligned}\text{Cov}(w_{t+h}, \mathcal{F}_{i,t}w_{t+h}) &= \sum_{s=0}^{+\infty}(1-\gamma)^s\gamma\ \text{Cov}(w_{t+h}, w_{t-s} + \varepsilon_{i,t-s}) \\ &= \sum_{s=0}^{+\infty}(1-\gamma)^s\gamma\ \text{Cov}(w_{t+h}, w_{t-s}) \\ &= \rho^h\gamma\ \frac{1}{1-(1-\gamma)\rho}\ \frac{\sigma_e^2}{1-\rho^2},\end{aligned}$$

and similarly,

$$\text{Cov}(w_{t+h}, \mathcal{F}_{i,t}w_{t+h}) = \rho^{h+1}\gamma\ \frac{1}{1-(1-\gamma)\rho}\ \frac{\sigma_e^2}{1-\rho^2},$$

$$\text{Cov}(\mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t}w_{t+h}) = \gamma^2\sum_{s=0}^{+\infty}\sum_{k=0}^{+\infty}(1-\gamma)^{k+s}\ \text{Cov}(w_{t-s} + \varepsilon_{i,t-s}, w_{t-k} + \varepsilon_{i,t-k}).$$

Suppose that $h = s - k$. When $h \geq 1$, $s = h + k$; when $h \leq -1$, $k = s - h$. Let $m = -h$, $k = s + m$. After substituting in the new index, and tedious derivation, we have

$$\begin{aligned}&\text{Cov}(\mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t}w_{t+h}) \\ &\quad = 2\sum_{h=1}^{+\infty}\sum_{k=0}^{+\infty}\gamma^2(1-\gamma)^{h+2k}\rho^h\ \frac{\sigma_e^2}{1-\rho^2} + \sum_{s=0}^{+\infty}\gamma^2(1-\gamma)^{2s}(A + \sigma_\varepsilon^2) \\ &\quad = \frac{A\gamma^2(1+\rho(1-\gamma)) + \sigma_\varepsilon^2\gamma^2(1-\rho(1-\gamma))}{(1-(1-\gamma)^2)(1-\rho(1-\gamma))}\end{aligned}$$

and

$$\begin{aligned}&\text{Cov}(\mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t-1}w_{t+h}) \\ &\quad = \gamma^2\sum_{s=0}^{+\infty}\sum_{r=0}^{+\infty}(1-\gamma)^{r+s}\ \text{Cov}(w_{t-s} + \varepsilon_{i,t-s}, w_{t-1-r} + \varepsilon_{i,t-1-r}).\end{aligned}$$

Let $h = r - s + 1$. When $h \geq 1$, $r = h + s - 1 \geq 0$; when $h \leq -1$, $m = -h$, $s = r + m + 1 \geq 2$; when $h = 0$, $s = r + 1$. After substituting in the new index and tedious derivation, we have

$$\begin{aligned}&\text{Cov}(\mathcal{F}_{i,t}w_{t+h}, \mathcal{F}_{i,t-1}w_{t+h}) \\ &\quad = \gamma^2\left(\sum_{h=1}^{+\infty}\sum_{s=0}^{+\infty}(1-\gamma)^{2s+h-1}\rho^h A + \sum_{m=1}^{+\infty}\sum_{r=0}^{+\infty}(1-\gamma)^{2r+m+1}\rho^m A\right.\end{aligned}$$

$$+ \sum_{r=0}^{+\infty} (1-\gamma)^{2r+1}(A+\sigma_\varepsilon^2) \Bigg)$$

$$= \gamma^2 \sigma_\varepsilon^2 \, \frac{1-\gamma}{1-(1-\gamma)^2} + \gamma^2 A \, \frac{\rho + 1 - \gamma}{[1-(1-\gamma)^2][1-\rho(1-\gamma)]}.$$

After substituting in all the components, we can derive

$$\beta_{BGMS} = \frac{\rho^h \left((2/\gamma)-1\right) - 1 - (\sigma_\varepsilon^2/\sigma_e^2)[1-\rho(1-\gamma)](1+\rho)}{2 + 2(\sigma_\varepsilon^2/\sigma_e^2)(1-\rho(1-\gamma))(1+\rho)}.$$

Denote $\kappa = (1-\rho(1-\gamma))\,(1+\rho)$.

A.1.3 Proof of Corollary 1

We have

$$\frac{\partial \beta_{CG}}{\partial \gamma} = -\frac{\rho^h}{\gamma^2} < 0,$$

$$\frac{\partial \beta_{KW}}{\partial \gamma} = \frac{\rho - 1}{(1-(1-\gamma)\rho)^2} < 0,$$

$$\frac{\partial \beta_{CG}}{\partial \rho} = h\rho^{h-1}\left(\frac{1}{\gamma} - \frac{1}{2}\right) > 0,$$

$$\frac{\partial \beta_{BGMS}}{\partial (\sigma_\varepsilon^2/\sigma_e^2)} = \frac{-2\kappa\rho^h\left((2/\gamma)-1\right)}{(2+2\kappa\sigma_\varepsilon^2/\sigma_e^2)^2} < 0.$$

A.1.4 Proof of Corollary 2

Let

$$\beta_{BGMS} = \frac{\rho^h((2/\gamma)-1) - 1 - \kappa\sigma_\varepsilon^2/\sigma_e^2}{2 + 2\kappa\sigma_\varepsilon^2/\sigma_e^2} < 0, \tag{A1}$$

$$\beta_{CG} = \rho^h\left(\frac{1}{\gamma} - \frac{1}{2}\right) - \frac{1}{2} > 0, \tag{A2}$$

$$\beta_{KW} = \rho^h - \frac{\gamma}{1-(1-\gamma)\rho} < 0. \tag{A3}$$

Inequalities (A1), (A2) and (A3) yield

$$\sigma_\varepsilon^2 > \frac{\sigma_e^2}{\kappa}\left[\rho^h\left(\frac{2}{\gamma} - 1\right) - 1\right], \quad \gamma < \frac{2\rho^h}{\rho^h + 1}, \quad \gamma > \frac{\rho^h(1-\rho)}{1-\rho^{h+1}},$$

respectively.

A.1.5 Proof of Proposition 2

To calculate

$$\zeta_k \equiv \frac{\partial(w_{t+k} - \overline{\mathcal{F}}_{t+k-1} w_{t+k})}{\partial e_t},$$

first we write both w_{t+k} and $\overline{\mathcal{F}}_{t+k-1}w_{t+k}$ in the moving-average representation of e_t:

$$w_{t+k} = \sum_{s=0}^{+\infty}\rho^s e_{t+k-s},$$

$$\overline{\mathcal{F}}_{t+k-1}w_{t+k} = \sum_{s=0}^{+\infty}(1-\gamma)^s\gamma w_{t+k-1-s} = \sum_{s=0}^{+\infty}(1-\gamma)^s\gamma\sum_{m=0}^{+\infty}\rho^m e_{t+k-1-s-m}.$$

To calculate $\partial w_{t+k}/\partial e_t$, note that when $s=k$, $e_{t+k-s}=e_t$. As a result,

$$\frac{\partial w_{t+k}}{\partial e_t} = \rho^k.$$

To calculate $\partial(\overline{\mathcal{F}}_{t+k-1}w_{t+k})/\partial e_t$, note that when $s+m=k-1$, $t+k-1-s-m=0$. Letting $m = k-1-s$, we have

$$\frac{\partial(\overline{\mathcal{F}}_{t+k-1}w_{t+k})}{\partial e_t} = \sum_{s=0}^{k-1}(1-\gamma)^s\gamma\rho^{k-1-s} = \gamma\,\frac{\rho^k-(1-\gamma)^k}{\rho-1+\gamma}.$$

So

$$\zeta_k = \frac{(\rho-1)\rho^k+\gamma(1-\gamma)^k}{\rho-1+\gamma}.$$

Now let us look for the parameter set such that the initial under-reaction but subsequent overshoot exists. Namely, for some $k_{IRF}\in(1,+\infty)$, $\zeta_k>0$ for $1\leq k<k_{IRF}$, and $\zeta_k<0$ for $k_{IRF}<k$.

When $\gamma\in[1,2]$, $\zeta_k<0$ holds for $k\geq 1$, so it has to be the case that $\gamma\in(0,1)$.

When $\gamma\in(1-\rho,1)$, $\zeta_k>0$ holds only when $(\rho-1)\rho^k+\gamma(1-\gamma)^k>0$. Rearranging,

$$\frac{\gamma}{1-\rho} > \left(\frac{\rho}{1-\gamma}\right)^k.$$

Take the log of both sides and rearranging gives

$$k < \frac{\ln(\gamma/(1-\rho))}{\ln(\rho/(1-\gamma))}.$$

Now k_{IRF} is the next integer greater than or equal to $\ln(\gamma/(1-\rho))\,/\,\ln(\rho/(1-\gamma))$. For k_{IRF} to be strictly larger than 1, the sufficient and necessary condition is

$$\frac{\gamma}{1-\rho} > \frac{\rho}{1-\gamma} > 1,$$

or equivalently, $\gamma(1-\gamma)>\rho(1-\rho)$.

When $\gamma\in(0,1-\rho)$, $\zeta_k>0$ holds only when $(\rho-1)\rho^k+\gamma(1-\gamma)^k<0$. Rearranging,

$$\frac{\gamma}{1-\rho} < \left(\frac{\rho}{1-\gamma}\right)^k.$$

Take the log of both sides and rearranging gives

$$k < \frac{\ln(\gamma/(1-\rho))}{\ln(\rho/(1-\gamma))}.$$

For k_{IRF} to be strictly larger than 1, the sufficient and necessary condition in this case is

$$\frac{\gamma}{1-\rho} < \frac{\rho}{1-\gamma} < 1,$$

or equivalently, $\gamma(1-\gamma) < \rho(1-\rho)$.

In summary, the condition for the existence of $k_{IRF} > 1$ is as follows: when $\gamma \in (1-\rho, 1)$, $\gamma(1-\gamma) > \rho(1-\rho)$ needs to hold; when $\gamma \in (0, 1-\rho)$, $\gamma(1-\gamma) < \rho(1-\rho)$ needs to hold. Next, I try to simplify this parameter condition depending on the value of ρ.

If $\rho \in \left(\frac{1}{2}, 1\right)$, then the condition becomes: when $\gamma \in (1-\rho, 1)$, $\gamma \in (1-\rho, \rho)$ needs to hold; when $\gamma \in (0, 1-\rho)$, $\gamma \in (0, 1-\rho)$ needs to hold. Additionally, I check that at $\gamma = 1-\rho$, $\ln(\gamma/(1-\rho)) / \ln(\rho/(1-\gamma)) > 1$ holds too. So if $\rho \in \left(\frac{1}{2}, 1\right)$, then $\gamma \in (0, \rho)$.

If $\rho \in \left(0, \frac{1}{2}\right)$, then the condition becomes: when $\gamma \in (1-\rho, 1)$, $\gamma \in (\rho, 1-\rho)$ needs to hold; when $\gamma \in (0, 1-\rho)$, $\gamma \in (0, \rho)$ needs to hold. Additionally, I check that at $\gamma = \rho$, $\ln(\gamma/(1-\rho)) / \ln(\rho/(1-\gamma)) > 1$ holds too. So if $\rho \in \left(0, \frac{1}{2}\right)$, then $\gamma \in (0, \rho)$ too.

In summary, the sufficient and necessary condition for the existence of such a k_{IRF} is $\gamma \in (0, \rho)$.

A.1.6 Proof of Corollary 3

I want to find γ in the set

$$\left(\frac{\rho^h(1-\rho)}{1-\rho^{h+1}}, \frac{2\rho^h}{1+\rho^h}\right) \cap (0, \rho).$$

I need to compare ρ with $\rho^h(1-\rho)/(1-\rho^{h+1})$ and $2\rho^h/(1+\rho^h)$ separately.

First, let us find the value of h such that $\rho^h(1-\rho)/(1-\rho^{h+1}) < \rho$ holds. Rearranging,

$$\rho^{h-1} - \rho^h + \rho^{h+1} < 1.$$

The left-hand side of this inequality is decreasing in h, and it holds when $h = 1$. So $\rho^h(1-\rho)/(1-\rho^{h+1}) < \rho$ holds for $h \geq 1$.

Second, let us find the value of h such that $2\rho^h/(1+\rho^h) > \rho$. When $h = 1$, this inequality holds; when $h = 2$, this inequality does not hold. So the relative magnitude of ρ and $2\rho^h/(1+\rho^h)$ depends on h. The set of γ that we are looking for is

$$\left(\frac{\rho^h(1-\rho)}{1-\rho^{h+1}}, \min\left\{\frac{2\rho^h}{1+\rho^h}, \rho\right\}\right).$$

A.1.7 Proof of Proposition 3

Since

$$\beta_{KW} = \frac{\text{Cov}(y_{t+h} - \mathcal{F}_t y_{t+h}, y_t)}{\text{Var}(y_t)},$$

we just need to derive the numerator and denominator separately. We can express y_t in terms of θ_t as

$$y_t = \sum_{j=1}^{N} x_{j,t} = \left(\sum_{j=1}^{N} a_j\right)\theta_t.$$

As a result,

$$\text{Var}(y_t) = \left(\sum_{j=1}^{N} a_j\right)^2 \text{Var}(\theta_t) = \left(\sum_{j=1}^{N} a_j\right)^2 \frac{\sigma_\eta^2}{1-\rho^2}.$$

By iterating backwards on the equation

$$\mathcal{F}_t x_{j,t+h} = \gamma_j x_{j,t} + (1-\gamma_j)\mathcal{F}_{t-1} x_{j,t+h-1},$$

we get

$$\mathcal{F}_t x_{j,t+h} = \sum_{s=0}^{+\infty} (1-\gamma_j)^s \gamma_j x_{j,t-s}.$$

And $\mathcal{F}_t y_{t+h}$ can also be expressed in terms of θ_t, as

$$\begin{aligned}\mathcal{F}_t y_{t+h} = \sum_{j=1}^{N} \mathcal{F}_t x_{j,t+h} &= \sum_{j=1}^{N}\sum_{s=0}^{+\infty} (1-\gamma_j)^s \gamma_j x_{j,t-s} \\ &= \sum_{j=1}^{N}\sum_{s=0}^{+\infty} (1-\gamma_j)^s \gamma_j a_j \theta_{t-s}.\end{aligned}$$

Note that for $s \geq 0$, $\text{Cov}(\theta_t, \theta_{t-s}) = \rho^s \sigma_\eta^2/(1-\rho^2)$. Now we can derive the main expression:

$$\begin{aligned}&\text{Cov}(y_{t+h} - \mathcal{F}_t y_{t+h}, y_t) = \text{Cov}(y_{t+h}, y_t) - \text{Cov}(\mathcal{F}_t y_{t+h}, y_t) \\ &= \text{Cov}\left(\sum_{j=1}^{N} a_j\theta_{t+h}, \sum_{j=1}^{N} a_j\theta_t\right) - \text{Cov}\left(\sum_{j=1}^{N}\sum_{s=0}^{+\infty}(1-\gamma_j)^s\gamma_j a_j \theta_{t-s}, \sum_{j=1}^{N} a_j\theta_t\right) \\ &= \left(\sum_{j=1}^{N} a_j\right)^2 \text{Cov}(\theta_{t+h}, \theta_t) - \left(\sum_{j=1}^{N} a_j\right)\sum_{j=1}^{N} a_j\gamma_j \sum_{s=0}^{+\infty}(1-\gamma_j)^s\ \text{Cov}(\theta_t, \theta_{t-s}) \\ &= \left(\sum_{j=1}^{N} a_j\right)^2 \rho^h \frac{\sigma_\eta^2}{1-\rho^2} - \left(\sum_{j=1}^{N} a_j\right)\sum_{j=1}^{N} a_j\gamma_j \sum_{s=0}^{+\infty}(1-\gamma_j)^s\rho^s \frac{\sigma_\eta^2}{1-\rho^2} \\ &= \left(\sum_{j=1}^{N} a_j\right)\frac{\sigma_\eta^2}{1-\rho^2}\left[\left(\sum_{j=1}^{N} a_j\right)\rho^h - \sum_{j=1}^{N}\frac{a_j\gamma_j}{1-(1-\gamma_j)\rho}\right] \\ &= \left(\sum_{j=1}^{N} a_j\right)\frac{\sigma_\eta^2}{1-\rho^2}\sum_{j=1}^{N} a_j\left(\rho^h - \frac{\gamma_j}{1-(1-\gamma_j)\rho}\right).\end{aligned}$$

After substituting in both numerator and denominator, we get the expression for β_{KW}:

$$\beta_{KW} = \frac{\sum_{j=1}^{N} a_j\left(\rho^h - \gamma_j/\left(1-(1-\gamma_j)\rho\right)\right)}{\sum_{j=1}^{N} a_j}.$$

A.2 When the variable follows an AR(2) process

This subsection explores whether adaptive expectations can simultaneously reconcile $\beta_{CG} > 0$, $\beta_{KW} < 0$ and $\beta_{BGMS} < 0$ when the actual variable follows an AR(2) process

$$w_t = \rho_1 w_{t-1} + \rho_2 w_{t-2} + e_t,$$

where e_t follows a normal distribution, $e_t \sim N(0, \sigma_e^2)$.

I will explore in the domain $\gamma \in (0, 2)$ whether a set of parameters $\{\gamma, \sigma_\epsilon^2\}$ can numerically lead to $\beta_{CG} > 0$, $\beta_{KW} < 0$ and $\beta_{BGMS} < 0$, given the estimated values for $\{\rho_1, \rho_2, \sigma_e^2\}$.

As an example, I use the set of values of $\{\rho_1, \rho_2, \sigma_e^2\}$ from the CPI during the period 1981Q3 to 2024Q2. In Table A1, I report the estimates for $\{\rho_1, \rho_2, \sigma_e\}$.

TABLE A1 AR(2) coefficients.

	CPI
ρ_1	1.27
ρ_2	−0.3
σ_e	0.55

Notes: This table reports the AR(2) coefficients for the CPI during the period 1981Q3 to 2024Q2.

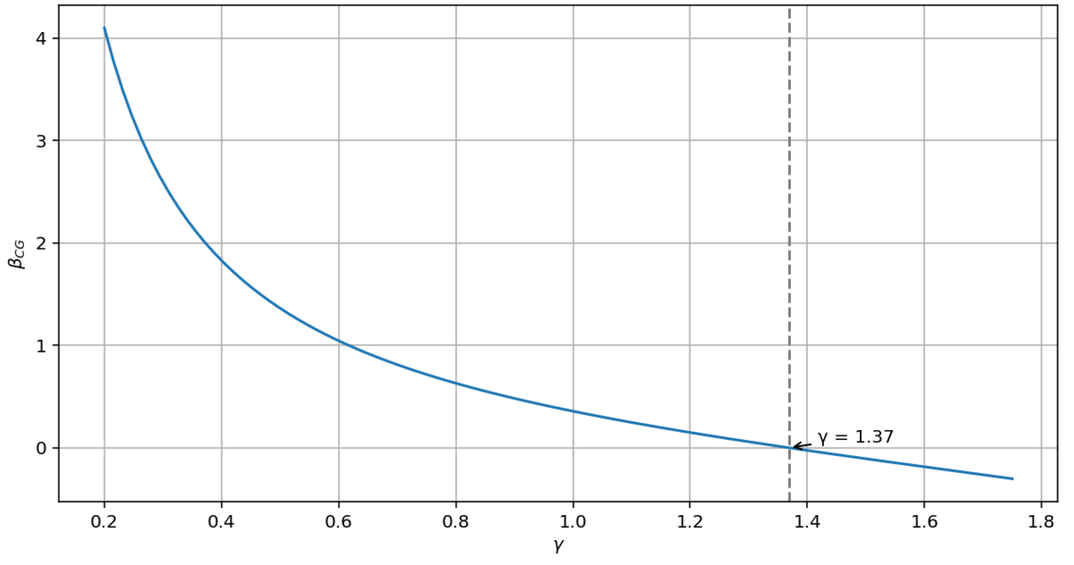
After obtaining the estimates for ρ_1, ρ_2 and σ_e^2, I simulate the CPI series and forecasts under adaptive expectations for various values of γ and σ_ϵ for T quarters, where T is large. Specifically, I simulate one series of actual CPI and 35 series of individual forecasts. First, I plot the relationship between γ and $\beta_{CG} > 0$ and $\beta_{KW} < 0$. Second, given chosen values of γ, I plot the relationship between σ_ϵ^2 and β_{BGMS}.

Figure A1 shows the relationship between β_{CG}, β_{KW} and γ. First, as in the AR(1) case, both coefficients are decreasing in γ. The larger γ, the more over-reaction or extrapolation there is in both measurements. Second, also the same as in the AR(1) case, there is a range of $\gamma \in [0.25, 1.37]$ where $\beta_{KW} < 0$ and $\beta_{CG} > 0$ hold simultaneously.

Figure A2 shows the relationship between β_{BGMS} and σ_ϵ^2. In this example, I choose $\gamma = 0.8$ from the range $[0.25, 1.37]$ above. As in the AR(1) case, β_{BGMS} is decreasing in σ_ϵ^2. When σ_ϵ^2 is larger than around 0.3, β_{BGMS} switches its sign.

In summary, the ability of adaptive expectations to simultaneously reconcile those three coefficients does not depend on the AR(1) assumption.

(a)



(b)

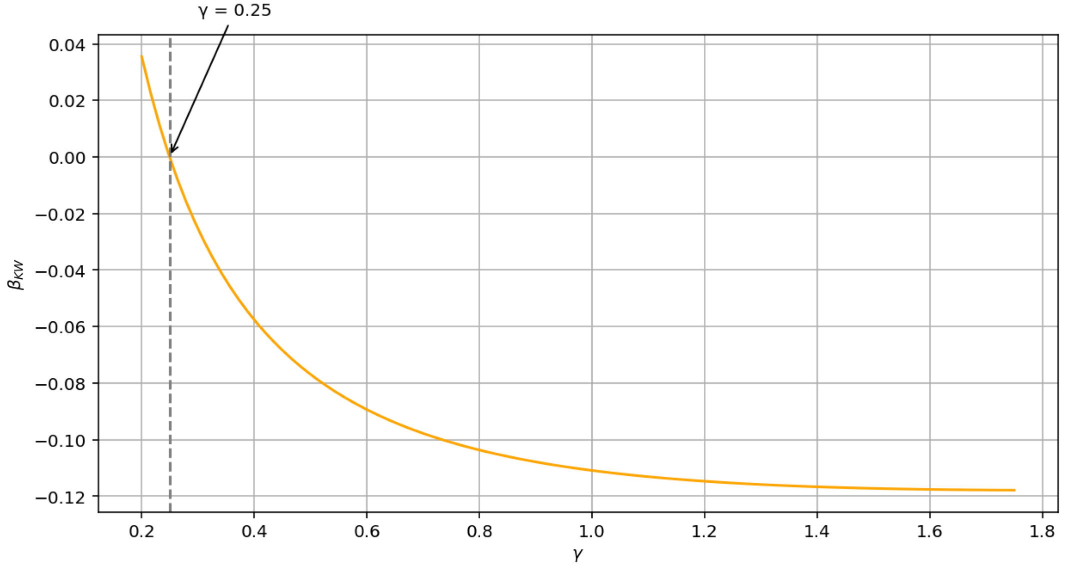


FIGURE A1 β_{CG} and β_{KW} as functions of γ in the AR(2) case. *Notes*: This figure plots β_{CG} and β_{KW} as functions of γ in the AR(2) case. The AR(2) persistence parameters adopt estimates from Table A1.

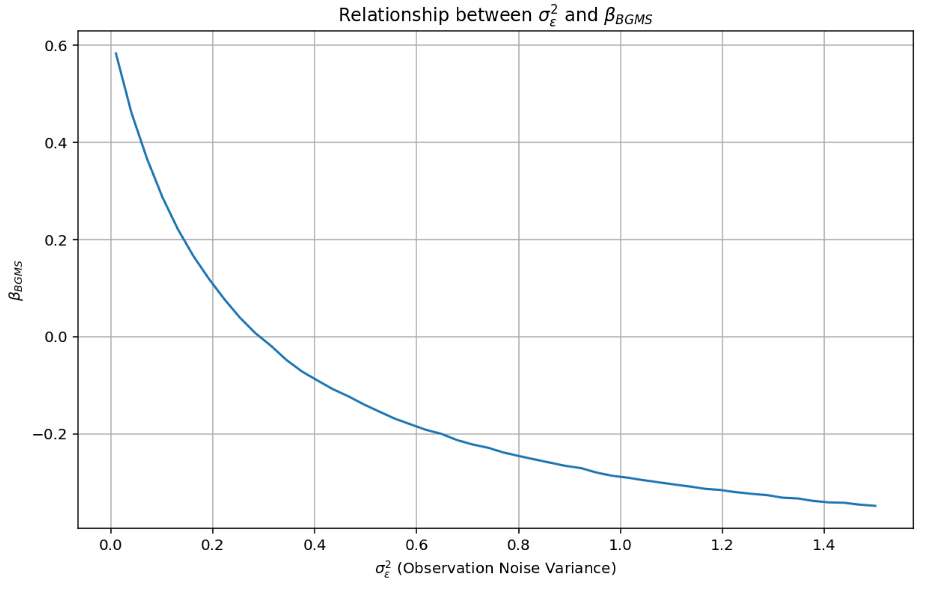


FIGURE A2 β_{BGMS} as a function of σ_ε^2 in the AR(2) case. *Notes*: This figure plots β_{BGMS} as a function of σ_ε^2 in the AR(2) case. The AR(2) persistence parameters adopt estimates from Table A1; $\gamma = 0.8$ lies in the overlap range from Figure A1.

A.3 Extra tables and figures

TABLE A2 Parameter estimates for AR(1) processes.

	CPI			PGDP		
	1981Q3–2024Q2	1981Q3–2019Q4	1981Q3–2007Q4	1968Q4–2024Q3	1968Q4–2019Q4	1968Q4–2007Q4
ρ	0.95	0.94	0.95	0.99	0.99	0.99
σ_e	0.75	0.69	0.58	0.52	0.48	0.52

Notes: This table reports the parameter estimates for the AR(1) process for different periods of time. Here, ρ is the persistence, and σ_e is the standard deviation for the innovation.

TABLE A3 Regression coefficients over different periods.

	CPI			PGDP		
	1981Q3–2024Q2	1981Q3–2019Q4	1981Q3–2007Q4	1968Q4–2024Q3	1968Q4–2019Q4	1968Q4–2007Q4
β_{CG}	0.72	0.31	0.56	1.07	0.97	0.98
	(0.36)	(0.19)	(0.19)	(0.34)	(0.35)	(0.35)
β_{BGMS}	−0.13	−0.28	−0.26	−0.14	−0.15	−0.13
	(0.10)	(0.09)	(0.06)	(0.08)	(0.09)	(0.09)
β_{KW}	−0.05	−0.17	−0.22	0.12	0.12	0.12
	(0.12)	(0.07)	(0.08)	(0.07)	(0.08)	(0.08)

Notes: This table reports the consensus regression results for different sample periods. Standard errors (in parentheses) are Newey–West with a lag of 4 quarters.

TABLE A4 Stationarity test.

	CPI			PGDP		
	$\beta_{CG,t}$	ρ_t	γ_t	$\beta_{CG,t}$	ρ_t	γ_t
ADF p-value	0.51	0.45	0.11	0.31	0.43	0.48
KPSS p-value	0.1	0.02	0.01	0.02	0.1	0.01
Time coverage	2001Q3–2023Q3			1988Q4–2023Q4		

Notes: This table reports the results from the augmented Dickey–Fuller (ADF) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. An ADF test p-value less than 0.05 means rejection of unit root. A KPSS test p-value less than 0.05 means rejection of stationarity.

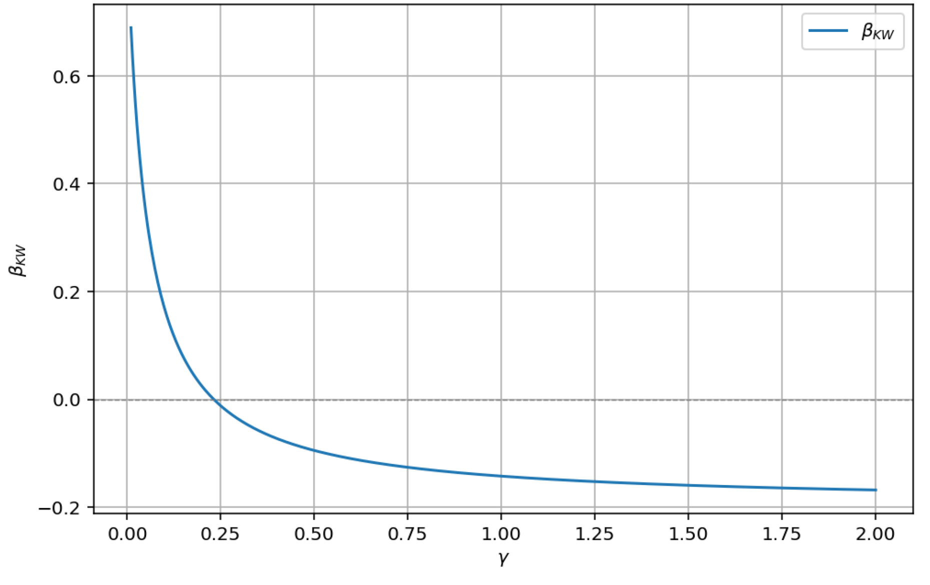


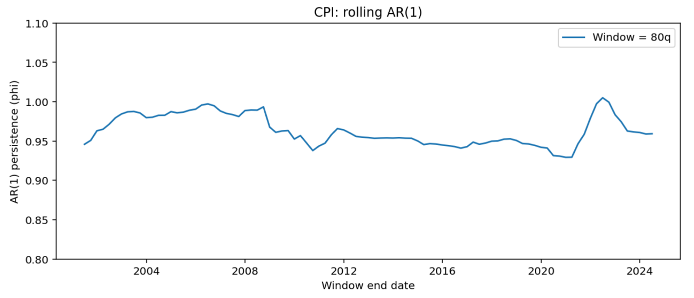
FIGURE A3 β_{KW} as a function of γ. *Notes*: This figure plots β_{KW} in equation (9) as a function of the adaptive learning parameter γ in its domain $[0, 2]$. Here, $\rho = 0.95$ is the AR(1) persistence of the CPI during 1981Q3 to 2007Q4. The smallest value of β_{KW} fails to match the empirical $\beta_{KW} = -0.22$.

A.4 Construction of variables

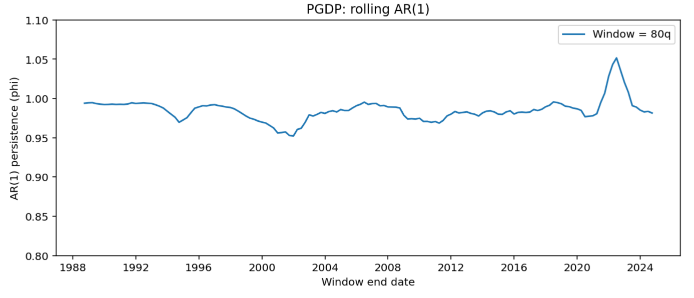
This subsection reports the construction of forecasts of one-year horizon from the SPF. All actual realizations are historically latest observations at the time of making forecasts and extracted from the real-time dataset by the Federal Reserve Bank of Philadelphia.

- PGDP inflation: in SPF, forecasters make forecasts on the level of PGDP, which need to be converted to growth rate. The one-year ahead inflation rate is calculated as $(PGDP5 - PGDP1)/PGDP1 * 100$. To calculate the forecast revision in Bordalo *et al.* (2020), the lagged forecast is calculated as $(PGDP6 - PGDP2)/PGDP2 * 100$.
- CPI: in SPF, forecasters make forecasts on the growth of CPI. The one-year ahead CPI inflation rate is calculated as $(CPI2 + CPI3 + CPI4 + CPI5)/4$. To calculate the forecast revision in Bordalo *et al.* (2020), the lagged forecast is calculated as $(CPI3 + CPI4 + CPI5 + CPI6)/4$.

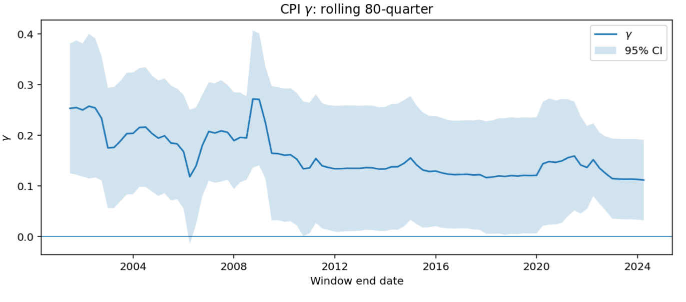
(a)



(b)



(c)



(d)

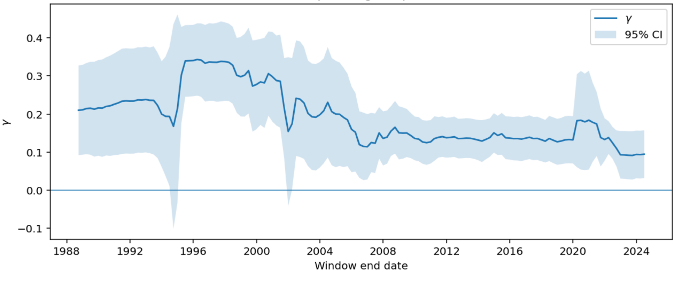
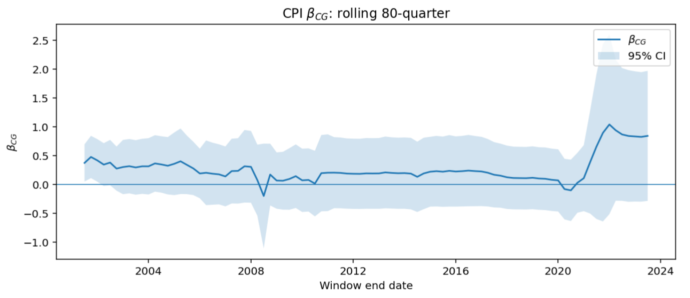
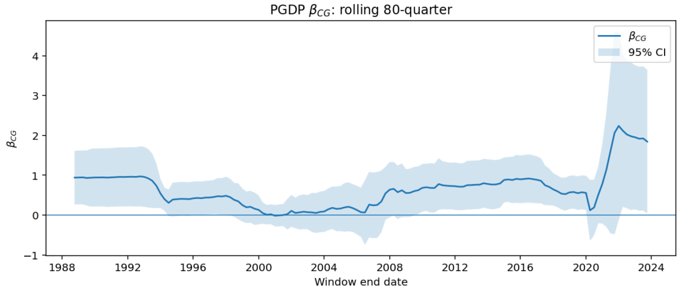


FIGURE A4 Rolling window estimates: ρ and γ. *Notes*: This figure plots the time series of 20-year rolling window estimates for ρ and γ.

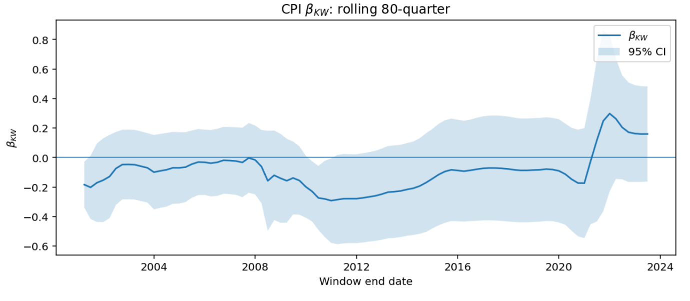
(a)



(b)



(c)



(d)

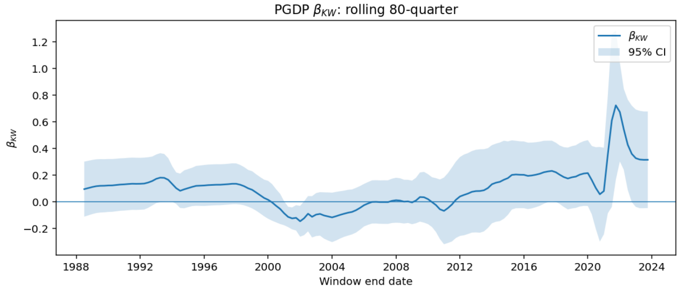


FIGURE A5 Rolling window estimates: β_{CG} and β_{KW}. *Notes*: This figure plots the time series of 20-year rolling window estimates for β_{CG} and β_{KW}.