Institute for
New Economic Thinking

Working Papers
ineteconomics.org/workingpapers

Market Participants Neither Commit Predictable Errors nor Conform to REH: Evidence from Survey Data of Inflation Forecasts

Roman Frydman* and Joshua Stillwagon**

Working Paper No. 163

September 1st, 2021

ABSTRACT

We develop a novel characterization of participants' forecasts with a mixture of normal variables arising from a Markov component. Using this characterization, we formulate five behavioral specifications, including four implied by the diagnostic expectations approach, as well as three implied by REH, and derive several new predictions for Coibion and Gorodnichenko's regression of forecast errors on forecast revisions. Predictions of all eight specifications are inconsistent with the observed instability of individual CG regressions' coefficients, based on inflation forecasts from 24 professionals. Our findings suggest how to build on key insights of the REH and behavioral approaches in specifying individuals' forecasts.

https://doi.org/10.36687/inetwp163

JEL Codes: C52, D80, D83, D91, E31, E71

Keywords: Diagnostic Expectations, Rational Expectations, Model Selection, Structural Change, Inflation Forecasts' Survey

---

* Department of Economics, New York University

** Department of Economics, Babson College, email: jstillwagon@babson.edu. The authors are participants in the Program on Knightian Uncertainty Economics at the Institute for New Economic Thinking (INET). They thank Halina Frydman for suggesting the mixture formulation, which serves as the mathematical basis for the argument in the paper, and for detailed comments on an earlier draft. They also thank Anders Rahbek for insightful and constructive comments that substantially improved the paper. The authors are grateful to INET for its ongoing support of their research. Correspondence to: Roman Frydman, Department of Economics, New York University, 19 West 4th Street, New York, NY 10003, USA; e-mail rf3@nyu.edu

1 Introduction

Coibion and Gorodnichenko (CG, 2015) proposed regressing market participants' forecast errors on their forecast revisions, as measured by survey data, to test predictions of alternative theoretical specifications implied by the rational expectations hypothesis (REH). CG's (pp. 2651, 2653) ingenious idea was that predictions of full- and limited-information rational expectations models "map" onto the constant term and the slope of their proposed regression.[1] This mapping is direct, based on the assumption that participants' forecasting strategies – how they form forecasts on the basis of available information – remain unchanging over time.

Estimating the slope as being positive, CG concluded that theoretical specifications implied by the limited-information REH models are consistent with survey data aggregated over participants' forecasts of a number of variables. In contrast, behavioral economists have amassed compelling empirical evidence that participants' forecasts of outcomes are inconsistent with REH.[2]

Here, we present the estimates of the coefficients of the CG regression that support the behavioral economists' conclusion that participants' forecasts of inflation, as measured by survey data, are inconsistent with REH, regardless of whether participants have access to full or limited information. However, we also find that the predictions of a number of specifications that have been proposed by behavioral economists to formalize participants' departure from REH are also inconsistent with survey data.

Our examination of whether behavioral specifications of participants' forecasts are consistent with the survey data focuses on the so-called diagnostic expectations (DE) approach that Gennaioli and Shleifer (GS) proposed in their influential 2018 book and several co-authored articles. (We sketch the DE approach in Section 4). DE is a particularly well-suited vehicle for this

[1] CG considered two departures from full-information rational expectations: the noisy information models, originated by Lucas (1973) (see Woodford (2003) for a review of the subsequent development of these models), and the sticky-information model proposed by Mankiw and Reis (2002).

[2] For extensive review of this evidence, see Barberis, et al. (1998, pp. 310-317) and Shleifer (2000).

examination.

GS proposed DE as a formalization of Kahneman and Tversky's finding that, in a variety of settings, a psychological mechanism, which they called the representativeness heuristic, drives subjects' assessment of uncertainty. Importantly, GS's specification of DE implies that participants overreact to "news" – new information about payoff-relevant outcomes – relative to the REH-implied forecasts.

This supposed regularity of overreaction led GS to offer DE as a general approach to specifying participants' forecasts in behavioral macroeconomic and finance models. GS's specification of DE aims to implement Thaler's idea that relying on the representativeness heuristic would turn the hodgepodge of early behavioral models into "something resembling [a] science... [of] *predictable errors*" in how individuals make assessments under uncertainty (Thaler, 2017, pp. 489-490, emphasis in the original).[3]

GS (p. 9) have suggested that DE could replace REH in specifying participants' forecasts in macroeconomic and finance models and in policy analysis. Indeed, GS's specification of DE assumes that participants' forecast errors move in a quantitatively predictable way with REH-implied forecasts. Having assumed such "predictability," GS argue that macroeconomic policy analysis based on DE models is not subject to Lucas's critique, because DE would capture how changes in macroeconomic policy would lead participants to revise how they forecast market outcomes. This would enable policymakers to analyze the effects of policy changes on the process driving these outcomes.

However, as we show here, the overreaction supposedly implied by DE is not a regularity. Rather, it is an artifact of GS's particular specification of DE, which rests on their assumption that how the representativeness heuristic impels participants to *deviate* from REH can be formalized with the *REH-implied* forecast revisions.[4]

[3] GS (pp. 137-142) provide an overview of "context-specific" specifications of participants' forecasts in early behavioral models. They argue that DE would not only remedy the shortcomings of these models, but could provide a general approach to specifying participants' predictable errors that could be used in a variety of contexts.

[4] GS base the so-called reference probability distribution, according to which participants

GS's assumption – that participants overreact in the same direction and in the proportionately (predictable) magnitude as the REH forecast revision – is at odds with the empirical evidence, including Kahneman and Tversky's regarding the influence of the representativeness heuristic. In a widely cited early behavioral paper, Barberis, *et al.* (1998, p. 318) have asserted, "If our model is to generate the [pattern] of returns documented in the empirical studies, the investor must be using the wrong model to form expectations" (p. 318). Based on extensive psychological and econometric evidence, they formulate such a "wrong" model by assuming that, while an economist's model specifies earnings to evolve according to a random walk, the investor "thinks that the world moves between two 'states' or 'regimes' and that there is a different model governing earnings in each regime." As is typical in the literature, Barberis, *et al.* formalize this assumption with the two-state Markov chain.[5] Our alternative specification of DE builds on this model.

Once we acknowledge the relevance of behavioral economists' findings, DE no longer implies the regularity of overreaction. Depending on the values of the model parameters and the realizations of payoff-relevant variables, DE overreacts in some periods and underreacts in others periods, relative to the REH-implied forecast.

GS (2018, p. 155) obtain the regularity of overreaction under the assumption that the probability density function (pdf) of outcomes underpinning DE's specification of participants' forecasts is normal. Our alternative specification of DE includes the Markov chain component, which allows for change in how participants forecast outcomes. We show that this extension results in the pdf of outcomes being a mixture of normal pdfs, thereby rendering it compatible with GS's (p. 155, Theorem 5.1) formulation of the DE approach. This novel characterization of participants' forecasting strategies underpins our theoreti-

assess an outcome's representativeness, on an "objective" process driving outcomes. Because this process, according to Muth's hypothesis, underpins REH, DE's overreaction follows immediately from GS's assumption that, like REH, participants' deviation from REH can be formalized as being based on an economist's model. For a formal demonstration, see Section 6.2.

[5] Hamilton (1988) originated modeling of change in REH models with Markov chains. See Hamilton (2008) for an extensive review of subsequent developments.

cal formulations of both behavioral and REH specifications of these strategies, as well the derivation of their predictions for the coefficients of the CG regression.

CG (2015) and Bordalo, *et al.* (2020) confronted the predictions of time-invariant REH and DE specifications, respectively, with survey data. While these predictions "map" directly onto the coefficients of the CG regression, we show these predictions are not generally valid when we recognize that either the process driving outcomes or participants' forecasting strategies allow for a Markov components.

Macroeconomic and finance models typically constrain the parameters of a Markov chain to remain unchanging over an infinite past and indefinite future. Thus, in the context of these models, the unconditional distribution of a model's parameters governed by a Markov chain eventually converges to a steady-state (stationary) probability distribution (Lawler, 2006, p. 15). Assuming stationarity, our mixture characterization enables us to derive predictions of the REH and behavioral specifications, including DE, of participants' forecasts for the coefficients of the CG regression.

Our analysis yields a number of novel implications and testable predictions. On *theoretical* grounds, GS's specification of DE does imply the regularity of overreaction, regardless of whether DE is constrained as being time-invariant or to have a Markov component. However, on *empirical* grounds, the Markov specification of DE predicts that the parameters of the CG regression are compatible with either the regularity of overreaction or underreaction.

We also show that representing participants' forecasts with full-information rational expectations (FIRE) in a model that specifies change with a Markov chain predicts the negative slope coefficients when the sum of the off-diagonal transition probabilities exceeds unity. This is the same prediction as that of GS's time-invariant specification of DE, which aims to formalize departures from FIRE.

Our approach also provides an alternative interpretation of the predictions of the limited-information REH specifications considered by CG (2015). As they point out, constraining the process driving outcomes to be time-invariant predicts that the constant term and the slope of the CG regression are, respectively, zero and positive, if information rigidities are present.

CG (p. 2651) interpreted this prediction as indicating that participants' forecasts, though consistent with REH, deviate from FIRE, owing to noisy information about the state of the economy. However, we show that representing participants' forecasts with FIRE in a model that allows for a Markov component exhibiting even moderate regime persistence, in the sense that the sum of the off-diagonal transition probabilities is less than unity, also predicts the positive slope of the CG regression.

We formulate two alternative specifications of DE: the time-invariant version proposed by GS and its counterpart involving a Markov chain. We also consider the latter specification under the assumption that the Markov chain persists in one state for a prolonged period of time. Furthermore, we derive predictions of a version of Barberis, *et al.'s* (1998) pre-DE behavioral model, as well as of our alternative specification of DE based on that model. Thus, we effectively test the predictions of five alternative behavioral specifications of participants' forecasts.

We also test the predictions of two FIRE specifications: a time-invariant and a Markov specification. Because the predictions of the latter are tantamount to testing predictions of the noisy information REH specification, our estimation effectively confronts three REH-implied specifications of forecasting strategies with survey data.

Our dataset consists of time-series of inflation forecasts from 1969 to 2014 by 24 individuals included in the US Survey of Professional Forecasters (SPF), with over 50 observations by each individual for the three-quarter ahead forecast revision. We estimate a CG regression for each of the forecasters.

Our analysis yields two main findings. First, each of the five behavioral specifications that we consider appear to be inconsistent with survey data for each of the forecasters. Our estimates of the 24 individual CG regressions also appear to be inconsistent with the widely used time-invariant FIRE and noisy information REH specifications, as well as those allowing the process driving outcomes to have a Markov chain component.

Our analysis points to a primary explanation of our findings. Despite their apparent differences, REH and behavioral specifications of participants' forecasting strategies typically rest on a shared premise: these strategies can be represented with a stationary stochastic process over an infinite past and indefinite future.

One of the central implications of our theoretical framework is that, once we represent *change* in the process driving outcomes and in how participants forecast them with a stationary Markov chain, the predicted constant and slope coefficients in the CG regression *do not change* over time. Thus, subjecting the coefficients of the CG regression to tests of structural change provides a hitherto unexplored way to test alternative models of expectations, including diagnostic expectations. We find that these coefficients undergo structural breaks, thereby revealing the inconsistency of the premise of stationarity with survey data on participants' forecasts.

Our findings should be interpreted neither as a rejection of the relevance of Muth's (1961) hypothesis in specifying participants' forecasts nor as a rejection of the behavioral findings that REH models are inconsistent with empirical evidence on how participants actually forecast outcomes. To build models that rest on Muth's hypothesis and yet recognize the relevance of behavioral findings requires acknowledging that the process driving market outcomes undergoes change that cannot be represented with a stationary stochastic process, such as a Markov chain. Our findings suggest that, as Knight (1921) emphasized, market participants recognize the uncertainty that such unforeseeable change engenders and revise their forecasting strategies accordingly.

The plan of the paper is as follows. Sections 2 and 3 provide a formal overview of the DE approach in the context of the Linda experiment and highlight the key steps in applying the approach in macroeconomic and finance models. Building on GS's formulation, Section 4 presents a general definition of overreaction in terms of the means of the "objective" and reference pdfs, which underpin an economist's specification of DE.

Using this definition, Sections 5 and 6 show that the regularity of overreaction is an artifact of GS's specification of the reference pdf as being based on the “objective” process formalized by an economists’ model. Section 7 formulates a Markov counterpart of GS’s time-invariant specification of DE and characterizes the “objective” and reference distributions with mixtures of normal pdfs. This characterization implies that, as with GS’s time-invariant specification, overreaction is assumed to be of the same sign and a fixed proportion of the REH-implied forecast revision. Section 8 formulates the specification of DE based on the behavioral economists’ empirical findings and shows that DE no longer implies the regularity of overreaction.

Relying on the characterization of the REH, reference, and DE specifications with mixtures of normal pdfs established in the foregoing sections, Sections 9, 10, 11 derive predictions of the theoretical formulations of participants’ forecasts for the coefficients of the CG regression. Section 12 presents the summary of these predictions, and Section 13 presents our findings that the predictions of all of the behavioral and REH specifications considered in the paper appear inconsistent with the survey data on forecasts of inflation by each of the 24 forecasters. Section 14 addresses the implications of our findings for building macroeconomic and finance models. The proofs are presented in Online Appendix A. The sketch of the econometric methodology and detailed estimates of the CG regression for each of the 24 forecasters are in Online Appendix B.

2 Diagnostic Expectations in the Linda Experiment

Here, we follow Gennaioli and Shleifer (2018) and provide an overview of the main concepts underpinning their DE approach in the context of the Linda experiment. The simplicity of the experiment enables us to highlight a difficulty overlooked by GS, but which is inherent in any application of the representativeness heuristic in economic models: events that in some contexts appear representative of other events may, in other contexts, appear unrepresentative of those events. As Kahneman and Tversky (1972, p. 431) acknowledged, “Representativeness, like perceptual similarity, is easier to assess than to characterize. In both cases, no general definition is available.”

2.1 An Overview of the Linda Experiment

The Linda experiment features a fictitious 31-year-old woman who currently works as a bank teller. As a college student, Linda engaged in "progressive" activities, including opposing discrimination, advocating for social justice, and participating in anti-nuclear demonstrations. We treat the set of 31-year-old women who graduated from college as a population, which we denote with W. We denote the subset of those who engaged in progressive activities while in college with $H^p \subset W$.

Tversky and Kahneman (TK, 1983 p. 297) presented the following statements to their experiment's subjects:

- Linda is a bank teller, places her among individuals in the set $T \subset W$.
- Linda is a bank teller who is also active in the feminist movement (the set F), which places her among the individuals comprising the intersection $T \cap F \subset W$.

Kahneman and Tversky asked the subjects whether it was more or less probable that Linda is among the bank tellers who are also active in the feminist movement (in $T \cap F$) than that she is among generic bank tellers (in T). An overwhelming majority of subjects responded that it is more probable that Linda is in $T \cap F$ than that she is in T. This finding was then replicated in many Linda-like experiments in a variety of contexts.

2.2 Representativeness in an Experimental Setting

TK (pp. 296-297, 299) hypothesized that their findings could be explained by subjects' reliance on a psychological mechanism, which they called the representativeness heuristic and operationalized in terms of the ratio of the relevant frequencies.[6]

Definition 1 *"An attribute is representative of a class if it is very diagnostic, that is, if the relative frequency of this attribute is* much higher *in that class than in a relevant reference class"*

[6] All citations only to page numbers refer to TK (1983).

For example, in the context of the Linda experiment, TK consider the event $T \cap F$ as an "attribute," H^p as a "class," and individuals who do not have a history of progressive activities, H^{np}, as a "reference class." The idea underpinning TK's operationalization of Definition 1 was that one would expect feminist bank tellers to be more prevalent among the individuals who, like Linda, have a progressive history, $f(T \cap F|H^p)$, than among the individuals who do not have that history, $f(T \cap F|H^{np})$.[7] It is this apparently much greater prevalence that TK referred to in describing $T \cap F$ as being "very diagnostic" of H^p, which they formalized with $\frac{f(T \cap F|H^p)}{f(T \cap F|H^{np})} >> 1$

We assume that the uncertainty about the events in the Linda experiment can be represented with a probability measure on the space $\Omega = H^p \cup H^{np}$. Thus, we operationalize Definition 1 in terms of the ratio of the conditional probabilities:

$$R(A|C, C^{ref}) = \frac{P(A|C)}{P(A|C^{ref})}, \tag{1}$$

where, in the context of our foregoing example, $A = T \cap F \subset \Omega$, $C = H^p$, and $C^{ref} = H^{np}$. According to Definition 1, A "is representative" of C if it is "very diagnostic," that is, if

$$R(A|C, C^{ref}) > d >> 1, \tag{2}$$

where d is some threshold value "much higher" than unity.

2.3 Diagnostic Probabilities

GS (pp. 144-152) introduce DE in the context of the Linda experiment. They represent subjects' assessment of uncertainty with a so-called distorted probability measure and specify how representativeness distorts subjective probabilities (p. 148) as follows:

$$P^{DE}(A|C) = P(A|C) \left[R(A|C, C^{ref}))\right]^{\theta} Z, \tag{3}$$

[7] $f(T \cap F|H^p) = \frac{n(T \cap F \cap H^p)}{n(H^p)}$, $f(T \cap F|H^{np}) = \frac{n(T \cap F \cap H^{np})}{n(H^{np})}$, and $n(\cdot)$ stands for a number of individuals in a respective set.

where $P^{DE}(\cdot|\cdot)$ specifies a distorted (subjective) probability on the space Ω, which we refer to as a diagnostic probability, $P(\cdot|\cdot)$ is the "objective" probability, and $\theta > 0$ formalizes the degree of distortion. Z ensures that (3) specifies a well-defined probability.

3 From the Laboratory to Real-World Markets

To operationalize how the representativeness heuristic "distorts" market participants' assessment of uncertainty, an economist would specify the probability distribution of outcomes (an analog of the attribute $T \cap F$) that he aims to explain in terms of a set of causal variables (an analog of the class H^P), usually called information available to participants. Because any formal economic model rests on the premise that it specifies the "objective" process driving outcomes, an economist, relying on Muth's (1961) hypothesis, can then represent a participant's "rational" assessment of uncertainty, and her REH forecasts, with the "objective" distribution, as specified by the economist's model.

However, there does not appear to be a theoretical argument that would enable an investigator – an experimental psychologist or an economist – to specify the reference class. By providing information to the subjects that Linda has a progressive history, H^P, TK (p. 300) aimed to influence them to compare her to those who do not have that history, thereby considering H^P as the relevant reference class.

In real-world market settings, by contrast, an economist has no way to influence participants' interpretation of the context within which they assess representativeness of uncertain events. However, empirical evidence on how market participants actually assess uncertainty provides a basis for specifying the reference class that participants might have considered relevant. Behavioral economists have provided compelling evidence that participants' forecasts do not conform to REH, and formalizing this evidence could provide the basis for specifying the reference class.

GS proposed DE to provide a unified approach to explain such behavioral findings. However, GS chose to formalize their argument – that the representativeness heuristic impels participants to overreact to information, relative to the REH forecasts – with a specification of the reference class of outcomes that is based on the "objective" probability distribution, which underpins REH. However, as we show in Section 8, once we specify the probability distribution of the reference class of outcomes on the basis of behavioral economists' findings, the supposedly "distorting" influence of the representativeness heuristic, as formalized with the analog of (3), does not result in the regularity of overreaction. DE overreacts to information in some periods and underreacts in others.

4 Representativeness in Macroeconomics and Finance Models

In contrast to the Linda experiment, the concept of representativeness in macroeconomic and finance models involves continuous random variables. To fix ideas, we consider a payoff-relevant variable $x_{t+1} = \ln \widetilde{x}_{t+1}$, and formalize an "attribute" (an analog of $A = T \cap F$ in (1)) with the measurable event, $x_{t+1} \in A \subset \mathbb{R}^+$, and a "class" (an analog of $C = H^p$) with an event $x_t \in C \subset \mathbb{R}^+$. We also operationalize the "reference class" (an analog of $C^{ref} = H^{np}$) with an event $x_t^{ref} \in C^{ref} \subset \mathbb{R}^+$.

GS (p. 154) define $x_{t+1} \in A$'s representativeness of x_t, relative to x_t^{ref}, in terms of the ratio of conditional probability density functions (pdfs), as follows:

$$R^{gs}(x_{t+1}|x_t, x_t^{ref}) = \frac{f(x_{t+1}|x_t)}{f^{ref}(x_{t+1}|x_t^{ref})} > 1, \; x_{t+1} \in A, x_t \in C, x_t^{ref} \in C^{ref} \tag{4}$$

where $f(x_{t+1}|x_t)$ is the "objective" (conditional) pdf of x_{t+1}, as hypothesized by an economist's model.[8] We refer to $f^{ref}(x_{t+1}|x_t^{ref})$ as a (conditional) reference pdf, which is assumed by an economist to characterize the reference class that participants consider relevant. We note that GS's (p. 154) specification

[8] In addition to x_t, an economist's model typically specifies the conditioning set to include other relevant information (such as realizations of the model's variables) up to time t. Allowing for such a larger information set would not alter any of our conclusions here.

of the reference class of outcomes specifies $x_t^{ref} = x_{t-1}$.

However, TK define representativeness in terms of probabilities (or, equivalently frequencies of discrete events), which for continuous variables can be written as

$$R(x_{t+1}|x_t, x_t^{ref}) = \frac{\int_A f(x_{t+1}|x_t)dx_{t+1}}{\int_A f^{ref}(x_{t+1}|x_t^{ref})dx_{t+1}} > 1,\ x_t \in A, x_t^{ref} \in C^{ref}. \tag{5}$$

As we state in the following proposition, there is the event $x_{t+1} \in A$ for which, conditional on the realizations of x_t and x_t^{ref}, GS's definition, in (4) implies that $x_{t+1} \in A$ is representative according to TK's definition, in (5).

Proposition 2 *If the ratio of "objective" and reference pdfs satisfies (4), there exists an event $x_{t+1} \in A$, which is representative of x_t, relative to x_t^{ref}, in the sense that (5) holds.*

Remark 3 *According to TK's experimentally-based Definition 1, formalized in (1) and (2), an event A is representative of C, relative to C^{ref}, if $R(x_{t+1}|x_t, x_t^{ref}) > \delta >> 1$. However, $R^{gs}(x_{t+1}|x_t, x_t^{ref}) > 1$, in (4), for $x_t \in C, x_t^{ref} \in C^{ref}$ does not, in general, imply that $R(x_{t+1}|x_t, x_t^{ref}) > \delta >> 1$. Because this discrepancy between GS's and TK's definitions of representativeness does not play a role in our argument here, we rely on GS's definition in (4).*

4.1 Tractable Specification

To render the operationalization in (4) tractable in specifying market participants' forecasts and in deriving the testable predictions of macroeconomic and finance models, GS (p. 155) characterize how x_{t+1} *actually* unfolds over time with a conditional normal pdf. Consequently, they specify the "objective" pdf of x_{t+1}, conditional on x_t, as

$$f(x_{t+1}|x_t) = \frac{1}{\sigma_{t+1|t}\sqrt{2\pi}} \exp\left[-\frac{(x_{t+1} - m_{t+1|t})^2}{2\left(\sigma_{t+1|t}\right)^2}\right],\ x_{t+1} \in A, x_t \in C, \tag{6}$$

where $m_{t+1|t}$ and $\left(\sigma_{t+1|t}\right)^2$ denote the conditional mean and the variance. GS (p. 155) also assume that the reference class that underpins participants' assessment of x_{t+1}'s representativeness can be characterized with the normal pdf:

$$f^{ref}(x_{t+1}|x_t^{ref}) = \frac{1}{\sigma_{t+1|t}^{ref}\sqrt{2\pi}} \exp\left[-\frac{(x_{t+1} - m_{t+1|t}^{ref})^2}{2\left(\sigma_{t+1|t}^{ref}\right)^2}\right], \ x_{t+1} \in A, \ x_t^{ref} \in C^{ref}, \tag{7}$$

where $m_{t+1|t}^{ref}$ and $\left(\sigma_{t+1|t}^{ref}\right)^2$ denote the conditional mean and variance.

4.2 Diagnostic Expectations

Using (4), GS (p. 154) specify the "distorted" pdf of x_{t+1} in the class x_t:

$$f^{de}(x_{t+1}|x_t) = f(x_{t+1}|x_t)\left[R^{gs}(x_{t+1}|x_t, x_t^{ref})\right]^{\theta} Z(\theta, x_t, x_{t-1}), \tag{8}$$

where, we refer to $f^{de}(x_{t+1}|x_t)$ as the diagnostic pdf, $\theta > 0$, and $Z(\theta, x_t, x_{t-1})$ is specified to ensure that $f^{de}(x_{t+1}|x_t)$ integrates to 1. We denote the conditional mean of a diagnostic density with $m_{t+1|t}^{de}$. GS call $m_{t+1|t}^{de}$ a diagnostic expectation (DE) of x_{t+1}, conditional on x_t.
GS's Proposition 5.1. (p. 155), which we restate here, provides the basis for their argument that DE implies the regularity of overreaction.

Proposition 4 *Suppose that, as specified in (6) and (7), the "objective," and reference (conditional) pdfs underpinning representativeness, in (4), are normal. Then, provided that* $(1+\theta)\left(\sigma_{t+1|t}^{ref}\right)^2 > \theta\left(\sigma_{t+1|t}\right)^2$, *there exists* $Z(\theta, x_t, x_{t-1})$ *that renders the diagnostic pdf,* $f^{de}(x_{t+1}|x_t)$ *in (8), a well-defined normal pdf with the following conditional mean and variance,*

$$m_{t+1|t}^{de} = m_{t+1|t} + \gamma\left(m_{t+1|t} - m_{t+1|t}^{ref}\right), \tag{9}$$

$$\left(\sigma_{t+1|t}^{de}\right)^2 = \frac{\gamma\left(\sigma_{t+1|t}^{ref}\right)^2}{\theta}, \tag{10}$$

where

$$\gamma = \theta \frac{\left(\sigma_{t+1|t}\right)^2}{\left(\sigma^{ref}_{t+1|t}\right)^2 + \theta\left[\left(\sigma^{ref}_{t+1|t}\right)^2 - \left(\sigma_{t+1|t}\right)^2)\right]} > 0. \tag{11}$$

Proof: GS (pp. 217-19).

4.2.1 REH-Implied Specification of Participants' Forecasts

The core premise of an economic model is that it formalizes an economist's reasonable (theoretically and empirically-based) understanding of the actual ("objective") uncertainty about payoff-relevant outcomes. Building on this premise, Muth (1961, p. 316) advanced the pathbreaking hypothesis that an economist could formally relate a participant's forecasts to "the way the economy works" by specifying them as being consistent with an economic model's specification of the process driving outcomes. Muth implemented his hypothesis in a model that assumed that how outcomes have unfolded over an infinite past and will unfold over an indefinite future can be represented with a stationary stochastic process. It was this implementation that came to be known as the rational expectations hypothesis (REH).

Adopting Muth's hypothesis, (9) the conditional mean and variance of the pdf characterizing the REH forecast are the same as their "objective" counterparts, that is, $m^{reh}_{t+1|t} = m_{t+1|t}$ and $\sigma^{reh}_{t+1|t} = \sigma_{t+1|t}$. This enables us to represent $m^{de}_{t+1|t}$ in (9) as

$$m^{de}_{t+1|t} = m^{reh}_{t+1|t} + \gamma\left(m^{reh}_{t+1|t} - m^{ref}_{t+1|t}\right), \tag{12}$$

where γ in (11) is defined accordingly. GS (p.155) refer to $m^{de}_{t+1|t} > m^{reh}_{t+1|t}$ ($m^{de}_{t+1|t} < m^{reh}_{t+1|t}$) as the "overreaction" ("underreaction") of DE, relative to the REH forecast. Proposition 4 shows that if $f(x_{t+1}|x_t)$ and $f^{ref}(x_{t+1}|x_t)$ are normal, then DE overreacts if and only if $m^{reh}_{t+1|t} > m^{ref}_{t+1|t}$.

5 Representing Deviations from REH as Driven by Revision of the REH Forecast

GS proposed DE as a new approach to specifying forecasts in behavioral-finance models that aimed to explain empirical findings that participants' forecasts do not conform to REH. However, their specification of the reference pdf shares a key feature with its REH counterpart: both are based on the "objective" process driving outcomes, as formalized by an economist's model. However, in contrast to the REH forecast, which is conditional on x_t, GS (p. 154) specified the mean of the reference pdf, $m^{ref}_{t+1|t}$ as conditional on x_{t-1}. We refer to this specification as REH-like and denote it with $m^{reh}_{t+1|t-1}$. We state this key assumption of GS's specification of DE as follows:

Assumption 5 *The "distorting" influence of the representativeness heuristic on participants' forecasts,* $m^{reh}_{t+1|t} - m^{ref}_{t+1|t}$*, is driven solely by the revision of its REH counterpart, which we formally state as follows*

$$m^{de}_{t+1|t} - m^{reh}_{t+1|t} = \gamma\left(m^{reh}_{t+1|t} - m^{ref}_{t+1|t}\right) = \gamma\left(m^{reh}_{t+1|t} - m^{reh}_{t+1|t-1}\right). \tag{13}$$

This assumption implies that the supposed regularity of overreaction is in fact generated by a well-known property of REH forecasts: *by design*, the revision of such a forecast is driven solely by the time-t realization of news about x_t.

6 Overreaction as an Artifact of the REH-like Specification of the Reference PDF

GS (p.174) illustrate their argument that DE implies an overreaction in the context of the following standard AR(1) model,

$$x_{t+1} = \rho x_t + \mu + \varepsilon_{t+1}, \tag{14}$$

where $0 < \rho < 1$ and μ are constants, and $\varepsilon_t \sim iidN(0, \sigma^2)$. In the context of this section, this model specifies the "objective" process driving a payoff-relevant variable x_t. Thus, according to Muth's hypothesis,

$$m_{t+1|t}^{reh(gs)} = E(x_{t+1}|x_t) = \rho x_t + \mu \tag{15}$$
$$= \rho^2 x_{t-1} + (1+\rho)\mu + \rho e_t, \tag{16}$$
$$\left(\sigma_{t+1|t}^{reh(gs)}\right)^2 = E\left[x_{t+1} - E(x_{t+1}|x_t)\right]^2 = \sigma^2, \tag{17}$$

where e_t, in (16), denotes the realization of ε_t.

Furthermore, according to Assumption 5, the mean and the variance of the reference pdf, in (7), are given by

$$m_{t+1|t}^{ref(gs)} = E(x_{t+1}|x_{t-1}) =$$
$$= \rho^2 x_{t-1} + (\rho+1)\mu, \tag{18}$$
$$\left(\sigma_{t+1|t}^{ref(gs)}\right)^2 = (1+\rho^2)\sigma^2. \tag{19}$$

Because the "objective" and reference pdfs are normal and $\left(\sigma_{t+1|t}^{ref(gs)}\right)^2 > \left(\sigma_{t+1|t}^{reh(gs)}\right)^2$, Proposition 4 holds, which together with Assumption 5, implies that

$$m_{t+1|t}^{de(gs)} - m_{t+1|t}^{reh(gs)} = \gamma^{(gs)}\left(m_{t+1|t}^{reh} - m_{t+1|t-1}^{reh}\right) = \left(m_{t+1|t}^{reh} - m_{t+1|t-1}^{reh}\right) = \gamma^{(gs)}\rho e_t, \tag{20}$$

where $\gamma^{(gs)} = \frac{\theta}{(1+\rho^2)(1+\theta)}$, and e_t is the realization of ε_t.

6.1 News

The difference between the payoff-relevant variable, such as x_t, and its REH forecast,

$$\eta_t = X_t - E(X_t|x_{t-1}), \tag{21}$$

is usually referred to as news about x_t, where $E(X_t|x_{t-1})$ is a conditional expectation of the "objective" process driving x_t. For the process in (14), and using (16), the realization of news, in (21), is given by

$$n_t^{(gs)} = e_t. \tag{22}$$

6.2 Representing Overreaction with the REH Forecast Revision

GS (p. 155) refer to $e_t > 0$ ($e_t < 0$) as good (bad) news about the payoff-relevant outcome x_t. Expressions (20) and (22) show that the supposed regularity of overreaction, relative to REH, implied by this news, is an artifact of GS's Assumption 5: good (bad) news leads participants to overreact in the same direction and in the proportionately (predictable) magnitude as the REH forecast revision.

7 Allowing for Change in the REH-like Specification of Reference PDF

The AR(1) process, in (14), exemplifies the typical structure of macroeconomic and finance models, an overwhelming majority of which assume away altogether change in the process driving outcomes. Macroeconomic and finance models that recognize that this process undergoes change typically represent it with a Markov chain. Constraining change with such probabilistic rules implies that the news, as defined in (21), comprises the realizations of ε_t as well as the realized state of the Markov chain at t. As we show here, DE involving a Markov component implies the regularity of overreaction to news. As with the time-invariant specification, this regularity is an artifact of GS's Assumption 5 that the "distorting" influence of the representativeness heuristic on participants' forecasts can be represented with the revision of the REH forecast.

7.1 A Markov Specification of the Change in the Process Driving An Outcome

We follow the prevailing practice in a particularly simple way by allowing the mean of the process, in (14), to change over time, which we formally state as follows:

$$x_{t+1} = \rho x_t + \mu_{t+1} + \varepsilon_{t+1}, \tag{23}$$

where μ_t evolves according to a Markov chain, which switches between two states, $\mu^{(1)}$ and $\mu^{(2)}$ with the transition probabilities p_{12} and p_{21} . Here $0 < \rho < 1$ is a constant, and $\varepsilon_t \sim iidN(0, \sigma^2)$. It follows from (23) that, while x_t and μ_{t-i} for $i = 0, 1...$ are dependent, x_t and μ_{t+i} for $i = 1, 2....$are independent.

Macroeconomic and finance models typically constrain the parameters of a Markov chain, such as $(\mu^{(1)}, \mu^{(2)}, p_{12}, p_{21})$, to remain unchanging over an infinite past and indefinite future. Thus, in the context of these models, the unconditional distribution of μ_t eventually converges to a steady-state (stationary) probability distribution (Lawler, 2006, p. 15). In accordance with the usual practice, we make the following assumption:

Assumption 6 *The distribution of the Markov process μ_t is stationary: $P(\mu_t = \mu^{(1)}) = \pi$, $P(\mu_t = \mu^{(2)}) = (1 - \pi)$, for all t.*

This assumption implies that for all t

$$E(\mu_t) = \pi\mu^{(1)} + (1 - \pi)\mu^{(2)}, \tag{24}$$

$$V(\mu_t) = \pi(1 - \pi)\left(\mu^{(1)} - \mu^{(2)}\right)^2, \tag{25}$$

where the expression for $V(\mu_t)$ is derived in the proof of Lemma 8 in Online Appendix.

7.2 "Objective" PDF as a Mixture of Normal PDFs

Allowing μ_t to evolve according to a Markov chain implies that the conditional "objective" and reference pdfs are no longer simply normal. However, the following lemma shows that the "objective" pdf implied by (23) is a mixture of the two normal pdfs with the following means and variances:

$$m_{t+1|t}^{(mk,i)} = E(x_{t+1}|x_t, \mu_{t+1} = \mu^{(i)}) = \rho x_t + \mu^{(i)}, \tag{26}$$

$$\left(\sigma^{(mk,i)}\right)^2 = \sigma^2 + E(\mu_t - \mu^{(i)})^2, \tag{27}$$

where "mk" in the superscript "(mk, i)" denotes that the process driving x_{t+1} has a Markov component, and "i " denotes whether $\mu_{t+1} = \mu^{(1)}$ or $\mu_{t+1} = \mu^{(2)}$.

Lemma 7 *Suppose that (23) characterizes the process driving x_{t+1}. Then, conditional on x_t, the "objective" pdf of x_{t+1}, denoted with $g^{reh(mk)}(x_{t+1}|x_t)$, is the following mixture of the two conditional normal pdfs with the means and variances in (26) and (27):*

$$g^{reh(mk)}(x_{t+1}|x_t) = \pi f^{(mk,1)}(x_{t+1}|x_t, \mu_{t+1} = \mu^{(1)}) + (1-\pi) f^{(mk,2)}(x_{t+1}|x_t, \mu_{t+1} = \mu^{(2)}), \tag{28}$$

Proof in Online Appendix A.

This lemma implies that the "objective" pdf is normal. Furthermore, from (28), (26), (27), (24), and (25), the (conditional) mean and variance of $g^{reh(mk)}(x_{t+1}|x_t)$ are given by

$$m_{t+1|t}^{reh(mk)} = \rho x_t + \mu^{(2)} + \pi \left(\mu^{(1)} - \mu^{(2)}\right) = \rho x_t + E(\mu_t), \tag{29}$$

$$.\left(\sigma_{t+1|t}^{reh(mk)}\right)^2 = \sigma^2 + V(\mu_t) \tag{30}$$

Using (23), $m_{t+1|t}^{reh(mk)}$, expressed in terms of $x_{t-1,}$ the realized state of μ_t, $\mu^{(j)}$ $j = 1, 2$, and e_t, can take one of the two values:

$$m_{t+1|t}^{reh(mk,j)} = \rho^2 x_{t-1} + \rho\mu^{(j)} + \rho e_t + E(\mu_t),\ j = 1, 2. \tag{31}$$

7.2.1 News When the Process Evolves According to a Markov Chain

Allowing for the Markov component in the process driving outcomes, in (23), implies that the news variable, in (21), is given by:

$$\eta_t^{(mk)} = X_t - E(X_t|x_{t-1}) = \mu_t - E(\mu_t) + \varepsilon_t. \tag{32}$$

This shows that the news comprises *both* the realization of ε_t (e_t) and the value that μ_t takes, relative to its expectation, $E(\mu_t)$. Depending on the realized state at t, the realization of $\eta_t^{(mk)}$ can take one of the two values:

$$\begin{aligned} n^{(mk,j)} &= \mu^{(j)} - E(\mu_t) + e_t, \\ &= \left(\mu^{(j)} - \mu^{(2)}\right) - \pi(\mu^{(1)} - \mu^{(2)}) + e_t \quad j = 1, 2. \end{aligned} \tag{33}$$

When $e_t > 0$, which GS's time-invariant specification would characterize as "good news," the news could be either good or bad, in the sense that $n_t^{(mk,j)} > 0$ or $n_t^{(mk,j)} < 0$, respectively, depending on the values of $\mu^{(1)}$ and $\mu^{(2)}$, the realized state at t, and the realization of ε_t. For example, suppose that $\mu^{(1)} > \mu^{(2)}$, $\mu^{(j)} = \mu^{(1)}$, and $e_t > 0$, then, from (33), $n^{(mk,1)} > 0$. However, $n^{(mk,2)} < 0$, when $\mu^{(j)} = \mu^{(2)}$ and the magnitude of $e_t < \pi(\mu^{(1)} - \mu^{(2)})$.

7.3 A Mixture Specification of the Reference PDF

GS assumed that, like REH, the reference pdf is based on the "objective" process, in (23). However, they specified the reference class of outcomes as x_{t-1}. Using the same approach as in Section 7.2, the following lemma shows that allowing for the Markov component in the "objective" process implies that the reference pdf is a mixture of the four normal pdfs with the following means:[9]

$$m_{t+1|t}^{(mk,i,j)} = E(x_{t+1}|x_{t-1}, \mu_{t+1} = \mu^{(i)}, \mu_t = \mu^{(j)}) = \rho^2 x_{t-1} + \rho\mu^{(j)} + \mu^{(i)}, \quad (34)$$

Lemma 8 *Suppose that (23) characterizes the process driving x_{t+1}. Then, conditional on x_{t-1}, the REH-like reference pdf of x_{t+1}, denoted with $g^{ref(mk)}(x_{t+1}|x_{t-1})$, is the following mixture of the four normal pdfs specified in (34),*

$$g^{ref(mk)}(x_{t+1}|x_{t-1}) = \sum_{i,j=1}^{2} p_{ji}\pi_j f^{(mk,i,j)}(x_{t+1}|x_{t-1}, \mu_{t+1} = \mu^{(i)}, \mu_t = \mu^{(j)}), \quad (35)$$

where p_{ji}, $j, i = 1, 2$ are transition probabilities and $\pi_j = P(\mu_{t+1} = \mu^{(j)})$ for all t. Furthermore, the mean and variance of $g^{ref(mk)}(x_{t+1}|x_{t-1})$ are given by

$$m_{t+1|t}^{ref(mk)} = \rho^2 x_{t-1} + (1+\rho)E(\mu_t), \quad (36)$$

$$\left(\sigma_{t+1|t}^{ref(mk)}\right)^2 = (1+\rho^2)\left[\sigma^2 + V(\mu_t)\right] + 2\rho\left\{E\left(\mu_{t+1}\mu_t\right) - \left[E(\mu_t)\right]^2\right\}. \quad (37)$$

Moreover,

$$(1+\theta)\left(\sigma_{t+1|t}^{ref}\right)^2 > \theta\left(\sigma_{t+1|t}^{reh}\right)^2 \quad (38)$$

holds, for any values of the model parameters $\left(\theta, \rho, \mu^{(1)}, \mu^{(2)}, p_{12,}p_{21}\right)$

Proof in Online Appendix A.

7.4 Overreaction

The normality of the "objective" and reference mixtures of pdfs and (38) ensures that Proposition 4 holds for the mixture specification of DE. Moreover, analogously to GS's time-invariant specification, DE's REH-like specification of the reference pdf that includes a mMarkov component is tantamount to

[9] Because the explicit expression for $\left(\sigma^{(mk,i,j)}\right)^2$ plays no role in our argument, we omit it to save space. The derivation of this expression is analogous to that for $\left(\sigma_{t+1|t}^{ref(mk)}\right)^2$ in (37) below.

assuming that the participants' overreaction, *relative to the REH forecast*, can be represented with *the revision of the REH forecast*. We state this conclusion and that of the previous section with a proposition:

Proposition 9 *Suppose that an economist assumes that while the reference pdf is based on the "objective" normal pdf, which underpins REH, in assessing an event's $x_{t+1} \in A$'s representativeness, participants consider x_{t-1} the reference class of outcomes. Then, DE overreacts to good (bad) news, in the sense that when $n_t > 0$ ($n_t < 0$), $m^{de}_{t+1|t} - m^{reh}_{t+1|t} > 0$ ($m^{de}_{t+1|t} - m^{reh}_{t+1|t} < 0$), where n_t and the conditional means are specified for either the time-invariant pdfs implied by GS's specification or the mixtures implied by the Markov specification.*

Proof in Online Appendix A.

8 The Irregularity of DE's Overreaction in Pre-DE Behavioral Models

GS (pp. 137-152) argue that their specification of the reference pdf, and thus of DE, formalizes Kahneman and Tversky's findings in the Linda-like experiments in a variety of contexts. However, as we discussed in Section 3, Kahneman and Tversky (1972, p. 431) emphasized that there appears to be no theoretical basis for specifying the reference class that participants might consider relevant in assessing the representativeness of uncertain events.

Moreover, GS's specification of the reference pdf, and thus of DE, as being based on the "objective" pdf underpinning REH appears to be at odds with behavioral economists' compelling empirical findings that participants' forecasts deviate from the predictions of the "objective" process driving outcomes. As Barberis, *et al.* (1998, p. 318) have asserted, "If our model is to generate the [pattern] of returns documented in the empirical studies, the investor *must* be using the wrong model to form expectations" (p.318, emphasis added). We show here that once we acknowledge the relevance of these findings, DE no longer implies the regularity of overreaction. Depending on the values of the parameters of both the REH and empirically-based reference pdfs, as well as the realizations of x_t, DE overreacts in some periods and underreacts in other periods.

8.1 Empirically-Based Specification of the Reference PDF

One of the main psychological mechanisms underpinning Barberis, *et al.*'s (1998) specification of participants' forecasts is the representativeness heuristic. They argue (p. 317) that Kahneman and Tversky's empirical findings and others do not provide a basis for specifying how the news drives the overreaction or underreaction of participants' forecasts. This assessment of empirical evidence stands in contrast to GS's Assumption 5 that owing to the "distorting" influence of the representativeness heuristic, the direction and magnitude of the error market participants commit are *solely and precisely* related to the news implied by the REH forecast.

Appealing to an extensive review of empirical evidence, Barberis, *et al.* (p. 318) argue that the "distortion" of participants' forecasts of stock returns arises primarily from participants' not basing their forecasts on the "objective" process driving earnings, as specified by an economist's model. While an economist's model assumes that earnings evolve according to a random walk, the investor "thinks that the world moves between two 'states' or 'regimes' and that there is a different model governing earnings in each regime." Barberis, *et al.* formalize this assumption with the two-state stationary Markov chain.

8.2 A Behavioral Markov (BM) Specification of DE

To facilitate comparison with GS's REH-like specification of the reference pdf, we use an AR(1) process, in (14), to characterize the 'objective" process driving x_t, which we restate here for convenience,

$$x_{t+1} = \rho x_t + \mu + \varepsilon_{t+1}, \tag{39}$$

However, we adapt Barberis *et al.'s* assumption that participants "think that the world moves between two 'states'" to our context by specifying their forecasts as being based on the following "wrong" version of the "objective" process in (39):

$$x_{t+1} = \rho_t^{(b)} x_t + \mu_{t+1}^{(b)} + \varepsilon_{t+1}, \tag{40}$$

where "b" in the superscript denotes that $\rho_t^{(b)}$ and $\mu_t^{(b)}$ specify the BM model. Each of them evolves according to a Markov chain, which switches between two states, $\rho^{(b,i)}$, and $\mu^{(b,i)}$ $i = 1, 2$ with the transition probabilities, $p_{12,}$ and q_{12}, respectively and $\varepsilon_t \sim iidN(0, \sigma^2)$. To simplify the presentation, we assume that $\rho_t^{(b)}$ and $\mu_t^{(b)}$ are independent. It follows from (40) that, while x_t and μ_{t-i} for $i = 0, 1...$ are dependent, x_t and μ_{t+i} for $i = 1, 2....$are independent. We also assume that, while x_t and $(\mu_{t-i}^{(b)},\ \rho_{t-1-i}^{(b)})$ for $i = 0, 1...$ are dependent, x_t and $(\mu_{t+i}^{(b)},\ \rho_{t+1-i}^{(b)})$ for $i = 1, 2...$are independent. Analogously to the specification in (23), we also assume that $\rho_t^{(b)}$ and $\mu_t^{(b)}$ are stationary Markov chains.

8.2.1 A Markov Specification of the Behavioral Reference PDF

We assume that the forecasting model, in (40), characterizes the reference process that participants consider relevant in assessing $x_{t+1} \in A$'s representativeness of x_t. A proof analogous to that of Lemma 8 shows that the reference pdf is is a mixture of the four normal pdfs:

$$g^{ref(b)}(x_{t+1}|x_t) = \sum_{i,j=1}^{2} \pi_\rho^{(j)} \pi_\mu^{(i)} f^{(b,i,j)}(x_{t+1}|x_t, \rho_t^{(b)} = \rho^{(b,j)}, \mu_{t+1}^{(b)} = \mu^{(b,i)}), \tag{41}$$

where $\pi_\rho^{(j)}$ and $\pi_\mu^{(i)}$, $j, i = 1, 2$, are components of the respective stationary distributions. Furthermore, the conditional mean and variance of $g^{ref(b)}(x_{t+1}|x_t)$ are given by

$$m_{t+1|t}^{ref(b)} = E(\rho_t^{(b)})x_t + E(\mu_t^{(b)}) \tag{42}$$

$$= E(\rho_t^{(b)})\rho x_{t-1} + \rho\mu + E(\mu_t^{(b)}) + \rho e_t, \tag{43}$$

$$\left(\sigma_{t+1|t}^{ref(b)}\right)^2 = \sigma^2 + V(\rho_t^{(b)})E(x_t^2) + V(\mu_t) \tag{44}$$

where $E(\rho_t^{(b)}), E(\mu_t^{(b)}), V(\rho_t^{(b)})$, and $V(\mu_t^{(b)})$ are the means and variances, which are specified analogously to (24) and (25).

8.2.2 A Behavioral Markov DE May Overreact or Underreact

Because the mixture in (41) is a normal pdf, and the expressions (17) and (44), show that $\left(\sigma_{t+1|t}^{ref(b)}\right)^2 > \left(\sigma_{t+1|t}^{reh(gs)}\right)^2$, Proposition 4 holds. Thus, the diagnostic expectation implied by GS's specification of the time-invariant REH pdf, and the BM specification of the reference pdf is given by

$$m_{t+1|t}^{de(b)} = m_{t+1|t}^{reh(gs)} + \gamma^{(b)} \left(m_{t+1|t}^{reh(gs)} - m_{t+1|t}^{ref(b)}\right) \tag{45}$$

$$= m_{t+1|t}^{reh(gs)} + \gamma^{(b)} \left\{\left[\rho - E(\rho_t^{(b)})\right] x_t + \mu - E(\mu_t^{(b)})\right\} \tag{46}$$

where, from (11), (17) and (44), $\gamma^{(b)} = \theta \frac{\sigma^2}{\sigma^2 + (1+\theta)\left[V(\rho_t^{(b)})E(x_t)^2 + V(\mu_t)\right]}$.
The expression in (46) shows that, according to the BM specification, whether DE overreacts, relative to its REH counterpart, depends on whether $\left[\rho - E(\rho_t^{(b)})\right] x_t + \mu - E(\mu_t^{(b)}) > 0$. The following lemma states this point explicitly:

Lemma 10 *Suppose that the specification of DE based on (39) and (40) characterizes how the representativeness heuristic leads participants away from forecasting according to REH. Letting*

$$\rho - E(\rho_t^{(b)}) > 0$$

implies that if and only if

$$x_t > \frac{E(\mu_t^{(b)}) - \mu}{\rho - E(\rho_t^{(b)})}.$$

holds, DE overreacts, that is, $m_{t+1|t}^{reh(gs)} > m_{t+1|t}^{ref(b)}$ *holds. Conversely, letting*

$$\rho - E(\rho_t^{(b)}) < 0$$

implies that if and only if

$$x_t < \frac{E(\mu_t^{(b)}) - \mu}{\rho - E(\rho_t^{(b)})}.$$

holds, DE underreacts, that is, $m^{reh(gs)}_{t+1|t} < m^{ref(b)}_{t+1|t}$ *holds.*

To be sure, the BM representation of participants' understanding of change is quite restrictive. However, relaxing GS's REH-like specification of the reference pdf illustrates a more general point. Once we acknowledge the relevance of behavioral economists' findings that participants' forecasts deviate from the predictions of an economist's model, DE no longer implies the regularity of overreaction. Depending on the values of the model parameters of both the REH and reference pdfs, $\left(\rho, \mu, \sigma^2, \rho^{(b,i)}, \mu^{(b,i)}, \pi^{(i)}_{\rho}, \pi^{(i)}_{\mu}\right)$, $i = 1, 2$, and the realizations of x_t, DE overreacts in some periods and underreacts in others periods.

9 Coibion and Gorodnichenko's Econometric Framework

Coibion and Gorodnichenko (CG, 2015) proposed a new regression-based framework for testing the predictions of sticky and noisy REH models based on survey data on participants' forecasts of macroeconomic variables. CG (pp. 2651, 2653) point out that the theoretical structure of these REH specifications of forecasts "map" directly onto the following regression relationship between *ex post* forecast errors and forecast revisions:

$$x_{t+h} - F_t(x_{t+h}) = \alpha + \beta \left[F_t(x_{t+h}) - F_{t-1}(x_{t+h})\right] + \upsilon_{t+1}, \qquad (47)$$

where $F_t(x_{t+h})$ denotes participants' time-t forecast of a variable x_t at time $t+h$, and υ_{t+1} is the error term implied by a theoretical specification of participants' forecasts. To simplify the presentation (without a loss of generality), here we set $h = 1$.

The CG regression is ideally suited not only for examining whether participants' forecasts, as measured by the survey data, are inconsistent with so-called full information rational expectations (FIRE), but also for determining which departures from FIRE appear to be consistent with that data. As CG emphasize:

> [Our] approach possesses multiple advantages over traditional tests of ...FIRE. First, we rely on the predictions of theoretical models of

> information rigidities to guide our choice of the relevant regressors. Second, models of information rigidities make specific predictions about the sign of the coefficient on forecast revisions, so that our specification provides guidance not only about the null of FIRE but also about alternative models (p. 2645).

Here, we rely on the CG regression, in (47), to test departures from FIRE implied by the REH-like Markov and Behavioral Markov specifications of DE, which we formulated in Sections 7.1 and 8.2. We also derive predictions of the early behavioral-finance specifications, such as that proposed by Barberis, *et al.* (1998), as well as of FIRE implied by a Markov specification of the "objective" process, in (23). Because the theoretical structure of these specifications does not directly match the relationship between participants' forecast error and forecast revision, in (47), we derive their predictions for α and β by relying on GS's assumption that forecasts can be represented as normally distributed.

10 Predictions of DE's Specifications for the CG Regression

Hypothesizing that the mean of the diagnostic pdf, $m^{de}_{t+1|t}$ in (8), represents participants' time-t forecasts of x_{t+1}, CG's regression, in (47), can be written as the relationship between DE's forecast error, $x_{t+1} - m^{de}_{t+1|t}$, and the revision of diagnostic expectations between $t-1$ and t:

$$x_{t+1} - m^{de}_{t+1|t} = \alpha + \beta(m^{de}_{t+1|t} - m^{de}_{t+1|t-1}) + \upsilon_{t+1}. \tag{48}$$

On the other hand, provided that the REH and reference pdfs are normal, the *theoretical* specification implied by (12) relates DE's forecast error, $x_{t+1} - m^{de}_{t+1|t}$, to $m^{reh}_{t+1|t} - m^{ref}_{t+1|t}$ as well as the REH forecast error, $f\epsilon^{reh}_{t+1|t} = x_{t+1} - m^{reh}_{t+1|t}$:

$$x_{t+1} - m^{de}_{t+1|t} = -\gamma\left(m^{reh}_{t+1|t} - m^{ref}_{t+1|t}\right) + f\epsilon^{reh}_{t+1|t}, \tag{49}$$

In order to derive the predictions of *any* specification of DE for α, β, in

(48), we note that (12) implies that

$$m^{de}_{t+1|t} - m^{de}_{t+1|t-1} = (1+\gamma)\left(m^{reh}_{t+1|t} - m^{reh}_{t+1|t-1}\right) - \gamma\left(m^{ref}_{t+1|t} - m^{ref}_{t+1|t-1}\right) \quad (50)$$

We also note that the outcome x_{t+1}, as well as the conditional means $m^{de}_{t+1|t}$, $m^{reh}_{t+1|t}$, and $m^{ref}_{t+1|t}$ in (48) and (50) are the time-t realizations of random variables, which we denote with $X_{t+1}, M^{de}_{t+1|t}$, $M^{reh}_{t+1|t}$, and $M^{ref}_{t+1|t}$, respectively.

To simplify the notation, we define the variables in (49) and (50) as follows:

$$Y_{t+1} = X_{t+1} - M^{de}_{t+1|t},\ Z_{1,t} = M^{de}_{t+1|t} - M^{de}_{t+1|t-1}, \quad (51)$$

$$Z_{2,t} = M^{reh}_{t+1|t} - M^{ref}_{t+1|t},\ FE^{reh}_{t+1|t} = X_{t+1} - M^{reh}_{t+1|t}. \quad (52)$$

According to the time-invariant AR(1) model, in (14), assuming that $\varepsilon_t \sim iidN(0,\sigma^2)$ immediately implies that all of the variables in (51) and (52) are normally distributed. However, when μ_t evolves according to a Markov chain, each of the variables in (51) is a mixture of normal pdfs. For example, an analogous argument to that in the proof of Lemma 7 shows that the pdf of $Z_{2,t} = M^{reh}_{t+1|t} - M^{ref}_{t+1|t} = M^{reh}_{t+1|t} - M^{reh}_{t+1|t-1}$is a mixture with two normal components, and thus it is normally distributed. We state this with a lemma.

Lemma 11 *Suppose that the REH-like Markov specification of DE, in Section 7.1, represents participants' forecasts. Then, the pdf of $Z_{2,t}$, denoted with $g^{reht(mk)}(z_{2,t})$, is the mixture of the two conditional normal pdfs:*

$$g^{reh(mk)}(z_{2,t}) = \pi f^{(mk,1)}(z_{2,t}|\mu_t = \mu^{(1)}) + (1-\pi) f^{(mk,2)}(z_{2,t}|\mu_t = \mu^{(2)}), \quad (53)$$

Proof in Online Appendix A.

The normality and stationarity of Y_{t+1}, $Z_{1,t}$, $Z_{2,t}$, and $Z_{3,t+1}$ provide a straightforward way to express the predictions of the REH-like and behavioral specifications of DE, in Sections 5, 7.1 and 8.2 for the coefficients of the CG regression, in (48). The standard expression for the conditional mean of jointly normal variables expresses these predictions in terms of the moments of the variables in (51):

$$
\begin{aligned}
E(Y_{t+1}|Z_{1,t}) &= -\gamma E(Z_{2,t}|Z_{1,t}) + E(FE^{reh}_{t+1|t}|Z_{1,t}) \\
&= -\gamma \left\{ E(Z_{2,t}) + \frac{Cov(Z_{1,t}, Z_{2,t})}{V(Z_{1,t})} \left[Z_{1,t} - E(Z_{1,t}) \right] \right\} \qquad (54) \\
&\quad + E(FE^{reh}_{t+1|t}|Z_{1,t})
\end{aligned}
$$

where $0 < \gamma < 1$ and $Cov(\cdot, \cdot)$ is the covariance.
While $E(FE^{reh}_{t+1|t}|Z_{1,t}) = 0$ for the time-invariant REH-like specification of DE, allowing for change with a Markov chain renders $E(FE^{reh}_{t+1|t}|Z_{1,t}) \neq 0$. Thus, the predictions of the Markov REH-like and behavioral specifications of DE for the coefficients in the CG regression in (48) must also take into account that

$$
E(FE^{reh}_{t+1|t}|Z_{1,t}) = E(FE^{reh}_{t+1|t}) + \frac{Cov(Z_{1t}, FE^{reh}_{t+1|t})}{V(Z_{1,t})} \left[Z_{1t} - E(Z_{1t}) \right]. \qquad (55)
$$

Denoting with the superscript "*de*" that the expressions for the coefficients are implied by a specification of DE, we summarize the argument in this section with a proposition:

Proposition 12 *Suppose that the REH and reference pdfs can be represented with the mixtures of normal pdfs, which arise from stationary Markov chains, and DE represents participants" forecasts. The following expressions characterize the predations of any such DE specification for the coefficients in the CG regression (48):*

$$
\begin{aligned}
\alpha^{de} &= E(FE^{reh}_{t+1|t}) \qquad (56) \\
&\quad -\gamma E(Z_{2t}) + \frac{Cov(Z_{1t}, FE^{reh}_{t+1|t}) - \gamma Cov(Z_{1t}, Z_{2t})}{V(Z_{1,t})} E(Z_{1t}) \\
\beta^{de} &= \frac{Cov(Z_{1t}, FE^{reh}_{t+1|t}) - \gamma Cov(Z_{1t}, Z_{2t})}{V(Z_{1,t})}. \qquad (57)
\end{aligned}
$$

In the remainder of this Section, we apply this proposition to derive the predictions of the REH-like and behavioral Markov specifications of DE, formulated in Sections 5, 7, and 8.2.

10.1 Predictions of the REH-like Specifications of DE

The REH-like time-invariant and Markov specifications of the DE assume that $m^{ref}_{t+1|t} = m^{reh}_{t+1|t-1}$. Thus, according to Proposition 12, their predictions for α and β in (48) involve the moments of

$$\begin{aligned} Z_{2,t} &= M^{reh}_{t+1|t} - M^{reh}_{t+1|t-1}, \\ Z_{1,t} &= M^{de}_{t+1|t} - M^{de}_{t+1|t-1} \\ &= (1+\gamma^{(mk)})\left(M^{reh}_{t+1|t} - M^{reh}_{t+1|t-1}\right) - \gamma^{(mk)}(M^{reh}_{t+1|t-1} - M^{reh}_{t+1|t-2}), \\ FE^{reh}_{t+1|t} &= X_{t=1} - E\left(FE^{reh}_{t+1|t}|Z_{1,t.}\right). \end{aligned} \tag{58}$$

The following corollary to Proposition 12 derives the relevant moments of these variables and the predictions of the REH-like Markov specification of the reference pdf, and thus of DE, denoted with the superscript "$de(mk)$."

Corollary 13 *Suppose that DE's REH-like Markov specification, in Section 7.1, characterizes participants' forecasts. Then,*

1. $\alpha^{de(mk)} = 0$ *at all* t.
2. $\beta^{de(mk)} < 0$ *if* $1 < p_{12} + p_{21} < (\frac{1}{\gamma^{(mk)}\rho} - 1)\left[(1+\gamma^{(mk)}(1-\rho)\right]$.
   (a) *If this condition is not satisfied, there are values of the model parameters,* $(\mu^{(1)}, \mu^{(2)}, p_{12}, p_{21}, \rho, \sigma^2)$ *for which* $\beta^{de(mk)} > 0$
3. *However, the sign and the magnitude of* $\beta^{de(mk)}$ *are unchanging over time.*

Proof in Online Appendix A.

According to Proposition 9, the REH-like Markov specification implies overreaction, that is, $m^{de}_{t+1|t} > m^{reh}_{t+1|t}$. However, Corollary 13 reveals that, because $Cov(Z_{1,t}, FE^{reh}_{t+1|t}) \neq 0$, the slope coefficient in the CG regression implied by that theoretical specification may be either positive or negative. Nevertheless, the REH-like Markov specification implies an unambiguous testable prediction: the constant equal to zero and the sign and magnitude that the slope takes remain unchanging over time, thereby suggesting either the regularity of overreaction or underreaction.

10.1.1 Predictions of GS's Time-Invariant Specification

The proof of Corollary 13 shows that constraining $\mu_t = \mu$, for all t implies that $E(FE^{reh}_{t+1|t}|Z_{1,t}) = 0$ $(Cov(Z_{1,t}, FE^{reh}_{t+1|t} = 0)$, which renders the predicted sign of the slope negative. We state this with a corollary.

Corollary 14 *GS's REH-like time-invariant specification of DE, in Section 5, implies the following predictions for the coefficients of the CG regression, in (48):*

1. $\alpha^{de(gs)} = 0$ at all t.

2. $\beta^{de(gs)} < 0$ *and its magnitude is unchanging over time.*

Bordalo *et al.* (2020) assume that the time-invariant REH-like specification of DE can explain survey data, which implies that $\beta^{de(gs)} < 0$. However, Corollary 13 shows that allowing for the Markov component in the process driving x_t may substantially alter this prediction: while there are values of the model parameters for which the slope of the CG regression is negative, there are also values of those parameters for which it is positive.

10.1.2 Predictions When the Markov Chain Persists in a Regime

According to Corollary 14, although the Markov specification of DE allows for change in how participants forecast outcomes, it nonetheless predicts that the constant term in the CG regression is unchanging over time. However, the persistence of a Markov chain in one state for a prolonged period of time might cause structural break(s) in α. For example, Engel and Hamilton (1990) formalized such persistence with a two-state Markov chain in which they constrained the probabilities of switching, p_{12} and p_{21}, to be small, implying that the process would be expected to stay in one state for a long time. During each of such subperiods, typically referred to as regimes, α would remain unchanged, but it would take a different sign and (generally) a different magnitude in one regime as compared with the other.

The following corollary states predictions of assuming regime persistence for the coefficients of the CG regression, denoted with $\alpha^{de(mp)}$ and $\beta^{de(mp)}$;

Corollary 15 *Suppose that the transition probabilities p_{12} and p_{21} are sufficiently small, so that the process driving μ_t, in (23), may stay in one of the regimes for a long period of time of time, and yet undergo intermittent structural breaks that can be detected by an econometric procedure. Then, the REH-like Markov specification of DE in Section 7.1 implies the following testable predictions for the coefficients of the CG regression, in (48), denoted with $\alpha^{de(mp)}$ and $\beta^{de(mp)}$:*

1. *$\alpha^{de(mp)}$ switches the sign (from positive to negative or vice versa) when the transition from $\mu_t = \mu^{(i)}$ to $\mu_t = \mu^{(j)}$ occurs, $i, j = 1, 2$, $i \neq j$.*
    (a) *However, the sign and magnitude of $\alpha^{de(mp)}$ are the same whenever μ_t returns to and persists within one of the regimes:*
2. *$\beta^{de(mp)} < 0$ and its magnitude is unchanging over time, regardless of whether μ_t persists in one of the regimes.*

Proof in Online Appendix.

This corollary shows that while the regime persistence may cause the constant term $\alpha^{de(mp)}$ to undergo intermittent structural breaks, this change is constrained in a way that can easily be tested. Although $\alpha^{de(mp)}$ differs across the two regimes, its positive (or negative) sign as well as its magnitude are the same within each regime. For example, as we show in the proof of this corollary, whenever a Markov chain is in regime in which $\left[\mu^{(i)} - E(\mu_t)\right] > 0$, $(\left[\mu^{(i)} - E(\mu_t)\right] < 0)$, $\alpha^{de(mp)} > 0$ $(\alpha^{de(mp)} < 0)$.

10.2 Predictions of the Behavioral Markov Specification of DE

The BM specification of DE, in Section 8.2, defines the variables in (51) and (52). The following corollary derives the moments of these variables and uses

Proposition 12 to derive predictions for the coefficients of (48), denoted with $\alpha^{de(b)}$ and $\beta^{de(b)}$:

Corollary 16 *DE's Behavioral Markov specification, in Section 8.2, implies the following predictions for the coefficients of the CG regression, in (48):*

1. *Either* $\beta^{de(b)} < 0$ *if* $\rho > \frac{\gamma^{(b)}E(\rho_t^{(b)}}{1+\gamma^{(b)}}$ *, or* $\beta^{de(b)} > 0$*, if* $\rho < \frac{\gamma^{(b)}E(\rho_t^{(b)}}{1+\gamma^{(b)}}$.

2. Either $\alpha^{de(b)} > 0$ *if* $\left\{\frac{\mu\rho}{1-\rho}\left[\rho - E(\rho_t^{(b)})\right] + \left[\mu - E(\mu_t^{(b)})\right]\right\}$
$\times \left[\rho\left(1+\gamma^{(b)}\right) - \gamma^{(b)}E(\rho_t^{(b)})\right] < 0,$
or $\alpha^{de(b)} < 0$*, if* $\left\{\frac{\mu\rho}{1-\rho}\left[\rho - E(\rho_t^{(b)})\right] + \left[\mu - E(\mu_t^{(b)})\right]\right\}$
$\times \left[\rho\left(1+\gamma^{(b)}\right) - \gamma^{(b)}E(\rho_t^{(b)})\right] > 0$

3. *However, the signs and magnitudes of* $\alpha^{de(b)}$ *and* $\beta^{de(b)}$ *are unchanging over time.*

11 Predictions of the Pre-DE Behavioral and REH Specifications of Participants' Forecasts

Here, we adopt the approach of the preceding section to derive predictions of the pre-DE behavioral specifications of participants' forecasts, such as Barberis *et al.*'s, as well as predictions of the REH-implied specifications. To this end, we note that any specification of forecasts, denoted with $m_{t+1|t}^{for}$, satisfies the following relationship:

$$x_{t+1} - m_{t+1|t}^{for} = \left(m_{t+1|t}^{reh} - m_{t+1|t}^{for}\right) + f\epsilon_{t+1|t}^{reh}, \tag{59}$$

Comparing this with (49) shows that by setting $m_{t+1|t}^{ref} = m_{t+1|t}^{for}$ and $\gamma = -1$, Proposition 12 can be formally recast as the statement of predictions of any specification of participants forecasts for the coefficients of the CG regression, (47). To this end, we redefine the variables in (51) and (52) as

$$Y_{t+1} = X_{t+1} - M_{t+1|t}^{for},\ Z_{1,t} = M_{t+1|t}^{for} - M_{t+1|t-1}^{for}, Z_{2,t} = M_{t+1|t}^{reh} - M_{t+1|t}^{for} \tag{60}$$

Noting that $FE_{t+1|t}^{reh} = X_{t+1} - M_{t+1|t}^{reh}$ and setting $\gamma = -1$ in Proposition 12, the following proposition states predictions of any specification of participants' forecasts, denoted with α^{for} and β^{for}, in terms of the moments of these variables.

Proposition 17 *Suppose that the process driving outcomes and participants' forecasts can be represented with mixtures of normal pdfs arising from stationary Markov chains. The following expressions characterize the predictions of any such specification for the coefficients in the CG regression (47):*

$$\alpha^{for} = E(FE^{reh}_{t+1|t}) + E(Z_{2t}) + \frac{Cov(Z_{1t}, FE^{reh}_{t+1|t}) + Cov(Z_{1t}, Z_{2t})}{V(Z_{1,t})} E(Z_{1t}) \tag{61}$$

$$\beta^{for} = \frac{Cov(Z_{1t}, FE^{reh}_{t+1|t}) + Cov(Z_{1t}, Z_{2t})}{V(Z_{1,t})}. \tag{62}$$

11.1 Predictions of Barberis *et al.*'s (1998) pre-DE Specification

The proof of Corollary 16 derives the moments underlying predictions in (61) and (62), denoted with $\alpha^{(beh)}$ and $\beta^{(beh)}$, which we state with the corollary to Proposition 17:

Corollary 18 *Suppose that, while the "objective" process driving outcomes is time-invariant, in (39), participants's forecasts are based on process, in (40), the mean of which evolves according to a two-state Markov chain. Such non-REH specification of forecasts implies the following predictions for the coefficients of the CG regression, in (47):*

1. *Either* $\beta^{(beh)} < 0$ *if* $\rho < \frac{\gamma^{(b)} E(\rho_t^{(b)}}{1+\gamma^{(b)}}$ *, or* $\beta^{(beh)} > 0$*, if* $\rho > \frac{\gamma^{(b)} E(\rho_t^{(b)}}{1+\gamma^{(b)}}$.

2. Either $\alpha^{(beh)} > 0$ *if* $\left\{\frac{\mu\rho}{1-\rho}\left[\rho - E(\rho_t^{(b)})\right] + \left[\mu - E(\mu_t^{(b)})\right]\right\}$
$\times \left[\rho\left(1+\gamma^{(b)}\right) - \gamma^{(b)} E(\rho_t^{(b)})\right] > 0$,
or $\alpha^{(beh)} < 0$*, if* $\left\{\frac{\mu\rho}{1-\rho}\left[\rho - E(\rho_t^{(b)})\right] + \left[\mu - E(\mu_t^{(b)})\right]\right\}$
$\times \left[\rho\left(1+\gamma^{(b)}\right) - \gamma^{(b)} E(\rho_t^{(b)})\right] < 0$

3. *However, the signs and magnitudes of* $\alpha^{(beh)}$ *and* $\beta^{(beh)}$ *are unchanging over time.*

11.2 Predictions of FIRE for the CG Regression

The time-invariant aggregate REH specifications tested by CG (2015) predict that $\alpha = 0$ and $\beta > 0$, if information rigidities are present. CG (p. 2651) interpreted this prediction as indicating that participants' forecasts, though consistent with REH, deviate from FIRE, owing to noisy information about the state of the economy. However, such an interpretation of the coefficients of the CG regression overlooks the implications for the regression's slope of recognizing that the process driving outcomes undergoes change.

Using Proposition 17, we show here that representing participants' forecasts with FIRE in a model that specifies change with a Markov chain exhibiting even moderate regime persistence ($p_{12} + p_{12} < 1$) implies $\beta > 0$ .When $p_{12} + p_{12} > 1$, FIRE predicts that $\beta < 0$, which is the same prediction as that implied by GS's time-invariant specification of DE in Corollary 14.

Hypothesizing that FIRE is based on the Markov specification, (23), defines $M^{for}_{t+1|t} = M^{fire}_{t+1|t}$, which, from (60), sets $Z_{2,t} = 0$ for all t and implies that $Y_{t+1|t}$ and $Z_{1,t}$ are given by

$$\begin{aligned} Y_{t+1|t} &= FE^{fire}_{t+1|t} = X_{t+1} - M^{fire}_{t+1|t} = X_{t+1} - E\left(X_{t+1}|x_t\right) \\ &= \left[\mu_{t+1} - E\left(\mu_t\right)\right] + \varepsilon_{t+1}, \qquad (63) \\ Z_{1,t} &= M^{fire}_{t+1|t} - M^{fire}_{t+1|t-1} = \rho\left[\mu_t - E\left(\mu_t\right)\right] + \rho\varepsilon_t \qquad (64) \end{aligned}$$

According to an argument analogous to Lemma 7, the pdfs of $Y_{t+1|t}$ and $Z_{1,t}$, are mixtures of normal pdfs. Thus, Proposition 17 implies the following predictions for the coefficients of the CG regression, denoted with $\alpha^{fire(mk)}$ and $\beta^{fire(mk)}$, which we state with a corollary:

Corollary 19 *Suppose that FIRE implied by the model (23) represents participants' forecasts. Such specification implies the following predictions for the coefficients of he CG regression, (47):*

1. $\alpha^{fire(mk)} = 0$ *at all* t.

2. *Either* $\beta^{fire(mk)} > 0$ *if* $p_{12} + p_{21} < 1$, $\beta^{fire(mk)} < 0$ *if* $p_{12} + p_{21} > 1$.[10]

   (a) *However, the sign and magnitude of* $\beta^{fire(mk)}$ *are unchanging over time.*

Proof in Online Appendix A.

12 Summary of the Predictions

Table 1 summarizes predictions for the CG regression of the five behavioral specifications of participants' forecasts, including four implied by the DE approach, as well as three alternative REH-implied specifications, and refers to the respective corollaries and CG (2015) for their derivations.

Table 1: Predictions of Theoretical Specifications of Participants' Forecasts

Model	Prediction for α	Prediction for β
REH-like Specifications of DE		
A: GS's Time-Invariant	$\alpha = 0$	$\beta < 0$
B: Involving a Markov Component	$\alpha = 0$	$\beta < 0$ or $\beta > 0$
C: Assuming Regime Persistence	$\alpha > 0$ and $\alpha < 0$	$\beta < 0$
D: Behavioral Markov Specification of DE	$\alpha > 0$ or $\alpha < 0$	$\beta < 0$ or $\beta > 0$
E: Pre-DE Behavioral Specification	$\alpha > 0$ or $\alpha < 0$	$\beta < 0$ or $\beta > 0$
REH-implied Specifications		
F: Time-Invariant FIRE	$\alpha = 0$	$\beta = 0$
G: FIRE Involving a Markov Component	$\alpha = 0$	$\beta < 0$ or $\beta > 0$
H: Noisy-Information	$\alpha = 0$	$\beta > 0$

Caption: A: Corollary 14, B: Corollary 13, C: Corollary 15, D: Corollary 16, E: Corollary 18, F: CG (2015), G: Corollary 19, H: CG (2015).

A finding of $\beta > 0$ in the CG regression has traditionally been interpreted as an "underreaction," and $\beta < 0$ as an "overreaction." However, once DE's reference pdf involves a Markov component (Model B), DE predicts CG estimates consistent with either under- or overreaction. Analogously for FIRE, if

[10] $\beta^{fire(mk)} = 0$ if $p_{12} + p_{12} = 1$. Because this case does not affect our conclusions in Section 13, we omit it from the corrollary.

the forecasted variable's law of motion involves a Markov component (model G), FIRE implies that either $\alpha = 0$ and $\beta < 0$, or $\alpha = 0$ and $\beta > 0$. Therefore, allowing for a Markov component in the process driving outcomes may render FIRE's prediction the same as that of DE with a Markov component or noisy-information REH model H.

13 Empirical Findings

Here, we test predictions in Table 1 by estimating individual CG regressions, based on survey data of inflation forecasts by 24 professionals.

13.1 Full-Sample Estimates of the CG Regression

We begin with the full-sample estimates, which the literature typically focuses on in assessing the empirical adequacy of theoretical specifications of expectations. Table B1 in Online Appendix B displays such estimates of the individual CG regressions. The estimates are based on data from 24 individuals in the Philadelphia Federal Reserve's Survey of Professional Forecasters with more than 50 observations of their three-quarter-ahead forecast revisions of the Gross Domestic Product's inflation.[11] Table 2 presents the summary of the tests of the coefficients of these individual regressions grouped across forecasters.

Table 2: Grouping of Individuals Based on Tests of Full-Sample Regressions' Estimates

Individuals	α	β	Consistent with
14/24	$\alpha = 0$	$\beta = 0$	Model F in Table 1
5/24	$\alpha = 0$	$\beta < 0$	Models A, B and G
2/24	$\alpha = 0$	$\beta > 0$	Models H, B and G
3/24	$\alpha \neq 0$	$\beta = 0$	No model in Table 1

Bordalo, *et al.* (2020) assert that "for individual forecasters the prevalent pattern is overreaction (p. 2779)." By contrast, we find that the estimates for

[11] This has been the most prominently studied variable and horizon in the literature examining aggregate-level survey data. As examples, see Coibion and Gorodnichenko (2012), Angeletos, Huo and Sastry (2021), and Bianchi, Ludvigson, and Ma (2021).

only five of the 24 individual regressions are consistent with $\alpha = 0$ and $\beta < 0$.[12] Even more surprising in Table 2 is the number of individual CG regressions that fail to reject time-invariant FIRE (model F). For 14 of the 24 forecasters, we cannot reject at even 10% that $\alpha = 0$ and $\beta = 0$. This is incongruent with the findings of CG (2015) and a number of other studies they cite that FIRE is inconsistent with individuals' forecasting. These interpretations of the empirical findings have, however, presumed the stationarity of the process driving outcomes and in how participants forecast them.

13.2 Time-Invariance of the Coefficients of the CG Regression

Models in Table 1 differ in a number of important respects. However, all of them rest on a common premise: the process driving outcomes and participants' revisions of their forecasting strategies can be represented with a stationary Markov chain. One of the central implications of our theoretical framework is that, although such representations do allow for *change* in the specification of individuals' forecasting strategies, they predict that the constant and slope coefficients in the CG regression *do not change* over time. Predictions of Models B, D, E, G, and H in Table 1 formalize this implication.

For example, according to Corollary 13, the REH-like specification of DE with a Markov component (Model B) predicts $\alpha = 0$ for all t and $\beta < 0$ or $\beta > 0$, depending on the values of transition probabilities and other parameters. However, under the stationarity assumption, Model B predicts that β must remain either positive or negative in both sign or magnitude for all t. Similarly, while the behavioral Markov specification (Model D) predicts that $\alpha > 0$ or $\alpha < 0$, and $\beta < 0$ or $\beta > 0$, it also predicts that the CG regression's coefficients are time-invariant both in sign and magnitude.

Model C is the only model among those in Table 1 predicting structural break(s) under the stationarity Assumption 6. These breaks could arise from intermittent switches between persistent regimes, as formalized with the low

[12] See Table B1 in Online Appendix B for the full-sample estimates and t-values for all 24 forecasters.

off diagonal transition probabilities. However, as Corollary 15 shows, while the regime persistence may cause the constant term α to undergo intermittent structural breaks, this change is constrained in a way that can easily be tested. Although α differs across the two regimes, its positive (or negative) sign and its magnitude are the same within each regime.

13.3 Structural Breaks in the Individual CG Regressions

According to Propositions 12 and 17, regardless of the specification of the "objective" and reference pdfs, the Assumption 6 of these processes' stationarity implies that the moments of the variables, in (51), (52), and (60), that underpin these predictions are time-invariant. Thus, subjecting the coefficients of the CG regression to tests of structural change provides a hitherto unexplored way to confront alternative models of expectations, including diagnostic expectations, with survey data on participants' forecasts. Moreover, the predictions of any of the specifications in Table 1 for the constant and the slope in the CG regression depend on different moments of the variables, in (51), (52), and (60). Whereas the prediction for α depends on the means and the covariances of these variables, the prediction for β depends only on the covariances. Thus, the tests for structural breaks require a procedure that allows the constant and slope of the CG regression to break at different times.

Consequently, we rely on the Multiplicative Indicator Saturation (MIS) procedure, which has been designed to detect breaks in the coefficients of the regression model at potentially different times. MIS is an extension of the Autometrics algorithm (Doornik, 2009) and the indicator saturation methods of Hendry, *et al.* (2008) and Castle, *et al.* (2015), whose consistency properties and appropriate size and power have already been demonstrated in previous studies under a range of conditions.[13] For an overview of the MIS methodology, see Online Appendix B.

By contrast, because the Bai-Perron (1998) procedure does not allow the constant and the slope to break at different times, on theoretical grounds, it

[13] For recent applications of MIS and further references, see Castle, *et al.* (2017).

is not suitable for testing the structural stability of predictions for α and β in the CG regression (47). Nonetheless, given its widespread use, we also report results of the Bai-Perron test.

Table 3: Structural Breaks in the CG Regressions Detected by MIS and Bai-Perron Procedures

Full Sample	Individuals	α Break	β Break	α or β	Bai-Perron
$\alpha = \beta = 0$	14/24	13/14	10/14[14]	14/14	7/14
$\alpha = 0; \beta < 0$	5/24	5/5	3/5	5/5	3/5
$\alpha = 0; \beta > 0$	2/24	2/2	1/2	2/2	2/2
$\alpha \neq 0; \beta = 0$	3/24	3/3	1/3	3/3	1/3

Columns 1 and 2 of Table 3 restate the results of the tests of the CG regressions grouped by number of individuals. Row 2 shows that 14 out of 24 individual CG regressions yielded a constant and the revision (slope) coefficient, not significantly different from zero, which is consistent with the time-invariant FIRE. Columns 3, 4, and 5 report the proportion of the individual regressions for those 14 forecasters that experience instability in the constant α, slope β, and either α or β. As column 5 row 2 shows, all of those 14 individual regressions experience a significant break in either the constant or the slope. The apparent prevalence of FIRE in the full-sample CG regressions is therefore completely overturned: the structural breaks detected by MIS in all of the individual regressions that did not reject time-invariant FIRE in the full sample are strongly inconsistent with that specification.

Of the 10 remaining individual regressions, all 10 experience a break in either α or β, as shown in column 5 of rows 3-5. This rejects all of our specifications in Table 1, with the possible exception of our persistent regime specification of DE (Model C).

In Table 3, column 4, there are nine individual regressions with a time-invariant β. However, as shown in Figures B1-B4, only five display a time-invariant $\beta < 0$, but all of those have the constant experiencing breaks while

[14] Two individual regressions in this row and one in the next row experienced a break in β within the first or last year of the sample, which could perhaps be viewed as outliers. All of these individual regressions however have breaks in the constant that are inconsistent with all theoretical specifications in Table 1

maintaining the same sign. However, Model C predicts α unchanging in both sign and magnitude, while a Markov chain persists in a regime and may switch intermittently between a fixed positive and a fixed negative value at the time the chain switches between the two regimes. Thus, the five regressions indicating a time-invariant $\beta < 0$ are inconsistent with Model C. In summary, MIS finds that all of our models in Table 1 are rejected based on the estimates of the 24 individual CG regressions.

The last column of Table 3 provides the summary for Bai-Perron tests relative to the full-sample estimates in columns 1 and 2. These results, like MIS, reveal that the full-sample estimates are misleading. Half of the cases that cannot reject time-invariant FIRE in the full sample experience breaks, and six of those experience a regime or regimes where either $\alpha \neq 0$ and/or $\beta \neq 0$ (see Table B2 in Online Appendix B for further details).. Similarly, three of the five full-sample estimates apparently consistent with time-invariant DE are no longer robust after Bai-Perron tests: they experience either sub-periods with a statistically significant $\alpha \neq 0$, or β loses significance or changes sign.

It is clear from Table 3 that the Bai-Perron procedure detected breaks in fewer individual regressions than did MIS. What is again surprising, however, is the significant number of regressions that still cannot reject FIRE after the Bai-Perron tests (eight of 24).[15] When contrasted with MIS, which rejects FIRE for all 24 regressions, this indicates a problem with Bai-Perron constraining α and β to break only simultaneously.

In particular, we find that more individual regressions experience a break in the constant than in the slope (23 vs. 15, as shown in Table 3). Moreover, as Figures B1-B4 show, breaks in the constant typically outnumber breaks in the slope for most individual regressions. This is consistent with our argument for using MIS: because the predictions for the CG constant depend, in part, on different moments of the relevant variables, the slope and constant may break at different points in time.

[15] Seven of these individuals experience no breaks, and one is found to have two significant breaks according to the Bai-Perron test, but we cannot reject $\alpha = 0$ and $\beta = 0$ in any of the three regimes.

13.4 Diversity of Forecasting Strategies

A number of papers have found significant diversity in how participants forecast outcomes.[16] A study that is of particular interest from the viewpoint of this paper is von Gaudecker and Wogrolly (2021) which documents significant diversity in households' beliefs about the stock market. They identify five separate groupings of forecasting strategies,[17] and then estimate panels of the CG regression for each group, based on the premise that individuals within these groups do not revise how they forecast outcomes.

However, our results suggest that the diversity of participants' forecasting strategies is substantially compounded by their revision of how they forecast outcomes at times and in ways that cannot be characterized with a stationary process, such as a Markov chain. As can be seen in Figures B1-B4 in Online Appendix B, the timing, direction, frequency, and magnitude of the breaks across individual regressions differ vastly.

14 Concluding Remarks: A Way Forward

The empirical inadequacy of all five alternative behavioral specifications of participants' forecasts casts doubts on the behavioral approach's core premise that market participants commit systematic, predictable errors, and that an economist can specify these errors precisely with a probability measure. As Lucas (1995, pp. 254-255) pointed out in his criticism of adaptive expectations, macroeconomic and finance models that violate Muth's (1961) hypothesis, as DE and other behavioral-finance models do, suffer from "glaring" inconsistency.[18] When an economist represents an individual's assessment of

[16]See Mankiw, *et al.* (2003), Reis (2020) and references therein.

[17]After determining the groupings, von Gaudecker and Wogrolly also examine differences across groups in terms of demographics, investment behavior, and response to returns and economic news.

[18]Lucas (1995, p. 255) recounts how the importance of ridding intertemporal models of such inconsistency persuaded macroeconomists to abandon the micro-founded models of the 1960s and embrace their REH counterparts. For an extensive discussion and formal illustration of this revolutionary development in macroeconomic theory, see Frydman and Goldberg (2007) and Frydman and Phelps (2013).

uncertainty about payoff-relevant outcomes in a way that is inconsistent with his own model's representation of this uncertainty, he contradicts his model's hypothesis: that it represents the actual uncertainty about these outcomes.

Lucas's argument that Muth's hypothesis should underpin the construction of logically coherent and empirically adequate macroeconomic and finance models appears persuasive. However, according to Muth's hypothesis, specifications of participants' forecasts crucially depend on an economist's model of the process driving outcomes. REH models' empirical difficulties, documented by behavioral economists and corroborated by our findings that the three REH-implied specifications based on stationary stochastic process are inconsistent with the survey data, suggest that economists must rethink how to represent this process.

One way to move beyond the prevailing approach to modeling is to specify how outcomes unfold over time and how participants forecast them with non-stationary stochastic processes. Because such representations would not imply that the coefficients of the CG regression are time-invariant, they might be consistent with structural breaks in those coefficients. If the process driving outcomes and individuals' forecast revisions could be represented with a single non-stationary process, according to Muth's hypothesis, participants would revise their forecasts at approximately the same time and in similar ways. However, as we documented in the preceding section, there is substantial diversity in the estimates and timing of breaks across the individual CG regressions, which suggests that individuals revise their forecasts at different times and in widely diverse ways.

It is nonetheless possible to reconcile the representation of such diversity, as well as the influence of both fundamental and psychological factors (such as market sentiment) in a consistent model that adheres to Muth's hypothesis. This, however, requires recognizing, as Knight (1921) argued, that the process driving outcomes undergoes unforeseeable change, which by definition cannot be represented *ex ante* with a single stochastic process, regardless of whether it is stationary or non-stationary.

REH implements Muth's hypothesis in a model that represents the uncertainty about payoff-relevant outcomes with a stationary stochastic process. This rules out, by design, behavioral economists' compelling findings that psychological and other non-fundamental factors have a substantial influence on participants' forecasts. Analogously, representing the change in the process driving outcomes with a non-stationary Markov chain and imposing Muth's hypothesis would also rule out the influence of non-fundamental factors on participants' forecasts.

Acknowledging that participants recognize that they face so-called Knightian uncertainty as a result of unforeseeable change would enable economists to represent the role of psychological factors, such as market sentiment, in a model that is consistent with Muth's hypothesis. In a novel paper, Ilut and Schneider (2014) show that this enhances our understanding of business cycles. They introduce Knightian uncertainty into a standard New Keynesian Model and formalize how confidence in future total productivity drives fluctuations in aggregate outcomes. They show that changes in confidence arising from Knightian uncertainty are empirically significant in explaining these fluctuations.

More broadly, recognizing that market participants face Knightian uncertainty would allow consistent representations of the autonomous role their forecasts play in driving outcomes, as argued by Phelps (1970) in his seminal micro-foundations volume.[19] Thus, recognizing that the future is open to change that cannot be specified *ex ante* with probabilistic rules would enable a synthesis of major advances in macroeconomic theory since the 1970s. Building macroeconomic and finance models in accordance with Knight's seemingly uncontroversial yet profound insight promises to enhance substantially our understanding of market outcomes and the role of economic policy.

References

[1] Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry (2021), "Imperfect Macroeconomic Expectations: Evidence and Theory," *NBER Macroeconomic Annual* 35(1), 1-86.

[19] For a formal demonstration, see Frydman, *et al.* (2019).

[2] Bai, Jushan, and Pierre Perron (1998). "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica* 66, 47-78.

[3] Bianchi, Francesco, Sydney C. Ludvigson, and Sai Ma (2020), "Belief Distortions and Macroeconomic Fluctuations," Working Paper No. w27406, NBER.

[4] Barberis, Nicholas C., Andrei Shleifer, and Robert Vishny.(1998), "A Model of Investor Sentiment," *Journal of Financial Economics,* 49, 307-343.

[5] Bordalo, Pedro, Gennaioli, Nicola, Ma, Yueran and Andrei Shleifer (2020), "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110, 2748–2782.

[6] Castle, Jennifer L., Jurgen Doornik, David F. Hendry and Felix Pretis (2015), "Detecting Location Shifts During Model Selection by Step-indicator Saturation," *Econometrics* 3, 240-264.

[7] Castle, Jennifer L., David F. Hendry, and Andrew B. Martinez (2017), "Evaluating Forecasts, Narratives and Policy Using a Test of Invariance," *Econometrics* 5, 1-27.

[8] Coibion, Oliver and Yuriy Gorodnichenko (2012), "What Can Survey Forecasts Tell Us about Information Rigidities?." *Journal of Political Economy* 120, 116-159.

[9] Coibion, Oliver and Yuriy Gorodnichenko (2015), "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review,* 105, 2644-2678.

[10] Doornik, Jurgen (2009), "Autometrics," in *The Methodology and Practice of Econometrics, A Festschrift in Honour of David F. Hendry*, Jennifer Castle and Neil Shephard (eds.), Oxford University Press.

[11] Engel, Charles and James D. Hamilton (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?", *American Economic Review*, 80, 689-713.

[12] Frydman, Roman and Michael D. Goldberg (2007), *Imperfect Knowledge Economics: Exchange Rates and Risk*, Princeton University Press.

[13] Frydman, Roman and Edmund S. Phelps (2013), "Which Way Forward

for Macroeconomics and Policy Analysis?," in Frydman, Roman and Edmund S. Phelps (eds.), *Rethinking Expectations; The Way Forward for Macroeconomics*, Princeton University Press, 1-48.

[14] Frydman, Roman, Johansen, Søren, Rahbek, Anders, and Morten Nyboe Tabor (2019), "The Knightian Uncertainty Hypothesis: Unforeseeable Change and Muth's Consistency Constraint in Modeling Aggregate Outcomes," INET Working Paper No. 92.

[15] Gennaioli, Nicola and Andrei Shleifer (2018), *The Crisis of Beliefs: Investor Psychology and Financial Fragility,* Princeton University Press.

[16] Hamilton, James D. (1988), "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates," *Journal of Economics Dynamics and Control*, 12, 385-423.

[17] Hamilton, James D. (2008), "Regime Switching Models," in Steven Durlauf and Lawrence Blume (eds.) *New Palgrave Dictionary of Economics*, 2nd edition, 53-57, Palgrave McMillan Ltd.

[18] Hendry, David F., Soren Johansen, and Carlos Santos (2008), "Automatic Selection of Indicators in a Fully Saturated Regression," *Computational Statistics* 23, 317-335.

[19] Ilut, L. Cosmin and Martin Schneider (2014), "Ambiguous Business Cycles," *American Economic Review*, 104, 2368-2399.

[20] Kahneman, Daniel and Amos Tversky (1972), "Subjective Probability: A Judgment of Representativeness," in *Cognitive Psychology,* 430-454.

[21] Knight, Frank H. (1921), *Risk, Uncertainty and Profit*, Houghton Mifflin.

[22] Lawler Gregory F. (2006), *Introduction to Stochastic Processes,* Chapman and Hall/CRC.

[23] Lucas, Robert E. Jr. (1973), "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review,* 63, 326-334.

[24] Lucas, Robert E. Jr. (1976), "Econometric Policy Evaluation: A Critique," In Brunner, Karl and Allan Meltzer (eds.), *The Phillips Curve and Labor Markets,* Carnegie-Rochester Conference Series on Public Policy, American Elsevier, 19–46.

[25] Lucas, Robert E. Jr. (1995), "The Monetary Neutrality," The Nobel Prize Lecture, The Nobel Foundation.

[26] Mankiw, Gregory N. and Ricardo Reis (2002), "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 117, 1295–1328.

[27] Mankiw, Gregory N., Reis Ricardo, and Justin Wolfers (2003), *NBER Macroeconomics Annual.*

[28] Muth, John F. (1961), "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29, 315-335.

[29] Phelps, Edmund S. (1970), *Microeconomic Foundations of Employment and Inflation Theory*, Norton.

[30] Reis, Ricardo (2020), "Comment or 'Imperfect Macroeconomic Expectations: Yes, But, We Disagree,'" *NBER Macroeconomics Annual.*

[31] Shleifer, Andrei (2000), "Inefficient Markets: An Introduction to Behavioural Finance," Oxford University Press.

[32] Thaler, Richard H. (2017), "From Cashews to Nudges: The Evolution of Behavioral Economics," The Nobel Prize Lecture, The Nobel Foundation.

[33] Tversky, Amos and Daniel Kahneman (1983), "Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *Psychological Review*, 90, 293-315.

[34] von Gaudecker, Hans-Martin, and Alex Wogrolly (2021), "Heterogeneity in Households' Stock Market Beliefs: Levels, Dynamics and Epistemic Uncertainty," *Journal of Econometrics.*

[35] Woodford, Michael. (2003), "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Aghion, Philippe, Frydman, Roman, Stiglitz, Joseph and Michael Woodford (eds.), 25–58, Princeton University Press.

FOR ONLINE PUBLICATION ONLY

15 Online Appendix A

Proof of Lemma 7

Using the law of total probability, we can express the pdf of x_{t+1}, conditional on x_t, as follows

$$g^{reh(mk)}(x_{t+1}|x_t) = \tag{65}$$
$$h^{(mk,1)}(x_{t+1,}\mu_{t+1} = \mu^{(1)}|x_t) + h^{(mk,2)}(x_{t+1,}\mu_{t+1} = \mu^{(2)}|x_t),$$

where $h^{(i)}\ (\cdot|\cdot)$, $i = 1, 2$ denote the respective pdfs implied by (23). Furthermore, we rewrite the above as

$$\begin{aligned} g^{reh(mk)}(x_{t+1}|x_t) &= \frac{h^{(mk,1)}(x_{t+1,}\mu_{t+1} = \mu^{(1)}, x_t)}{g(x_t)} \\ &\quad + \frac{h^{(mk,2)}(x_{t+1,}\mu_{t+1} = \mu, x_t)}{g(x_t)} \\ &= \frac{h^{(mk,1)}(x_{t+1}|\mu_{t+1} = \mu^{(1)}, x_t)P(\mu_{t+1} = \mu^{(1)}|x_t)g(x_t)}{g(x_t)} \\ &\quad + \frac{h^{(mk,2)}(x_{t+1}|\mu_{t+1} = \mu^{(2)}, x_t)P(\mu_{t+1} = \mu^{(2)}|x_t)g(x_t)}{g(x_t)}, \end{aligned}$$

where $g(x_t)$ is the marginal pdf of x_t. Using the assumed independence of μ_{t+1} and X_t and the stationarity of a Markov chain $\{\mu_t\}$ shows that the "objective" pdf is the mixture of the two normal pdfs, in (**??**):

$$g^{reh(mk)}(x_{t+1}|x_t) = \pi f^{(1,1)}(x_{t+1}|x_t, \mu_{t+1} = \mu^{(1)}) + (1-\pi) f^{(1,2)}(x_{t+1}|x_t, \mu_{t+1} = \mu^{(2)}). \tag{66}$$

where $\pi = P(\mu_{t+1} = \mu^{(1)})$ for all t.

Proof of Lemma 8

Using the law of total probability, we can express the pdf of x_{t+1}, conditional on x_{t-1}, as follows

$$g^{ref(mk)}(x_{t+1}|x_{t-1}) = \sum_{i,j=1}^{2} h^{(i,j)}(x_{t+1,}\mu_{t+1} = \mu^{(i)}, \mu_t = \mu^{(j)}|x_{t-1}), \qquad (67)$$

where $h^{(i,i)}\ (\cdot|\cdot)$, $i, j = 1, 2$ denote the respective pdfs implied by (23).

Analogously to the steps from (65) to (66) in the proof of Lemma 7, the above can be expressed as

$$\begin{aligned} &g^{ref(mk)}(x_{t+1}|x_{t-1}) \\ &= \sum_{i,j=1}^{2} f^{(i,j)}(x_{t+1}|\mu_{t+1} = \mu^{(i)}, \mu_t = \mu^{(j)}, x_{t-1}) P\left(\mu_{t+1} = \mu^{(i)}, \mu_t = \mu^{(j)}|x_{t-1}\right) \end{aligned}$$

Using the assumption that μ_{t+1} and μ_t are independent of x_{t-1}, we rewrite the above as

$$\begin{aligned} g^{ref(mk)}(x_{t+1}|x_{t-1}) = &\sum_{i,j=1}^{2} f^{(i,j)}(x_{t+1}|\mu_{t+1} = \mu^{(i)}, \mu_t = \mu^{(j)}, x_{t-1}) \\ \times P\left(\mu_{t+1} = \mu^{(i)}|\mu_t = \mu^{(j)}\right) P(\mu_t = \mu^{(j)}). \end{aligned}$$

Noting that $P\left(\mu_{t+1} = \mu^{(i)}|\mu_t = \mu^{(j)}\right) = p_{ji}$ is the transition probability, this shows that the reference pdf is the mixture of the four pdfs, in (**??**):

$$g^{ref(mk)}(x_{t+1}|x_{t-1}) = \sum_{i,j=1}^{2} p_{ji}\pi_j f^{(i,j)}(x_{t+1}|x_{t-1}, \mu_{t+1} = \mu^{(i)}, \mu_t = \mu^{(j)}), \qquad (68)$$

where $\pi_j = P(\mu_t = \mu^{(j)})$ for all t.

The mean of $g^{ref(mk)}(x_{t+1}|x_{t-1})$ is the weighted average of the means of the components pdfs, $m_{t+1|t}^{(mk,i,j)}$ in (34):

$$m_{t+1|t}^{ref(mk)} = \rho^2 x_{t-1} + \sum_{i,j=1}^{2} p_{ji}\pi_j \left(\rho\mu^{(j)} + \mu^{(i)}\right)$$
$$= \rho^2 x_{t-1} + \sum_{i=1}^{2} p_{1i}\pi_1 \left(\rho\mu^{(1)} + \mu^{(i)}\right) + \sum_{i=1}^{2} p_{2i}\pi_2 \left(\rho\mu^{(2)} + \mu^{(i)}\right),$$

which, using $p_{ii} = (1 - p_{ij})$, $i, j = 1, 2, i \neq j$ and $\pi_2 = (1 - \pi_1)$, can be written as

$$\begin{aligned} m_{t+1|t}^{ref(mk)} &= E(X_{t+1}|x_{t-1}) = \rho^2 x_{t-1} + \pi_1(1+\rho)\mu^{(1)} - \pi_1 p_{12}(\mu^{(1)} - \mu^{(2)}) \\ &\quad +\pi_2(1+\rho)\mu^{(2)} + \pi_2 p_{21}(\mu^{(1)} - \mu^{(2)}) \\ &= \rho^2 x_{t-1} + (1+\rho)\mu^{(2)} + \{p_{21} + \pi\left[1 + \rho - (p_{12} + p_{21})\right]\}\left(\mu^{(1)} - \mu^{(2)}\right), \end{aligned}$$

where $\pi \equiv \pi_1 = P\left(\mu_t = \mu^{(1)}\right)$ for all t. Noting that $\pi = \frac{p_{21}}{p_{21}+p_{12}}$, implies that $p_{21} - \pi\left(p_{12} + p_{21}\right) = 0$. Thus, the conditional mean of the reference pdf is given by

$$\begin{aligned} m_{t+1|t}^{ref(mk)} &= \rho^2 x_{t-1} + (1+\rho)\left[\mu^{(2)} + \pi(\mu^{(1)} - \mu^{(2)})\right] \\ &= \rho^2 x_{t-1} + (1+\rho)E(\mu_t) \end{aligned} \tag{69}$$

To compute the variance of the reference pdf, in(68), we note that (23) and (69) imply that

$$\begin{aligned} \left(\sigma_{t+1|t}^{ref(mk)}\right)^2 &= E\left\{\left[X_{t+1} - E(X_{t+1}|x_{t-1})\right]^2 || x_{t-1}\right\} \\ &= E\left\{\rho\left[\mu_t - E\left(\mu_t\right)\right] - + \left[\mu_{t+1} - E\left(\mu_t\right)\right]\right\}^2 \\ &= (1+\rho^2)\left[\sigma^2 + V(\mu_t)\right] \\ &\quad +2\rho\left\{E\left(\mu_{t+1}\mu_t\right) - \left[E\left(\mu_t\right)\right]^2\right\} \end{aligned} \tag{70}$$

Finally, Proposition 4 requires that $(1+\theta)\left(\sigma_{t+1|t}^{ref(mk)}\right)^2 > \theta\left(\sigma_{t+1|t}^{reh(mk)}\right)^2$, which from (30) and (70) follows if

$$\left[1+\rho^2(1+\theta)\right]\sigma^2 + V(\mu_t) + 2(1+\theta)\rho E\left(\mu_{t+1}\mu_t\right) - \left[E\left(\mu_t\right)\right]^2 > 0 \qquad (71)$$

We now show that

$$\left[1+\rho^2(1+\theta)\right]V(\mu_t) + 2(1+\theta)\rho E\left(\mu_{t+1}\mu_t\right) - \left[E\left(\mu_t\right)\right]^2 > 0 \qquad (72)$$

holds for any values of the model parameters $\left(\theta, \rho, \mu^{(1)}, \mu^{(2)}, p_{12,}p_{21}\right)$. To this end we express $E\left(\mu_{t+1}\mu_t\right) - \left[E\left(\mu_t\right)\right]^2$ in terms of $V(\mu_t)$ and $\left(\mu^{(1)} - \mu^{(2)}\right)^2$:

$$\begin{aligned}
E(\mu_{t+1}\mu_t) &= E\left[\mu_t E(\mu_{t+1}|\mu_t)\right] \\
&= \mu^{(1)}\left(\mu^{(1)}(1-p_{12}) + \mu^{(2)}p_{12}\right)\pi \\
&\quad + \mu^{(2)}\left[\mu^{(2)}(1-p_{21}) + \mu^{(1)}p_{21}\right](1-\pi) \\
&= \left(\mu^{(1)}\right)^2\pi - \left(\mu^{(1)}\right)^2 p_{12}\pi + \mu^{(1)}\mu^{(2)}p_{12}\pi \\
&\quad + \left(\mu^{(2)}\right)^2(1-\pi) - \left(\mu^{(2)}\right)^2 p_{12}\pi + \mu^{(1)}\mu^{(2)}p_{12}\pi \\
&= E(\mu_t^2) - p_{12}\pi\left(\mu^{(1)} - \mu^{(2)}\right)^2, \qquad (73)
\end{aligned}$$

where we used $p_{21}(1-\pi) = p_{12}\pi$. This shows that

$$E\left(\mu_{t+1}\mu_t\right) - \left[E\left(\mu_t\right)\right]^2 = V\left(\mu_t\right) - p_{12}\pi\left(\mu^{(1)} - \mu^{(2)}\right)^2 \qquad (74)$$

Furthermore, using $\pi = \frac{p_{21}}{p_{12}+p_{21}}$, we express $V\left(\mu_t\right)$ as follows:

$$\begin{aligned}
V\left(\mu_t\right) &= \left(\mu^{(1)}\right)^2\pi + \left(\mu^{(1)}\right)^2(1-\pi) - \left[\mu^{(1)}\pi + \mu^{(1)}(1-\pi)\right]^2 \\
&= \pi(1-\pi)\left(\mu^{(1)} - \mu^{(2)}\right)^2 \qquad (75) \\
&= p_{12}\pi\frac{1}{p_{12}+p_{21}}\left(\mu^{(1)} - \mu^{(2)}\right)^2, \qquad (76)
\end{aligned}$$

which implies that

$$E\left(\mu_{t+1}\mu_t\right)-\left[E\left(\mu_t\right)\right]^2=V\left(\mu_t\right)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2$$
$$=p_{12}\pi\frac{1-p_{12}-p_{21}}{p_{12}+p_{21}}\left(\mu^{(1)}-\mu^{(2)}\right)^2 \quad (77)$$

Substituting (76) into (74) enables us to rewrite the condition (72) as follows

$$p_{12}\pi\frac{1}{p_{12}+p_{21}}\left(\mu^{(1)}-\mu^{(2)}\right)^2$$
$$\times\left[1+\rho^2(1+\theta)+2\rho(1+\theta)\left(1-p_{12}-p_{21}\right)\right]>0 \quad (78)$$

Finally, we note that both roots of the quadratic equation in the square brackets are negative, and thus (78) holds for any $\rho>0$. Via (**??**), this shows that $(1+\theta)\left(\sigma_{t+1|t}^{ref(mk)}\right)^2>\theta\left(\sigma_{t+1|t}^{reh(mk)}\right)^2$ holds for any values of the model parameters, $\left(\theta,\sigma^2\rho,\mu^{(1)},\mu^{(2)},p_{12,}p_{21}\right)$.

Proof of Proposition 9

The argument for the time-invariant REH-like specification is presented in Section 6.2. Here we focus on DE's specification with a Markov component.

The expression for the news, in (33), implies that

$$\mu^{(j)}+e_t=n^{(mk,j)}+E(\mu_t)\quad j=1,2.$$

Substituting this into (31) implies that $m_{t+1|t}^{reh(mk,j)}$, expressed in terms of x_{t-1}, the realized state of μ_t and e_t, can take one of two values:

$$m_{t+1|t}^{reh(mk,j)}=\rho^2x_{t-1}+\rho n^{(mk,j)}+(1+\rho)E(\mu_t)\quad j=1,2,$$

which, using

$$m_{t+1|t-1}^{reh(mk)}=\rho^2x_{t-1}+(1+\rho)E(\mu_t), \quad (79)$$

implies that

$$m_{t+1|t}^{reh(mk,j)}-m_{t+1|t-1}^{ref(mk)}=\rho n^{(mk,j)},\ j=1,2.$$

Proof of Lemma 11

Noting that $m^{reh}_{t+1|t}$ and $m^{reh}_{t+1|t-1}$, in (29) and (79), are time-t and $t-1$ realizations, respectively, of $M^{reh}_{t+1|t}$ and $M^{reh}_{t+1|t-1}$ implies that

$$Z_{2,t} = M^{reh}_{t+1|t} - M^{reh}_{t+1|t-1} = \rho \left[\mu_t - E\left(\mu_t\right)\right] + \rho\varepsilon_t.$$

An argument analogous to the proof of Lemma 7 shows that the pdf of $Z_{2,t}$, denoted with $g^{reht(mk)}(z_{2,t})$, is the mixture of the following two normal pdfs, conditional on the value of μ_t:

$$\begin{aligned}&g^{reh(mk)}(z_{2,t})\\ =&\pi f^{(mk,1)}(z_{2,t}|\mu_t = \mu^{(1)}) + (1-\pi) f^{(mk,2)}(z_{2,t}|\mu_t = \mu^{(2)}),\end{aligned}$$

where, where, the means and variances of the i's component, $i = 1, 2$, are given

$$a^{(mk,i)}_{z_{2,t}} = E(Z_{2,t}|\mu_t = \mu^{(i)}) = \rho\left[\mu^{(i)} - E(\mu_t)\right] \tag{80}$$

$$\begin{aligned}v^{(mk,i)}_{z_{2,t}} &= V(Z_{2,t}|\mu_t = \mu^{(i)}) \\ &= \rho^2\left[\sigma^2 \ + E(\mu_t^2) - 2E(\mu_t)\mu^{(i)} + \left(\mu^{(i)}\right)^2\right]\end{aligned} \tag{81}$$

Consequently, the mean and variance of $Z_{2,t}$, for all t, are given by

$$\begin{aligned}E(Z_{2,t}) &= 0,\\ V(Z_{2,t}) &= \rho^2\left[\sigma^2 + V\left(\mu_t\right)\right].\end{aligned}$$

Proof of Corollary 13

We first consider $Z_{2,t}$ in (58). It follows from (29) and (23) that

$$\begin{aligned}M^{reh(mk)}_{t+1|t} &= E(X_{t+1}|X_t) = \rho X_t + E\left(\mu_t\right)\\ &= \rho^2 X_{t-1} + \rho\mu_t + E\left(\mu_t\right) + \rho\varepsilon_t, &(82)\\ M^{reh(mk)}_{t+1|t-1} &= E(X_{t+1}|X_{t-1}) = \rho E(X_{t+1}|X_t) + E\left(\mu_t\right)\\ &= \rho^2 X_{t-1} + (1+\rho) E\left(\mu_t\right) &(83)\\ &= \rho^3 X_{t-2} + \rho^2\mu_{t-1} + (1+\rho) E\left(\mu_t\right) + \rho^2\varepsilon_{t-1}, &(84)\end{aligned}$$

which, from (82) and (83), implies that

$$Z_{2,t} = M^{reh(mk)}_{t+1|t} - M^{reh(mk)}_{t+1|t-1} = \rho \left[\mu_t - E\left(\mu_t\right)\right] + \rho\varepsilon_t \tag{85}$$

We now consider $Z_{1,t} = M^{de}_{t+1|t} - M^{de}_{t+1|t-1}$:

$$\begin{aligned} Z_{1,t} = &(1+\gamma^{(mk)}) \left(M^{reh(mk)}_{t+1|t} - M^{reh(mk)}_{t+1|t-1} \right) \\ &-\gamma^{(mk)} (M^{reh(mk)}_{t+1|t-1} - M^{reh(mk)}_{t+1|t-2}). \end{aligned} \tag{86}$$

In order to relate $M^{reh(mk)}_{t+1|t-1} - M^{reh(mk)}_{t+1|t-2}$ to μ_t and ε_t, an argument analogous to those in Sections 7.2 and 7.3 shows that $g^{reh(mk)}(x_{t+1}|x_{t-2})$ is a mixture of eight normal pdfs with the mean

$$m^{reh(mk)}_{t+1|t-2} = \rho^3 x_{t-2} + (1+\rho+\rho^2) E\left(\mu_t\right),$$

which, from (79), shows that

$$M^{reh(mk)}_{t+1|t-1} - M^{reh(mk)}_{t+1|t-2} = \rho^2 \left[\mu_{t-1} - E\left(\mu_t\right)\right] + \rho^2\varepsilon_{t-1}, \tag{87}$$

Substituting (85) and (87) into (86) yields

$$\begin{aligned} Z_{1,t} = &(1+\gamma^{(mk)})\rho \left[\mu_t - E\left(\mu_t\right)\right] - \gamma^{(mk)}\rho^2 \left[\mu_{t-1} - E\left(\mu_t\right)\right] \\ &+(1+\gamma^{(mk)})\rho\varepsilon_t - \gamma^{(mk)}\rho^2\varepsilon_{t-1}. \end{aligned} \tag{88}$$

Noting that

$$FE^{reh(mk)}_{t+1|t} = \left[\mu_{t+1} - E\left(\mu_t\right)\right] + \varepsilon_{t+1}, \tag{89}$$

and $V(Z_{1,t}) > 0$, the moments of $Z_{1,t}$, $Z_{2,t}$ and $FE^{reh(mk)}_{t+1|t}$, which underpin the predictions in (56) and (57), are given by

$$
\begin{aligned}
E(Z_{1,t}) &= 0,\ E(Z_{2t}) = 0,\ \ E(FE^{reh(mk)}_{t+1|t}) = 0, && (90)\\
Cov(Z_{1,t}, Z_{2,t}) &= (1+\gamma^{(mk)})\rho^2\sigma^2 + (1+\gamma^{(mk)})\rho^2 V(\mu_t) && (91)\\
&\quad -\gamma^{(mk)}\rho^3\left\{E\left(\mu_{t-1}\mu_t\right) - \left[E\left(\mu_t\right)\right]^2\right\},\\
Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t}) &= (1+\gamma^{(mk)})\rho\left[E\left(\mu_{t+1}\mu_t\right) - \left[E\left(\mu_t\right)\right]^2\right] && (92)\\
&\quad -\gamma^{(mk)}\rho^2\left\{E\left(\mu_{t+1}\mu_{t-1}\right) - \left[E\left(\mu_t\right)\right]^2\right\}
\end{aligned}
$$

where $V(\mu_t) = E(\mu_t^2) - \left[E\left(\mu_t\right)\right]^2$.

According to Proposition 12,

$$
sign(\beta) = sign\left[Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t}) - \gamma Cov(Z_{1t}, Z_{2t})\right]. \qquad (93)
$$

We now show that whether $sign\left[Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t}) - \gamma Cov(Z_{1t}, Z_{2t})\right] < 0$ or > 0 depends on the values of the model parameters $\left(\mu^{(1)}, \mu^{(2)}, p_{12}, p_{21}, \rho, \gamma^{(mk)}, \sigma^2\right)$. Substituting (74) into (91) yields

$$
\begin{aligned}
Cov(Z_{1,t}, Z_{2,t}) &= (1+\gamma^{(mk)})\rho^2\sigma^2 + \rho^2\left[(1+\gamma^{(mk)}) - \gamma^{(mk)}\rho\right]V(\mu_t)\\
&\quad +\gamma^{(mk)}\rho^3 p_{12}\pi\left(\mu^{(1)} - \mu^{(2)}\right)^2\\
&= \rho^2\left[(1+\gamma^{(mk)}) - \gamma^{(mk)}\rho\right]\left[V(\mu_t) - p_{12}\pi\left(\mu^{(1)} - \mu^{(2)}\right)^2\right] && (94)\\
&\quad +\rho^2(1+\gamma^{(mk)})\left[\sigma^2 + p_{12}\pi\left(\mu^{(1)} - \mu^{(2)}\right)^2\right] && (95)
\end{aligned}
$$

In order to derive an analogous expression for $Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t})$, (92), we consider

$$
\begin{aligned}
E\left(\mu_{t+1}\mu_t\right) - E\left(\mu_{t+1}\mu_{t-1}\right) &= E\left\{E\left[\mu_{t+1}(\mu_t - \mu_{t-1})|\left(\mu_t, \mu_{t-1}\right)\right]\right\}\\
&= E\left(\mu_t - \mu_{t-1}\right)E\left[\mu_{t+1}|\left(\mu_t, \mu_{t-1}\right)\right]\\
&= E\left\{\left(\mu_t - \mu_{t-1}\right)E\left[\mu_{t+1}|\mu_t\right]\right\}
\end{aligned}
$$

Because $\mu_t - \mu_{t-1}$ takes two non-zero values, $\mu^{(1)} - \mu^{(2)}$ with the probability $p_{21}(1-\pi)$ and $\mu^{(2)} - \mu^{(1)}$ with the probability $p_{12}\pi$

$$
\begin{aligned}
&E\left\{\left(\mu_t-\mu_{t-1}\right)E\left[\mu_{t+1}|\mu_t\right]\right\}\\
=&\left(\mu^{(1)}-\mu^{(2)}\right)E\left[\mu_{t+1}|\mu_t=\mu^{(1)}\right]p_{21}(1-\pi)\\
&+\left(\mu^{(2)}-\mu^{(1)}\right)E\left[\mu_{t+1}|\mu_t=\mu^{(2)}\right]p_{12}\pi\\
=&\left(\mu^{(1)}-\mu^{(2)}\right)\left[\mu^{(2)}(1-p_{21})+\mu^{(1)}p_{21}\right]p_{12}\pi\\
&+\left(\mu^{(2)}-\mu^{(1)}\right)\left[\mu^{(1)}(1-p_{12})+\mu^{(2)}p_{12}\right]p_{12}\pi\\
=&\{\left(\mu^{(1)}-\mu^{(2)}\right)\mu^{(2)}-\left(\mu^{(1)}-\mu^{(2)}\right)\mu^{(2)}p_{21}+\left(\mu^{(1)}-\mu^{(2)}\right)\mu^{(1)}p_{21}\\
&+\left(\mu^{(2)}-\mu^{(1)}\right)\mu^{(1)}-\left(\mu^{(2)}-\mu^{(1)}\right)\mu^{(1)}p_{12}+\left(\mu^{(2)}-\mu^{(1)}\right)\mu^{(2)}p_{12}\}p_{12}\pi\\
=&\left\{-\left(\mu^{(1)}-\mu^{(2)}\right)^2+\left(\mu^{(1)}-\mu^{(2)}\right)^2p_{21}+\left(\mu^{(1)}-\mu^{(2)}\right)^2p_{12}\right\}p_{12}\pi\\
=&\left(\mu^{(1)}-\mu^{(2)}\right)^2p_{12}\pi\left(p_{12}+p_{21}-1\right),
\end{aligned}
$$

which (via (73)) shows that

$$
E(\mu_{t+1}\mu_{t-1})-\left[E\left(\mu_t\right)\right]^2=V\left(\mu_t\right)-p_{12}\pi\left(p_{12}+p_{21}\right)\left(\mu^{(1)}-\mu^{(2)}\right)^2. \tag{96}
$$

Substituting (73) and (96) into (92) yields

$$
\begin{aligned}
Cov(Z_{1,t},FE^{reh(mk)}_{t+1|t})=&(1+\gamma^{(mk)})\rho\left[V\left(\mu_t\right)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right]\\
&-\gamma^{(mk)}\rho^2\left[V\left(\mu_t\right)-p_{12}\pi\left(p_{12}+p_{21}\right)\left(\mu^{(1)}-\mu^{(2)}\right)^2\right]\\
=&\rho\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]\left[V\left(\mu_t\right)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right] && (97)\\
&-\gamma^{(mk)}\rho^2p_{12}\pi\left(1-p_{12}-p_{21}\right)\left(\mu^{(1)}-\mu^{(2)}\right)^2 && (98)
\end{aligned}
$$

We are now ready to derive

$$
\delta=Cov(Z_{1,t},FE^{reh(mk)}_{t+1|t})-\gamma Cov(Z_{1t},Z_{2t}) \tag{99}
$$

From (95) this difference includes one unambiguously negative term:

$$
\delta_1=-\gamma^{(mk)}\rho^2(1+\gamma^{(mk)})\left[\sigma^2+p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right]<0 \tag{100}
$$

Using (97) and (94) yields

$$\begin{aligned}
&\rho\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]\left[V\left(\mu_t\right)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right]\\
&-\gamma^{(mk)}\rho^2\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]\left[V(\mu_t)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right]\\
=&\,\rho(1-\gamma^{(mk)}\rho)\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]\left[V\left(\mu_t\right)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right],
\end{aligned}$$

which, combined with (98), yields the second term of (99):

$$\begin{aligned}
\delta_2 =&\,\rho(1-\gamma^{(mk)}\rho)\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]\left[V\left(\mu_t\right)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right] \\
&-\gamma^{(mk)}\rho^2p_{12}\pi\left(1-p_{12}-p_{21}\right)\left(\mu^{(1)}-\mu^{(2)}\right)^2
\end{aligned} \quad (101)$$

We now show that there are values of the parameters $\gamma^{(mk)}, \rho, p_{12}$, and p_{21} for which $\delta_2 < 0$, thereby implying (via (100) that

$$\delta = \delta_1 + \delta_2 = Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t}) - \gamma Cov(Z_{1t}, Z_{2t}) < 0.$$

Substituting (**76**) into (101) expresses δ_2 as

$$\begin{aligned}
\delta_2 =&\,\rho(1-\gamma^{(mk)}\rho)\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]\left[V\left(\mu_t\right)-p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\right]\\
&-\gamma^{(mk)}\rho^2p_{12}\pi\left(1-p_{12}-p_{21}\right)\left(\mu^{(1)}-\mu^{(2)}\right)^2\\
=&\,\rho(1-\gamma^{(mk)}\rho)\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]p_{12}\pi\frac{1-p_{12}-p_{21}}{p_{12}+p_{21}}\left(\mu^{(1)}-\mu^{(2)}\right)^2\\
&-\gamma^{(mk)}\rho^2p_{12}\pi\left(1-p_{12}-p_{21}\right)\left(\mu^{(1)}-\mu^{(2)}\right)^2\\
=&\,\rho p_{12}\pi\left(\mu^{(1)}-\mu^{(2)}\right)^2\left(1-p_{12}-p_{21}\right)\left(\mu^{(1)}-\mu^{(2)}\right)^2\\
&\times\left\{(1-\gamma^{(mk)}\rho)\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]\frac{1}{p_{12}+p_{21}}-\gamma^{(mk)}\rho\right\}\\
=&\,\frac{\rho p_{12}\pi}{p_{12}+p_{21}}\left(\mu^{(1)}-\mu^{(2)}\right)^2\left(1-p_{12}-p_{21}\right) &(102)\\
&\times\left\{(1-\gamma^{(mk)}\rho)\left[(1+\gamma^{(mk)})-\gamma^{(mk)}\rho\right]-\gamma^{(mk)}\rho\left(p_{12}+p_{21}\right)\right\} &(103)
\end{aligned}$$

In order to uncover the conditions under which $\delta_2 < 0$, we note that the term in (102) is negative if and only if

$$p_{12} + p_{21} > 1. \tag{104}$$

Furthermore, the term in (103) is positive if

$$1 < p_{12} + p_{21} < (\frac{1}{\gamma^{(mk)}\rho} - 1)\left[(1 + \gamma^{(mk)}(1 - \rho)\right]. \tag{105}$$

Thus, if $p_{12} + p_{21}$ satisfies (105), $\delta_2 < 0$..

However, although the right bound in (105) is greater than 1 for any values of $0 < \gamma^{(mk)}, \rho < 1$, if $\gamma^{(mk)}\rho. > 1/3$, there are values of $1 < p_{12} + p_{21} < 2$ such that

$$p_{12} + p_{21} > (\frac{1}{\gamma^{(mk)}\rho} - 1)\left[(1 + \gamma^{(mk)}) - \gamma^{(mk)}\rho\right] > 1, \tag{106}$$

which implies that the term in (103) is negative, and thus $\delta_2 > 0$. We also note that if $p_{12}+p_{21} < 1$, both (102) and (103) are positive, thus, $\delta_2 > 0$. This shows that the condition (105) is necessary and sufficient for

$$\delta = \delta_1 + \delta_2 = Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t}) - \gamma Cov(Z_{1t}, Z_{2t}) < 0$$

Finally if the condition (105) is not satisfied and

$$\begin{aligned}\delta_2 =& \frac{\rho p_{12}\pi}{p_{12}+p_{21}}\left(\mu^{(1)} - \mu^{(2)}\right)^2(1 - p_{12} - p_{21}) \\ &\times\left\{(1 - \gamma^{(mk)}\rho)\left[(1 + \gamma^{(mk)}) - \gamma^{(mk)}\rho\right] - \gamma^{(mk)}\rho\,(p_{12} + p_{21})\right\} \\ &> -\delta_1 = \gamma^{(mk)}\rho^2(1 + \gamma^{(mk)})\left[\sigma^2 + p_{12}\pi\left(\mu^{(1)} - \mu^{(2)}\right)^2\right]\end{aligned}$$

then

$$\delta = \delta_1 + \delta_2 = Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t}) - \gamma Cov(Z_{1t}, Z_{2t}) > 0.$$

Proof of Corollary 15

We formalize regime persistence by constraining μ_t to take the same value from $t-2$ to $t+1$.Suppose (without a loss of generality) that

$$\mu_{t-2} = \mu_{t-1} = \mu_t = \mu_{t=1} = \mu^{(1)}$$

Imposing this constraint in (88), (85), and (89) specifies $Z_{1,t}$, $Z_{2,t}$, and $FE^{reh(mk)}_{t+1|t}$ as follows

$$\begin{aligned} Z_{1,t} &= (1+\gamma^{(mk)})\rho\left[\mu^{(1)} - E\left(\mu_t\right)\right] - \gamma^{(mk)}\rho^2\left[\mu^{(1)} - E\left(\mu_t\right)\right] \\ &\quad +(1+\gamma^{(mk)})\rho\varepsilon_t - \gamma^{(mk)}\rho^2\varepsilon_{t-1}, \\ Z_{2,t} &= \rho\left[\mu^{(1)} - E\left(\mu_t\right)\right] + \rho\varepsilon_t, \\ FE^{reh(mk)}_{t+1|t} &= \left[\mu^{(1)} - E\left(\mu_t\right)\right] + \varepsilon_{t+1}. \end{aligned}$$

This immediately implies that

$$E(Z_{1,t}) = \rho(1+\gamma^{(mk)} - \gamma^{(mk)}\rho^2)\left[\mu^{(1)} - E\left(\mu_t\right)\right], \tag{107}$$

$$E(Z_{2,t}) = \rho\left[\mu^{(1)} - E\left(\mu_t\right)\right], \tag{108}$$

$$E(FE^{reh(mk)}_{t+1|t}) = \left[\mu^{(1)} - E\left(\mu_t\right)\right], \tag{109}$$

$$Cov(Z_{1,t}, Z_{2,t}) = (1+\gamma^{(mk)})\rho^2\sigma^2, \tag{110}$$

$$Cov(Z_{1,t}, FE^{reh(mk)}_{t+1|t}) = 0, \tag{111}$$

which (via Proposition 12) indicates the following predictions for the coefficients of (48) in each of the regimes:

$$\alpha^{de(mp)} = \left(1-\gamma^{(mk)}\rho\right)\left[\mu^{(i)} - E\left(\mu_t\right)\right] \tag{112}$$

$$\times\left[\frac{(1+\gamma^{(mk)})\rho^2\sigma^2}{V(Z_{1,t})} + (1+\gamma^{(mk)} - \gamma^{(mk)}\rho^2)\right], i = 1,2$$

$$\beta^{de(mp)} = -\gamma^{(mk)}(1+\gamma^{(mk)})\rho^2\sigma^2 < 0 \tag{113}$$

Thus, if a Markov chain is in the regime for which $\left[\mu^{(i)} - E\left(\mu_t\right)\right] > 0$, $(\left[\mu^{(i)} - E\left(\mu_t\right)\right]) < 0$, $\alpha^{de(mp)} > 0$ ($\alpha^{de(mp)} < 0$).

Proof of Corollary 16

As in the GS time-invariant model, in Section 5,

$$Z_{2,t} = \rho\varepsilon_t,$$
$$FE^{reh}_{t+1|t} = \varepsilon_t,$$

which, using (39), (42), (40), and (43, enables us to express $Z_{1,t}$, in (51) as follows:

$$Z_{1,t} = \left[1 + \gamma^{(b)}(1-\rho\right] \rho\varepsilon_t - \gamma^{(b)} \left[\rho - E(\rho_t^{(b)})\right] \varepsilon_t$$
$$-\gamma^{(b)} E\left(\rho_t^{(b)}\right) \left\{ \left[\rho - E(\rho_t^{(b)})\right] \rho x_{t-1} + \left[\mu - E(\mu_t^{(b)})\right] \right\}.$$

These expressions imply that

$$E\left(Z_{2,t}\right) = 0,$$
$$E\left(Z_{1,t}\right) = \gamma^{(b)} E\left(\rho_t^{(b)}\right) \left\{ \frac{\rho\mu}{1-\rho} \left[\rho - E(\rho_t^{(b)})\right] + \left[\mu - E(\mu_t^{(b)})\right] \right\},$$
$$Cov(Z_{1,t}, Z_{2,t}) = \left[\rho\left(1 + \gamma^{(b)}\right) - \gamma^{(b)} E(\rho_t^{(b)})\right] \rho\sigma^2,$$
$$E\left(FE^{reh}_{t+1|t}\right) = 0$$
$$Cov(Z_{1,t}, FE^{reh}_{t+1|t} = 0,$$

where we used $E(x_{t-1}) = \frac{\mu}{1-\rho}$. Then, Proposition 12 implies the predictions for the coefficients of the CG regression stated in the corollary.

Proof of Corollary 19

From (63) and (64), $Y_{t+1|t}$

$$E\left(Y_{t+1|t}\right) = 0,$$
$$E\left(Z_{1,|t}\right) = 0,$$

which, from (61), implies that

$$\alpha^{fire(mk)} = 0 \ at\ all\ t$$

(63) and (64) also imply that

$$Cov(Y_{t+1|t}, Z_{1,|t}) = \rho \left\{ \left(E\mu_{t+1}\mu_t \right) - \left[E\left(\mu_t\right) \right]^2 \right\},$$

which, using (77), enables us to express

$$Cov(Y_{t+1|t}, Z_{3,|t}) = \frac{\rho p_{12} \pi}{p_{12} + p_{21}} \left(\mu^{(1)} - \mu^{(2)} \right)^2 \left(1 - p_{12} - p_{21} \right).$$

Then, (62) implies predictions for $\beta^{fire(mk)}$ stated in the corollary.

16 Online Appendix B

16.1 Full-Sample Individual CG Regressions

Whereas the forecast error is always observed anytime a forecast is reported, the data point for the forecast revision requires two consecutive forecasts submitted by an individual. Table B1 is sorted by the number, N, of observations available for individuals in the survey with more than 50 observations for revisions. $\hat{\alpha}$ and $\hat{\beta}$ in the table are, respectively, the full-sample estimate of the constant and the slope in the CG regression, (47). These results are summarized in Table 2 in the body of the paper.

Table B1: Full-Sample Individual-Level Estimates

N	$\hat{\alpha}$	$\hat{\beta}$	N	$\hat{\alpha}$	$\hat{\beta}$	N	$\hat{\alpha}$	$\hat{\beta}$
94	0.006 [1.24]	−0.041 [−0.08]	68	0.001 [0.24]	−0.002 [−0.84]	57	0.011 [2.08]	−0.319 [−1.39]
90	−0.004 [−1.21]	−0.261 [−1.67]	65	−0.007 [−1.37]	−0.476 [−4.87]	56	0.001 [0.41]	0.149 [0.54]
82	−0.004 [−1.30]	−.215 [−1.31]	65	0.001 [0.40]	0.477 [1.11]	54	0.003 [1.09]	−0.367 [−2.79]
80	−0.003 [−0.58]	−0.313 [−1.44]	63	0.001 [.21]	0.196 [0.60]	53	0.001 [0.18]	−0.281 [−1.94]
78	−0.004 [−0.73]	−0.053 [−0.17]	62	−0.001 [−0.22]	−0.199 [−0.51]	52	−0.008 [−3.44]	−0.075 [−0.45]
78	−0.002 [−0.52]	−0.007 [−0.03]	61	0.002 [0.06]	−0.212 [−3.08]	52	0.001 [0.16]	1.266 [2.45]
78	0.001 [0.31]	−0.117 [−0.62]	61	−0.005 [−1.80]	−0.007 [−0.05]	52	0.003 [1.66]	0.135 [0.37]
70	−0.002 [−0.84]	0.053 [0.16]	59	0.002 [0.12]	0.266 [2.58]	51	0.000 [0.10]	0.067 [0.49]

Caption: t-values are displayed in brackets under the parameter estimates.

16.2 Overview of MIS

MIS was first proposed by Ericsson (2012) as an extension of robust estimation methods which, respectively, detect outliers and mean shifts: impulse indicator saturation (IIS), developed in Hendry, *et al.* (2008), and step indicator saturation (SIS), developed in Castle, *et al.* (2015). The general idea of MIS is to multiply each regressor by a step indicator for each observation that is equal to unity up until time j and zero thereafter. This allows for breaks in the regressors' coefficients separately and at any point in time. Combined with IIS and SIS, in the context of the CG regression yields:

$$x_{t+1} - F_t(x_{t+h}) = \alpha + \beta[F_t(x_{t+h}) - F_{t-1}(x_{t+h})] + \Sigma_{i=1}^{T-1}\beta^i\mu 1_{t<i}[F_t(x_{t+h}) - F_{t-1}(x_{t+h})]$$
$$+\Sigma_{i=1}^{T}\delta 1 + \Sigma_{i=2}^{T-1}\beta^i\mu 1_{t<i} + error.$$

The impulse indicators $\Sigma_{i=1}^{T}\delta 1$ (one for each observation) allow for an outlier at any point in time. The step indicators $\Sigma_{i=2}^{T-1}\beta^i\mu 1_{t<i}$, allow for a differential shift in the constant, relative to the end-of-sample constant.[20] The multiplicative indicators $\Sigma_{i=1}^{T-1}\beta^i\mu 1_{t<i}[F_t(x_{t+h}) - F_{t-1}(x_{t+h})]$ allow for a differential slope coefficient at any point in time, relative to the end-of-sample estimate.

The significant multiplicative step and impulse indicators are selected by the Autometrics tree search algorithm (Doornik, 2009). After the multiplicative indicators have been determined by the algorithm, a model-selection bias correction is applied (Hendry and Krolzig, 2005). This correction eliminates the well-known bias originally documented by Lovell (1983).[21]

16.3 MIS Estimates of CG Regression

Figures B1-B4 below display MIS estimates of statistically significant breaks in the constant and revision coefficient of individual regressions based on survey data from the 24 forecasters. These results have been grouped in rows 2-5 and columns 3-5 of Table 3.

All of the individual regressions experience breaks in either the constant or revision coefficient, which are significant at 1%. We can see, however, that the constant and the slope break at different times, including frequent cases where the constant changes but the slope does not. As discussed in Section

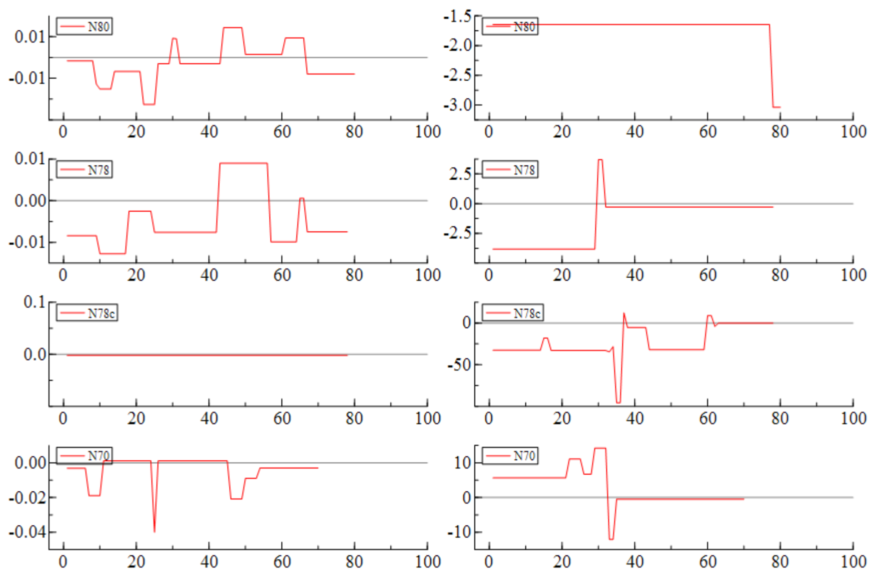
[20] The impulse indicator and step indicator, as specified, are identical in the first observation, so the latter is summed only from the second observation.

[21] The correction reduces the absolute value of the coefficients. The size of the adjustment depends inversely on the t-values and significance level used for selection. This reduces the selection bias, because this adjustment is larger when there is a greater probability of Type II error.

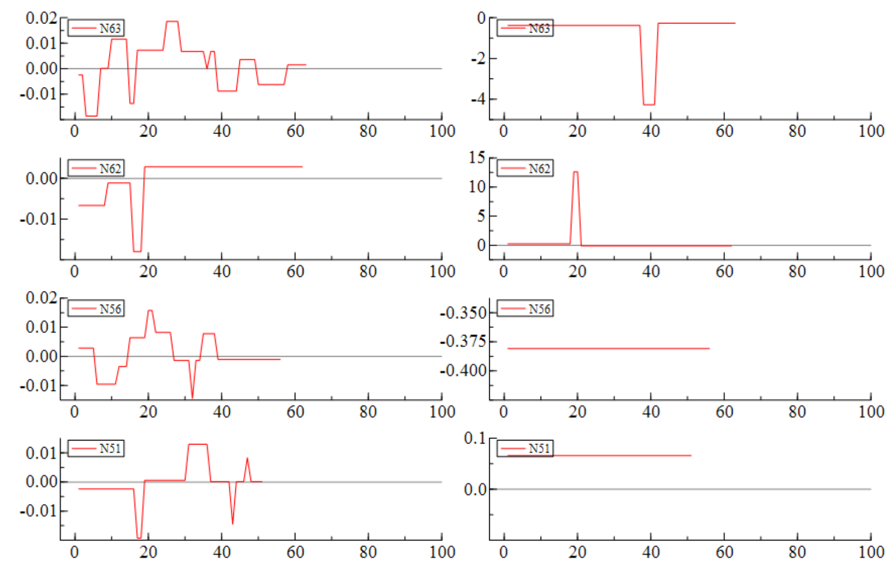
13.3, this supports our argument for using MIS, rather than the Bai-Perron test, as a procedure for testing the stability of the individual coefficients in the CG regression.

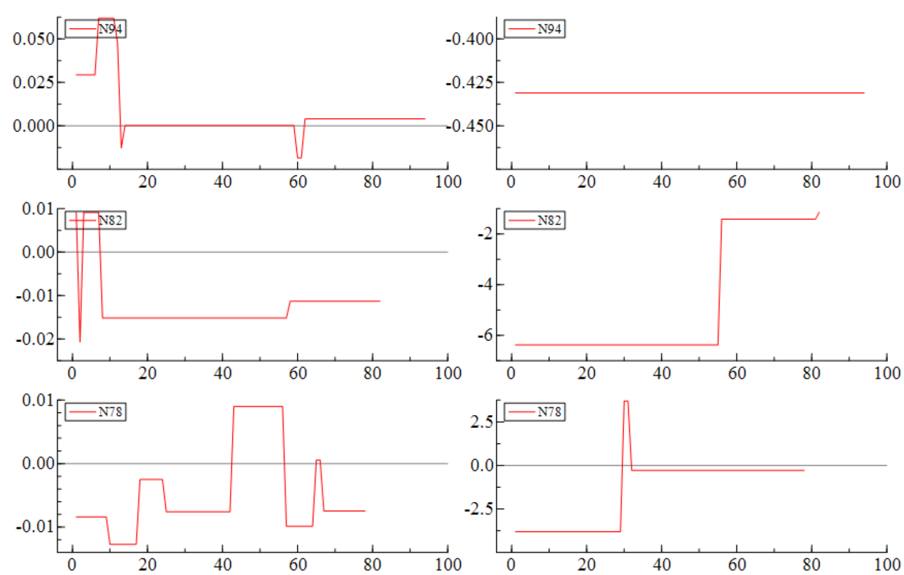
Figure B1: MIS Estimates of Significant Breaks in CG Regressions Not Rejecting FIRE in Full Sample

α β



α β





α β

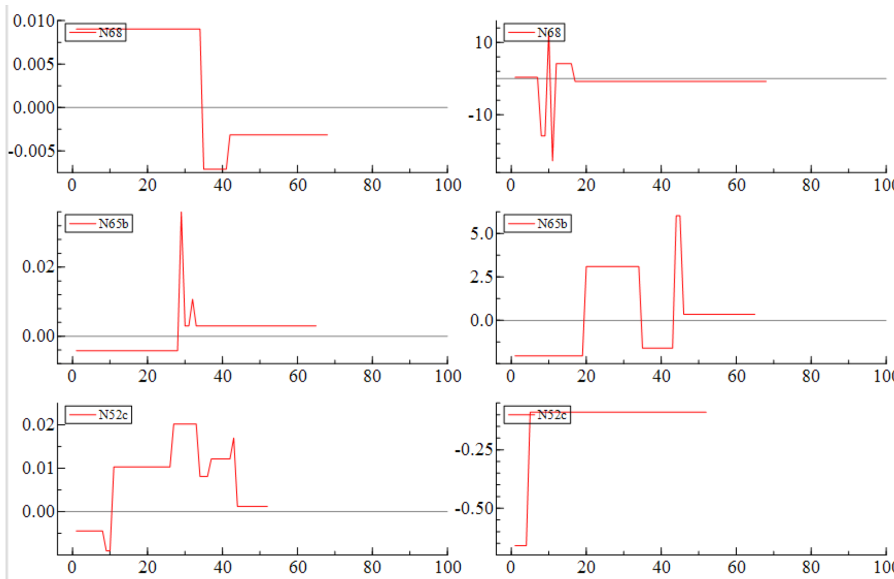


Figure B1 presents the MIS estimates of structural breaks in the 14 individual regressions that have not rejected $\alpha = \beta = 0$ based on the full sample. As summarized in row 2 of Table 2, all 14 of these CG regressions experience breaks in either the constant (13/14) or the slope (8/14).

Figure B2: MIS Estimates of Significant Breaks in CG Regressions Apparently Consistent with DE Based on the Full Sample

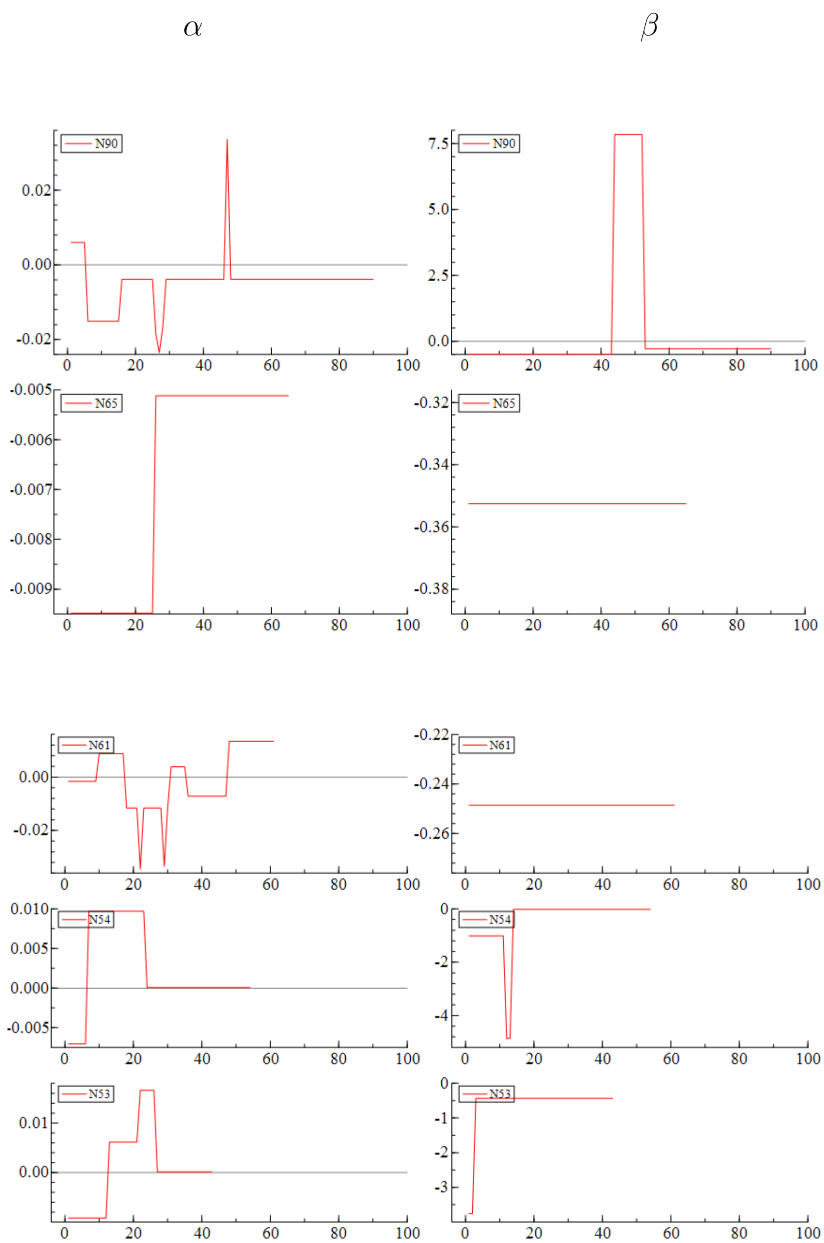
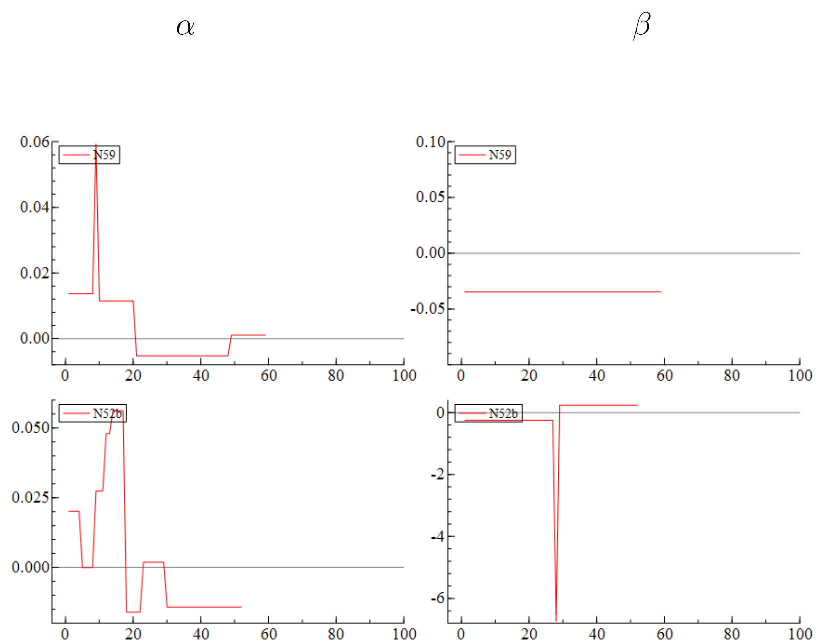


Figure B2 displays the MIS estimates of significant, at 1%, structural breaks in five individual regressions that are consistent with $\alpha = 0$ and $\beta < 0$, based on the full sample. As summarized in row 3 of Table 3, all regressions experience breaks in either α or β, which rejects time-invariant DE. Among them are two cases where MIS estimates indicate a time-invariant $\beta < 0$, which, assuming regime persistence, could render it consistent with Model C in Table 1. However, these cases also experience breaks in α that are inconsistent with Model C's predictions: α switches between two different values with the same sign.

Figure B3: MIS Estimates of Significant Breaks in Regressions Apparently Consistent with Noisy-Information REH Based on the Full Sample



As shown in Figure B3, neither individual regression is consistent with noisy information after MIS. One indicates a sign change in β, while the other that $\beta = 0$. Both detect significant breaks in α, some of which maintain the same sign. Thus, these two cases are not consistent with any of our models in Table 1.

Figure B4: MIS Estimates of Significant Breaks in Individual Regressions Indicating $\alpha \neq 0$ and $\beta = 0$ Based on the Full Sample

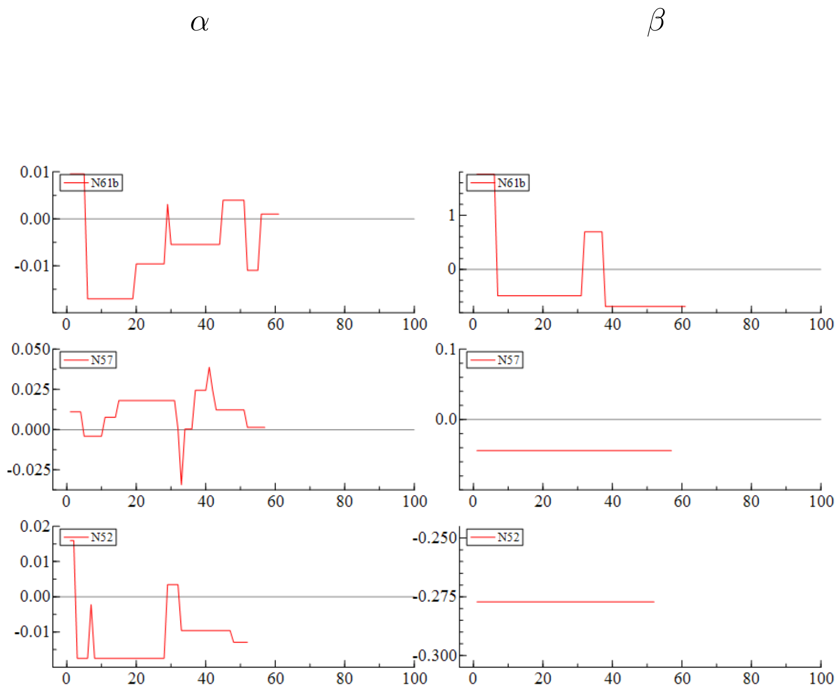


Figure B4 presents MIS estimates of the 3 remaining CG regressions. These full-sample regressions were not consistent with any of the models in Table 1. However, we need to test for breaks, because the regression could indicate breaks in α, which with a time-invariant $\beta < 0$ could be consistent with Model C. As Figure B4 shows, one of the CG regressions is indeed consistent with time-invariant $\beta < 0$, but indicates breaks in α, while maintaining the same sign. Again, none is consistent with any of our eight specifications in Table 1.

16.4 Bai-Perron Test for Structural Breaks in the CG Regression

The Bai-Perron Procedure determines the significant break dates subject to a user-input significance level and a trimming parameter which dictates the minimum duration and maximum number of breaks. A new regression is then estimated within those breaks.

MIS also detects breaks in each parameter at a user-input significance level. However, it does not constrain the constant and slope coefficient(s) to break simultaneously.[22]

We largely follow Bai-Perron (2003) with a 5% significance level, 15% trimming parameter (corresponding to a maximum of five breaks), and using the HAC standard errors with the Andrews' automatic kernel bandwidth estimator. We also use one-lag of pre-Whitening.

We recall that we used 1% significance for MIS. The 5% significance level used here for Bai-Perron should detect more breaks than if 1% were used for Bai-Perron. Nonetheless, Bai-Perron finds fewer individual regressions experiencing breaks than MIS (13 for Bai-Perron vs. all 24 for MIS).

Table B2 provides the estimates produced by the Bai-Perron procedure. For ease of interpretation, the model classifications are color coded. Red indicates a full sample or sub-sample estimates consistent with time-invariant FIRE ($\alpha = \beta = 0$), green indicates DE ($\alpha = 0$ and $\beta < 0$), and blue Noisy Information RE ($\alpha = 0$ and $\beta > 0$). The light green and light blue indicate, respectively, a significant β and α. The white indicates $\alpha \neq 0$ and $\beta = 0$.

[22] MIS also has the advantage of controlling for outliers using the impulse indicator saturation of Hendry, *et al.* (2008).

Table B2: Bai-Perron Estimates of Structural Breaks in the Individual CG Regressions

N	Regime 1 α	Regime 1 β	Reg 2 α	Reg 2 β	Reg 3 α	Reg 3 β	Reg 4 α	Reg 4 β	Reg 5 α	Reg 5 β	Reg 6 α	Reg 6 β
93	.025 (1.93)	.321 (.246)	.005 (.81)	-.004 (-.02)	-.002 (-.52)	-.584 (-8.93)	.005 (.88)	-.448 (-1.78)				
90	-.004 (-1.21)	-.261 (-1.67)										
82	.004 (.45)	-.678 (-3.51)	-.007 (-4.21)	.308 (.80)	-.004 (-2.29)	-0.877 (-6.54)	-.011 (-2.48)	-0.046 (-.19)	-.0003 (-0.04)	.607 (.43)	-.005 (-.22)	-1.641 (-.22)
80	-0.003 (-0.58)	-0.313 (-1.44)										
78	-.004 (-.73)	-.053 (-.17)										
78	-.005 (-1.36)	-.258 (-1.57)	.012 (4.14)	-.627 (-2.95)	-.006 (-2.56)	-.551 (-2.77)						
78	-.004 (-1.33)	-.288 (-.69)	0.014 (1.36)	0.300 (.53)	.0001 (.04)	-.498 (-1.34)						
70	-.002 (-.84)	0.053 (.16)										
68	.002 (.37)	1.075 (1.60)	.008 (5.57)	-1.33 (-6.05)	-.004 (-2.09)	-.279 (-.85)						
65	-0.007 (-1.37)	-.476 (-4.87)										
65	.002 (1.62)	-0.983(-2.45)	-.005 (-5.64)	-1.634 (-2.89)	-.008 (-.008)	2.004 (.003)	0.012 (5.23)	1.423 (0.36)	-0.002 (-.55)	-.218 (-2.66)	0.004 (.93)	1.057 (7.65)
63	.0009 (.21)	.196 (.60)										
62	-.001 (-.22)	-.199 (-.51)										
61	-.004 (-1.36)	-.273 (-2.92)	.013 (8.51)	.445 (1.89)								
61	.007 (2.17)	.705 (1.67)	-.015 (-10.02)	-.442 (-8.69)	-.004 (-2.27)	-.425 (-2.44)						
59	.012 (3.38)	-.073 (-.38)	-.011 (-8.12)	-.290 (-2.62)	.0007 (.14)	-.460 (-1.68)						
57	.011 (2.08)	-.319 (-1.39)										
56	.001 (.41)	.149 (.54)										
54	-.003 (-.92)	-.379 (-4.14)	0.013 (3.17)	.128 (.66)	-.003 (-4.47)	-.555 (-3.84)	.010 (8.56)	-.451 (-.46)	-.001 (-1.61)	-.743 (-1.99)		
53	-.0003 (-.33)	-.546 (-3.04)	-.016 (-3.28)	-1.098 (-1.42)	.009 (3.81)	-.417 (-2.91)	-.0009 (-.39)	-.319 (-2.66)				
52	-.008 (-3.44)	-.075(-.45)										
52	.007 (1.24)	-.139 (-.32)	.045 (3.71)	-.238 (-.59)	-.005 (-.63)	-.113 (-.25)	-.010 (-3.17)	.798 (4.15)				
52	-.009 (-3.13)	-.385 (-2.09)	.005 (2.83)	.505 (1.68)								
51	.0002 (.10)	.067 (.49)										

The Bai-Perron test did not detect breaks for eight individual regressions that did reject FIRE in the full sample (the first eight individual regressions displayed in Figure B1). Therefore, it was unable to reject time-invariant FIRE for one-third of the regressions. Similarly, the Bai-Perron procedure did not detect breaks for two of the five individual regressions consistent with $\alpha = 0$ and $\beta < 0$, as well as for the two consistent with $\alpha \neq 0$ and $\beta = 0$, based on the full-sample estimates. By contrast, MIS finds breaks in either α or β for all individual regressions.

References

[1] Bai, Jushan, and Pierre Perron (2003), "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics* 18, 1-22.

[2] Castle, Jennifer L., Jurgen Doornik, David F. Hendry and Felix Pretis (2015), "Detecting Location Shifts During Model Selection by Step-indicator Saturation," *Econometrics* 3, 240-264.

[3] Doornik, Jurgen (2009), "Autometrics," in *The Methodology and Practice of Econometrics, A Festschrift in Honour of David F. Hendry*, Jennifer Castle and Neil Shephard (eds.), Oxford University Press.

[4] Ericsson, Neil R. (2012). "Detecting Crises, Jumps, and Changes in Regime," Working Paper, Federal Reserve Board of Governors, Washington, D.C., USA.

[5] Hendry, David F., Soren Johansen, and Carlos Santos (2008), "Automatic Selection of Indicators in a Fully Saturated Regression," *Computational Statistics* 23, 317-335.

[6] Hendry, David F. and Hans-Martin Krolzig (2005), "The Properties of Automatic Gets Modelling," *The Economic Journal* 115, C32-C61.

[7] Lovell, Michael C. (1983). "Data Mining," *Review of Economics and Statistics* 65, 1-12.