Journal of Economic Dynamics & Control 170 (2025) 105005

Contents lists available at ScienceDirect

Journal of Economic Dynamics and Control

journal homepage: www.elsevier.com/locate/jedc

Judgment can spur long memory

Emilio Zanetti Chini [1]

University of Bergamo, Department of Economics, Via dei Caniana, 2 - 24127, Bergamo, Italy

ARTICLE INFO

This paper is dedicated to the memory of Francesca Rossi (1979—2024)

JEL classification:
E17
E3
C12
C22
C44
C53

Keywords:
Belief formation
Dynamic systems
Power spectrum
Filtering

ABSTRACT

We arrive at this conclusion by using a new family of models—the Long Memory Dynamic Judgmental Protocol (LMDJP)—where robust filtering and fractionally integrated auto-regressions are combined in an environment characterized by several players—namely, Forecast Producer, Forecast User, and Reality. Namely, we show that if judgment is parametrized as a deformation Likelihood function according to Lq-Likelihood methods, such a deformation affects (sometimes dramatically) the Power Spectrum, consequently inducing over-rejection in formal tests for no LM-effects based on the last. Our simulated and empirical evidence reveals that knowledge of the fractional integration parameter matters for the p-values of tests for spurious LM and, secondly, that the role of LM in belief formation is ambiguous.

1. Introduction

The systematic tendency of economic agents to over- or under-predict the future value of an economic variable (or forecasting bias) is one of the most well-known and puzzling phenomena in modern Macroeconomics and Finance.

For the past couple of decades, economic literature has identified the informational channel as one of the main drivers of this issue. Specifically, economic agents may incur costs in data processing, either during the forecasting phase or the evaluation phase (Mankiw and Reis, 2002; Sims, 2003). Alternatively, they might have incentives to alter forecast communication due to private information, agency costs, or other forms of strategic behavior (Ehrbeck and Waldmann, 1996; Gallo et al., 2002), *inter alia*.

On the other hand, a purely econometric approach would suggest that any deficiencies in forecasting activity are due to poor model specification. Historically, one of the most common sources of dynamic mis-specification in time series analysis has been the presence of long-run time series dynamics, identified by the decay of sample auto-covariance or, equivalently, the stability of its power spectrum. This feature is known as *long-range dependence* or *long-memory* (LM). LM has been well-documented in macroeconomics, at least since the contribution by Diebold and Rudebusch (1989). Decades of econometric and statistical literature provide

E-mail address: emilio.zanettichini@unibg.it.

[1] The Author would like to thank Juan Francisco Rubio-Ramirez and Luca Benati as Co-Editor and Associated Editor and two anonymous Reviewers, whose work allowed to improve the quality of this paper. He also thanks Francesco Ravazzolo and Alessandra Luati for their support. On multiple instances, he praises Barbara Annicchiarico, Alessandra Bonfiglioli, Simone Giannerini, Liudas Giraitis, Greta Goracci, Riccardo J. Lucchetti and Salvatore Piccolo for their remarks in occasion of several scientific meetings whereas this paper was presented. Moreover, he thanks Katsumi Shimotsu for sharing his MatLab routines, which showed precious for the enhancement of this paper; the author takes full responsibility for the outcome given here. Lastly, the Author wishes to express gratitude to the doctors, nurses, and staff at Pavia's Policlinico "S. Matteo," without whose uncompensated care this article would not have been written.

https://doi.org/10.1016/j.jedc.2024.105005
Received 6 March 2024; Received in revised form 18 November 2024; Accepted 21 November 2024

Available online 28 November 2024

various explanations for this phenomenon and models that enable researchers to forecast economic variables where this feature is relevant.

Here, we introduce a new, micro-founded explanation for the rise of LM: *judgment*. This is defined (Svensson, 2005) as anything outside the scope of an economic or statistical model—or, in econometric terms, as the non-sample information introduced in the estimation phase by a forecaster or decision maker. According to this definition, whenever reality does not confirm the prediction, judgmental bias (JB) occurs. JB differs considerably from traditional estimation bias: the former results from strategic interactions among several economic agents, whereas estimation bias does not necessarily have an economic explanation. Interestingly, JB can be significant even in the absence of estimation bias. For this reason, judgment has often been considered outside the purview of econometric practice and relegated to behavioral and cognitive sciences. It is not surprising that literature on JB in mainstream economic theory-based modeling, such as Dynamic Stochastic General Equilibrium (DSGE)—where judgment can be seen as a by-product of agents' heterogeneity—is scarce (Monti, 2010; Gelain and Manganelli, 2020), while literature on the effects of judgment on structural Vector AutoRegression (SVAR) modeling is almost nonexistent.[2]

Despite this literature gap, it is already possible to directly elicit the (unobserved) component of judgment in data using a robust signal extraction technique introduced by Zanetti Chini (2023) called the Judgmental (or Deformed Kalman) Filter (JF). This method is the solution to a dynamic game between two players: a forecast producer (FP, who is a professional forecaster) and a forecast user (FU, such as a central banker). Each player has a subjective utility function dependent on future realizations and available information, which causes the other to believe that the genuine forecasts and decisions have not been accurately communicated. The evidence from the Survey of Professional Forecasters (SPF) data and judgments estimated through this method is not only significant and pervasive but also time-varying and correlated with the business cycle. However, the long-run dynamics and effects of judgment remain unexplored.

A simple investigation of the survey data on U.S. real GDP is sufficient to motivate our research. In Fig. 1, real observations are compared with OLS (implemented in a dynamic system via the Kalman filter) and judgmental filter-based one-step-ahead estimates, either statically via simple scatter plots or dynamically via the autocorrelation function. In all cases, there is a clear divergence (in terms of static regression) between judgmental-filter-based forecasts and optimal forecasts: while in real observations the former are less dispersed than the latter, the opposite seems to hold when considering one-quarter-ahead forecasts. On the other hand, the autocorrelations of real data are quite similar to those of a classical AR(2) model (assumed to be the true GDP), with a cutoff after the second lag, while a completely different dynamic appears in professional forecasts. These forecasts decay considerably more slowly. Moreover, the path of this slow decay differs according to the initial level of judgment assumed in the estimation phase: if the data are assumed to have low judgment in the initial recursion, the autocorrelations of all the series under investigation (real data, filter-based estimates, innovations, and their spread) are U-shaped. In contrast, for high initial judgment, a differentiation occurs: real observations and filter-based forecasts exhibit permanently non-zero autocorrelations, while the other two display the previously seen U-shaped behavior.

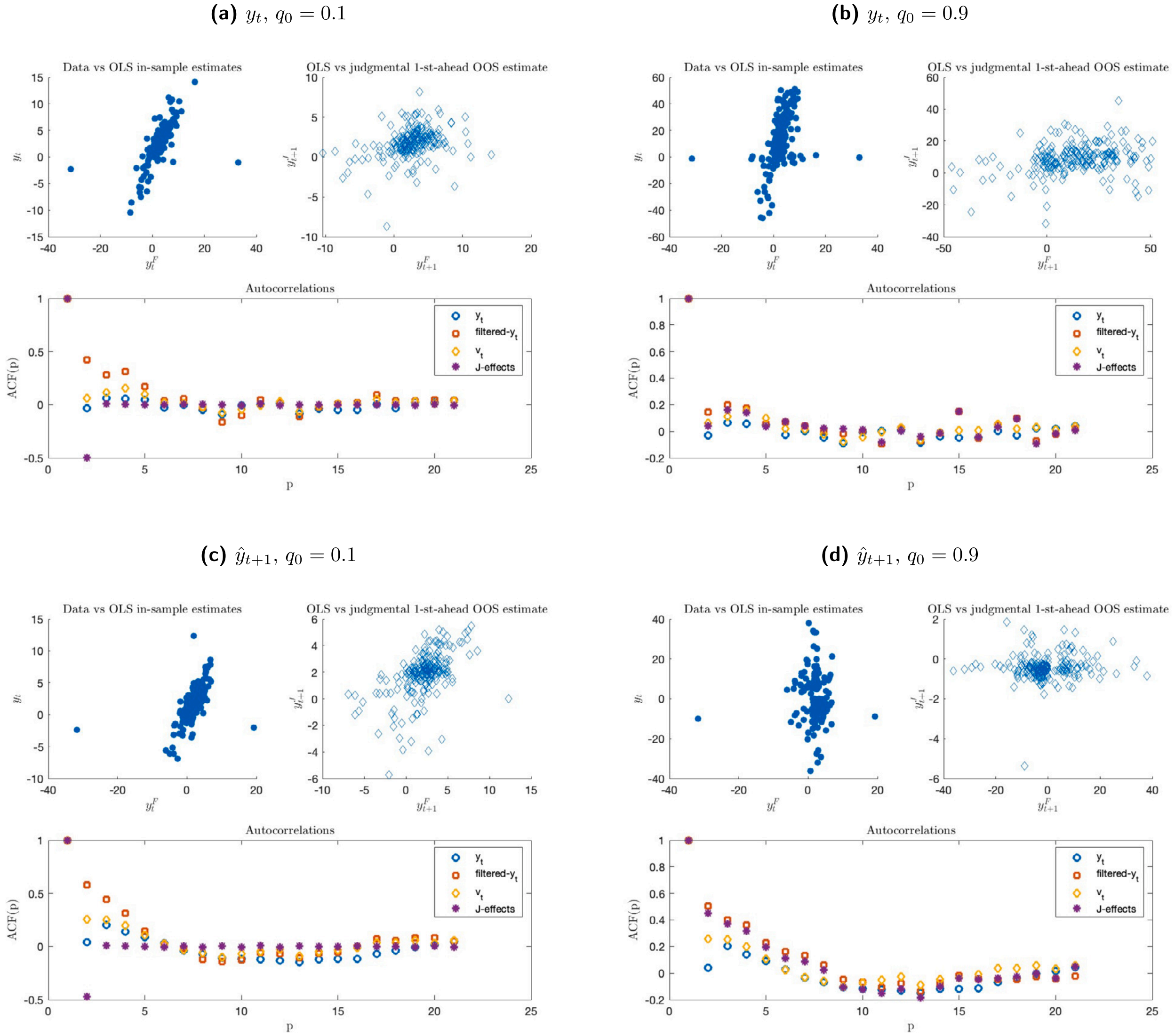
This stylized fact aligns with the findings emphasized by Afrouzi et al. (2023): a high AR coefficient regressor of a forecast on real data, in addition to results from extensive experiments, indicates persistence in expectations' overreaction. However, the statistical characterization of LM goes far beyond a simple AR coefficient magnitude. Thus, the following questions arise: *(i) Are there connections between judgment and LM? (ii) If yes, what is the effect of that relationship?* Our evidence allows us to answer positively to the first question: JB is associated with significant changes in the spectral density of the process. But LM implies, by definition, that the process has a stable spectrum. Hence, the association 'judgment–unstable spectrum' is symptomatic of spurious LM: the process, apparently long-range dependent if looking only at observed data, is in fact a classical stationary time series characterized by a significant amount of judgment. This has important consequences for econometric analysis: an improper assessment and modeling of judgment may lead econometricians to use an estimator for LM where it is unnecessary, and thus to incorrect inferences. This paper aims to bridge these two seemingly separate concepts by introducing a new family of models named Long Memory Dynamic Judgmental Protocol (LM-DJP).

The paper is organized as follows: Section 2 provides the contribution and the paper's status in the literature; Section 3 describes the model; estimation methods are summarized in Section 4; Section 5 reports the Monte Carlo exercise; real data applications are illustrated in Section 6; and Section 7 concludes. A supplement provides mathematical details, proofs, and further results; finally, Table 1 lists the acronyms.

2. Literature

The literature on LM is characterized by multiple definitions and, historically, was often confused with other features like non-stationarity. Not strangely, it is of the most conspicuous in the field of time series analysis and finds its origins in Natural Sciences and Engineering; see, *inter alia*, Beran et al. (2013) and Hassler (2018) for an updated overview. After seminal contributions by Hurst (1951), Granger (1966) and Mandelbrot and Van Ness (1968) in discrete and continuous-time, respectively, as well as the introduction of the fractional differentiation by Hosking (1981), the research on the topic in and allowed the econometric literature to enlighten several related issues: aggregation (Granger, 1980; Baillie, 1996; Zaffaroni, 2004), conditional heteroskedastic aspects (Baillie et al., 1996), maximum-likelihood estimation (Sowell, 1992), generalized integration (Granger and Ding, 1996), semi-parametric methods (Shimotsu and Phillips, 2005, 2006; Abadir et al., 2007; Shimotsu, 2010; Arteche, 2020), relationships with non-linear models

[2] To the best of our knowledge, the only exception is Brenna and Budrysb (2024), who use Bayesian inference to model judgmental dynamics.



NOTE: This figure displays the evidence of existence of a relationship of judgment on LM in Survey of Professional Forecasts. Each panel displays (i) in top left sub-panel, the association among levels y_t and simple OLS (in-sample) estimate $\hat{y}_t$ from an AR(2); (ii) in the top-right sub-panel, the analogue association among OLS-based out-of-sample forecasts ($\hat{y}_{t+1}$) obtained via standard Kalman filter and Judgmental-filter based forecasts y^F_{t+1} by Zanetti Chini (2023); (iii) in the bottom sub-panel, the estimated Autocorrelation Function for data (y_t), the judgmental-filtered series (y^F_{t+1}), its innovation component (υ_t) and the "judgment effects" (these last are defined as $y^F_{t+1} - \upsilon_t$). We consider two hypothetical scenarios for the initial level of judgment: low (10%, left column) and high (90%, right column). Moreover, we consider two kind of data: y_t (that real data observations, in top rows) and $\hat{y}_{t+1}$, that is one-quarter-ahead forecasts (bottom row).

Fig. 1. The Effects of Judgment in Long Memory on US survey data.

(Diebold and Inoue, 2001; Banerjee and Urga, 2005; Davidson and Sibbertsen, 2005), asymptotic analysis (Robinson, 1995; Robinson and Hidalgo, 1997), efficient estimation, testing and inference (Robinson, 1994; Dahlhaus, 2006), volatility modeling (Hurvich et al., 2005; Christensen and Nielsen, 2007; Corsi, 2009; Asai et al., 2012), factor modeling (Ergemen, 2023). This paper takes inspiration by the peculiar strand of literature that focuses on the role of economic agents' learning from their macroeconomic system and how this influences the long-run dynamics of economic time series by generating LM (Chevillon and Mavroeidis, 2017); and how the network structure of an economic system can be *per se* sufficient to cause LM without using other explanations (Schennach, 2018).

The literature on judgment is not systematic and spreads across several disciplines; see, for example, Goodwin and Fildes (2022). The Bayesian approach to Probability and Statistics (foundational in this framework) investigates how better collecting and summarizing opinions through opinion pooling or forecast combination; see, among many others, Genest and Schervish (1985); DeGroot and Mortera (1991); Geweke and Amisano (2011); Garratt et al. (2023). Game-theoretic literature also deals with judgment, but exclusively in theoretical aspects or centered on Bayesian calibration and claim validation; see Olszewski (2015). The most successful methods in frequentist Econometrics—see Hall and Mitchell (2007); Kapetanios et al. (2015) *inter alia*—connect elderly Bayesian intuitions with modern econometric modeling. However, this literature only partially addresses the generic research question of how

Table 1
List of Acronyms.

Acronym	Explanation	Acronym	Explanation
LM	long-memory	JB	judgmental bias
DSGE	dynamic stochastic general equilibrium	SVAR	structural vector autoregression
JF	judgmental filter	DKF	deformed Kalman Filter
FP	forecast producer	FU	forecast user
R	reality	FED	Federal Reserve Bank
SPF	survey of professional forecasters	OLS	ordinary least-squares
RGDP	real gross domestic product	JP	judgmental protocol
AR(p)	autoregression of order p	LM-DJP	long-memory dynamic judgmental protocol
DJP	dynamic judgmental protocol	ARFIMA	autoregressive fractionally integrated moving-average
ARMA	autoregressive moving-average	MLqE	maximum L_q-likelihood estimator
MLE	maximum likelihood estimator	AVAS	additivity and variance stabilization
ACE	alternating conditional expectation	LW	local Whittle likelihood estimator
SU-LMDJP	seemingly-unrelated LMDJP	ELW2	modified Whittle Likelihood Estimator
ELW	exact local Whittle likelihood estimator	GNP	gross national product
2SFEWL	two-step feasible Whittle likelihood estimator	DGP	data generating process

NOTE: This table reports the list of acronyms used in this paper in order of appearance in text.

to treat opinions, as it does not go into detail about the characterization of structural relationships or discuss long-term/strategic behavior effects. Operation Research-oriented literature (Lawrence et al., 2006; Lamont, 2002; Phillips et al., 2004; Clark and Friesen, 2009; DellaVigna, 2009; Cipriani and Guarino, 2014) provides some intuitions and evidence of the multiplicity of dangers associated with judgmental forecasting, such as motivational and strategic biases—mainly herding and (anti-)contrarianism—as well as on the inner nature of judgment, considering heuristics like representativeness, availability, anchoring, and adjustment as potential bias sources. In this literature, however, there is no unifying perspective on the rise of JB; an exception is represented by Clements (2018). Several strategies can enhance and/or mitigate judgmental bias. In particular, we follow the intuition by Manganelli (2009) to use a combination of statistical and judgmental factors in phase of estimation. However, this paper differs from the latter reference for the fact that judgment is the endogenous product of a dynamic system.

This paper provides a link among these two strands of literature for the first time, to the best of our knowledge. This link is ensured by two key ideas: a dynamic game among three agents (Forecast Producer, Forecast Users and Reality) that generalizes the static model by Vovk and Shafer (2005) and a robust estimation approach by Ferrari and Yang (2010). The economic intuition relies on the established literature on rational inattention; see Coibion and Gorodnichenko (2012, 2015); Bordalo et al. (2020); Ilut and Valchev (2023) *inter alia*. Finally, our modeling technique is based on a linear model. However, the literature that studies the connections among LM and nonlinear models is large and the degree of LM by Threshold ARMA, see Li et al. (2012) among others.

3. Theoretical framework

This section explains how to model judgment and long memory (LM) in time series. We adopt the idea that JB distorts any statistical functional based on the data available to the investigator. Theoretically, we need to verify a proportional relationship between JB and the degree of deformation: a low judgment should be associated with minor deformation of the functional under consideration, while a high judgment should produce significant distortion.

Definition 1. Let $y \in \mathbb{R}$ and q be the tuning parameter of an L_q-transform defined as follows:

$$L_q(y) = \begin{cases} \log(y) & \text{if } q = 1, \\ \frac{y^{1-q}}{1-q} & \text{otherwise.} \end{cases} \tag{1}$$

We define *data deformation* as any realization of $L_q(\cdot)$ with $q \neq 1$, and *judgment* as the realization of $1-q$ given $L_q(\cdot)$.

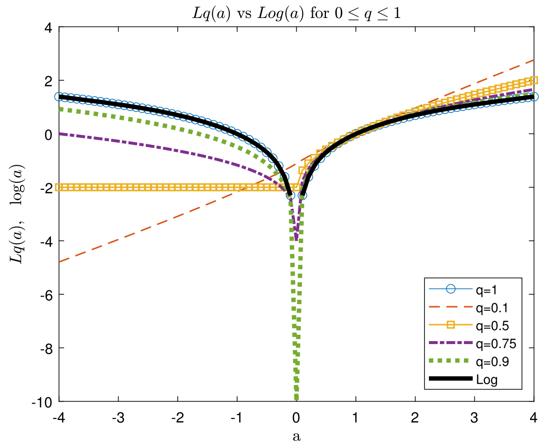
Example 1. In $L_q(y)$, q represents the deformation of $\log(y)$, and $1-q$ represents the judgment on y.

The way in which q distorts the "true" values of $\log(y)$ can be observed in Fig. 2. One of the most universally adopted functions linking data (i.e., y) to an unknown parameter is the Likelihood. Notably, a small deviation can significantly distort the log function. Moreover, if $q > 1$, the deformed function becomes discontinuous.

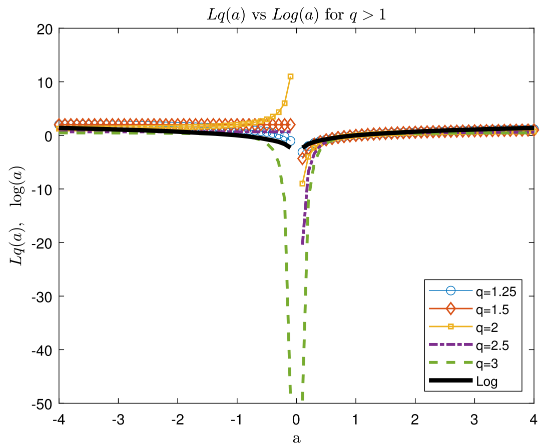
According to this definition, q is responsible for any contamination of a logarithmic-based transform of y—for example, the Likelihood—so it can be viewed as a proxy for judgment in the sense of Svensson (2005). Thus, the L_q-Likelihood (or Deformed Likelihood) estimator (MLqE) naturally arises from (1). In its original formulation, this estimator assumes q as given. The estimation of q is discussed in the Supplement.

The L_q-transform defined in (1) is a reparametrization of the classical Box-Cox transform. The Box-Cox transform is fundamental to the theory of generalized linear models, originally designed to handle potentially nonlinear or non-stationary datasets and is thus widely adopted in Statistics and Applied Sciences. Specifically, the Box-Cox transform generates a response for which the disturbance's

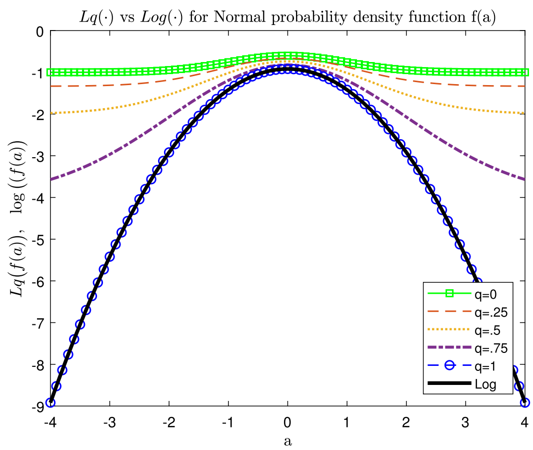
(a) Case 1: 0<q<1



(b) Case 2: q>1

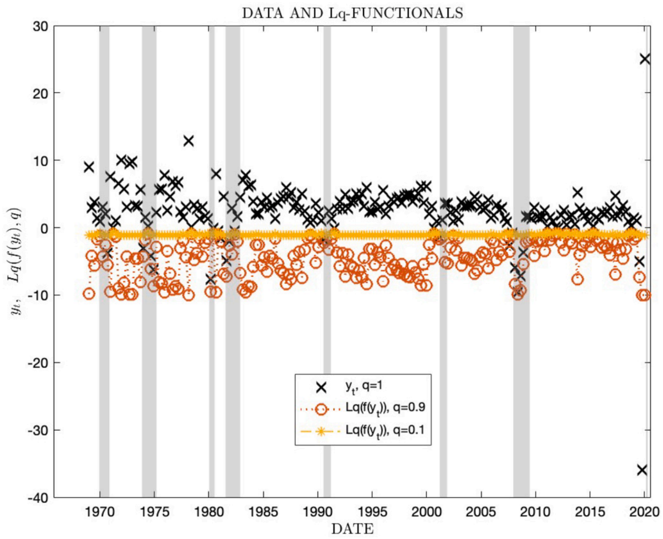


(c) The Deformed Logarithm of a Standard Normal density



NOTE: This figure displays the Lq-function applied to the numerical sequence $a = [-4; 4]$ for different values of q and compares it with the natural logarithm function.

Fig. 2. The Deformed Logarithm function.



NOTE: This figure displays the Lq-function applied to an DJS-AR(2) in US real data on GDP.

Fig. 3. The Deformed Likelihood in data.

variance is constant and approximately normally distributed. This feature can be limiting in datasets characterized by high time-varying conditional means and/or variances and non-normal distributions. In particular, the response of the L_q-functional is quite sensitive to different values of q at extreme observations. As shown in Fig. 3, the last two observations in the sample (2020:Q1–Q2) are approximately -35 and +20, respectively. When applying the L_q functional with $q = 0.9$ (low judgment), the response is around -10 for both observations, while it is -1 if $q = 0.1$ (high judgment). Conversely, the downturns in the rest of the sample are considerably lower in magnitude, and the estimated function with $q = 0.9$ tends to overestimate them, while the high-judgment estimating function remains stable around -1. This extreme oscillation of the L_q transform raises questions about the feasibility of using the Box-Cox transform. Various alternative functions have been developed to address this problem; see Atkinson et al. (2021) for a discussion.

To evaluate the effectiveness of our L_q-functional approach, we compare it with two alternative functionals: the alternating conditional expectation (ACE) and the additivity and variance stabilization (AVAS) methods by Breiman and Friedman (1985) and Tibshirani (1988), respectively. These are two of the most commonly used nonparametric methods for estimating a linear model with transformations applied to both the response variable and the explanatory variables:

$$g(y_t, k) = f(X_t, \beta) + \epsilon_t, \tag{2}$$

where k is a parameter vector defining a spline transform and disturbances ϵ_t are not necessarily normally distributed. As shown in Fig. 4, the effect of the transform on response variables and covariates (a set of four variables: industrial production, consumer price index of inflation, money velocity, and house prices) does not vary considerably between the two methods. The estimated transform on the covariate is linear, and applying ACE/AVAS to judgmental-filtered series (discussed in later sections of the paper) does not alleviate the overestimation issue observed with L_q-type functionals. Specifically, the "judgment effect" (i.e., the difference between judgmental-filtered series and the innovation part of the process) systematically overestimates the cyclical pattern of the data.

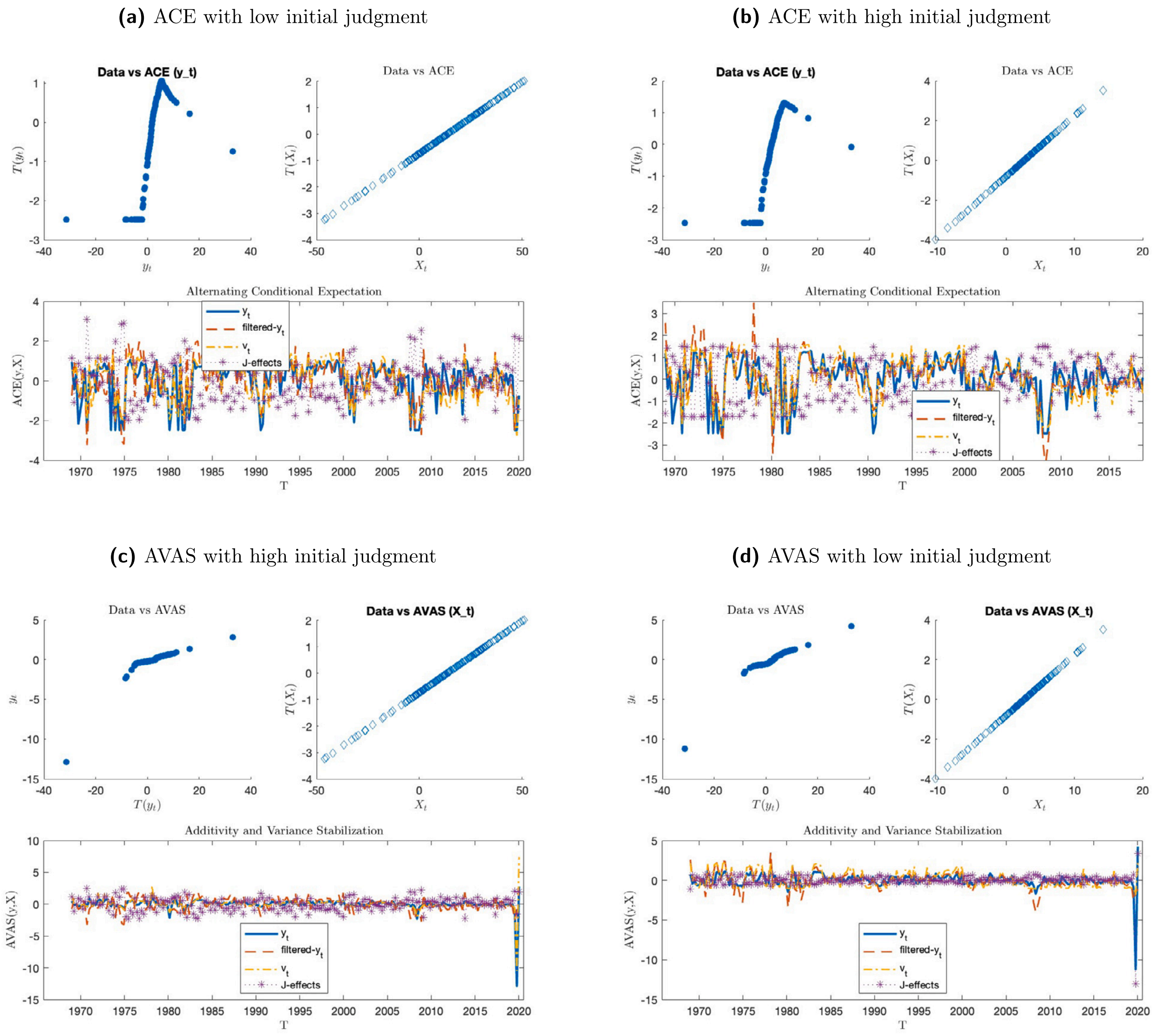
Ferrari and Yang (2010) shows that the Likelihood estimator based on the L_q-functional (see the Supplement for a summary) can be represented as a weighted mean where weights are proportional to the judgment: the lower the q, the more distorted the Likelihood becomes. This characteristic supports our confidence in using the L_q-type transform.

After defining the notions of deformation and judgment, we need to characterize their relationship and formation according to a simple microeconomic model. This can be formalized as follows:

Assumption 1. (i) There exists an Economy $\mathcal{E} := \{FP, FU, R\}$, where FP and FU are defined previously, and R represents Reality. (ii) Let $Z = [y^{FP}, y^{FU}, y^R]$ be the set of possible outputs of FP, FU, and R. Let $\Theta = [\theta^{FP}, \theta^{FU}]$ denote the parameter space containing the relevant information explaining the deliberations of FP and FU. Additionally, let $m_\Theta(R, Z)$ denote the "true" model relating Z to R. Finally, let a denote a Bayesian action defined over a quadratic loss function derived from $m_\Theta(\cdot)$. Then, there exists an a such that

$$\frac{\partial m_\Theta(\cdot)}{\partial a} = 0, \quad \frac{\partial m_\Theta(\cdot)}{\partial \theta^{FU}} = 0, \quad \text{and} \quad \frac{\partial m_\Theta(\cdot)}{\partial \theta^{FP}} = 0. \tag{3}$$

(iii) Let K^{FU} and K^{FP} denote the capital owned by FU and FP, respectively. Additionally, let $U^i(R, Z)$ for $i = \{FU, FP\}$ be the utility function associated with $m_\Theta(R, Z)$, and $U^i(R, R)$ be the utility function associated with the trivial model $m_\Theta(R, R)$ (i.e., a model with no judgment or differences in estimation among agents). Let $d_K = K^{FP} - K^{FU}$ denote the difference between the values



NOTE: This figure displays alternative transformations (ACE and AVAS) applied to the same DJS-AR(2) model used in Fig. 1 to US real data on GDP. Each panel display (i) in top left sub-panel, the association among levels y_t and ACE-transform of same y_t; (ii) in the top-right sub-panel, the analogue association among covariate X_t and ACE-transform of X_t; (iii) in the bottom sub-panel, the estimated ACE/AVAS function of the judgmental-filtered series y^F_{t+1}, its innovation component υ_t and the "judgment effects" defined as $y^F_{t+1} - \upsilon_t$. We consider two hypothetical scenarios for the initial level of judgment: low (10%, left column) and high (90%, right column).

Fig. 4. Alternative transformations.

of Z and R, and $D(Z,R) = U(R,R) - U(R,Z)$ denote the discrepancy function (i.e., the difference between the actual utility and the maximum possible utility in the system $\mathcal{E}$). Then, there exists $K^i > 0$ such that

$$K^i_t = K^i_{t-1} + D(Z,R). \tag{4}$$

(iv) Let $c(\Theta, K^i)$ be the cost functions for each agent, respectively. Then

$$\frac{\partial c^{FU}(\Theta,K)}{\partial K^{FU}_t} = \frac{\partial c^{FP}(\Theta,K)}{\partial K^{FP}_t} \quad \text{and} \quad \left\{ \frac{d_K}{D(Z,R)} \right\}_{t=1}^{T} = o_p(1). \tag{5}$$

In part (i), FP may represent external experts or internal officers of the Central Bank, though the latter have no role in the final decision. Part (ii) implies that FU's choice does not affect Reality when the process repeats, nor does FP respond to FU throughout the process, assuming no strategic behavior by the agents. Part (iii) indicates that each player can increase their capital if the discrepancy D is properly designed. Part (iv) means that Reality is impartial to the judgments of FU and FP. In other words, while both players face the same marginal cost, as the game progresses, Reality is always able to nullify the relative capital gain of the players.

Remark 1. A special case of $D(\cdot)$ is $H(R) := D(R, R)$ (i.e., no divergence between players' utility outputs and R), known as the *entropy function*. This function describes how the loss of each agent is minimized (or equivalently, their utility is maximized) for all possible convex functions and all players in $\mathcal{E}$. The form of $D(\cdot)$ and $H(\cdot)$ characterizes the dynamics of the system; see Zanetti Chini (2023) for details.

Definition 2. Let y_t be a time series to be analyzed; $X_t \in \mathbb{R}^k$ be a set of explanatory variables; Ψ and Ξ be the parameters of FU and FP, respectively; $U(\cdot)$ be the utility function that depends on Ψ or Ξ; $\mathcal{D}(X, Y)$ be the Bregman-type divergence between X and Y; $\mathcal{H}(\cdot,\cdot)$ be the entropy measured in the system for the same variables; d be the order of fractional differencing; and q be the deformation parameter. Then we define the Long-Memory Judgmental Protocol (LMJP) as the n-tuple $\mathcal{LMJP} := \{y_t, X_t, \Psi, \Xi, U, D, H, q, d\}$, which operates according to the following steps: for $t = 1, \dots, T$,

1. FU and FP acquire, evaluate, and interpret data from a collection of explanatory variables (X_t) to conduct out-of-sample inference on a target variable (y_t) using parameters Ψ and Ξ, respectively, while forming their own utility functions ($U(\Psi)$ and $U(\Xi)$).
2. FP makes (possibly biased) projections $\hat{y}_t$ of the objective variable under alternative assumptions about the explanatory factors and/or scenarios.
3. FU receives the projection sets, adds judgment $\tilde{\hat{y}}_t$ to them, and selects the best projection based on their objectives and information.
4. FU releases the official prediction for the next period, $\hat{y}_{t+1|t}$.
5. Reality reveals itself and is characterized by a fractional differencing parameter d.

Remark 2. A more restrictive version of the LMJP can be established by omitting X (i.e., any explanatory variables) from step 1 in Definition 2. In this case, the judgmental protocol consists of FU/FP forecasts on possible values of y_t in $t+1$ with no relation to any measurable knowledge apart from y_t. This implies that any deviations from the optimal forecasts $y_{t+1|t}$ are attributed solely to judgment rather than additional information. Consequently, the n-tuple $\{y_t, X_t, \Psi, \Xi, U, D, H, q, d\}$ defines a *seemingly unrelated LMJP* (SU-LMJP).

Special attention must be given to the initial value of q, denoted q_0, since it is generally assumed that $q_0 \neq \hat{q}$. Specifically, we rely on the following assumptions:

Assumption 2. (i) $q \in (0, 1]$;
(ii) The value of q_0 is independent of subsequent realizations of q_t.

Assumption 1 part (i), combined with Definition 1, implies that 0 represents pure judgment and 1 represents no judgment. Assumption 2 part (ii) is necessary to ensure ease of statistical treatment. Without this assumption, we would implicitly assume the existence of strategic behavior among the players in the judgmental protocol, which is left for future research.

Assumption 3. (i) d is the fractional integration parameter of a fractionally integrated autoregressive moving average (ARFIMA) model:

$$\begin{aligned} \Phi(L)(1-L)^d y_t &= \Theta(L)u_t, \quad u_t \sim \text{i.i.d.}(0, \sigma_u^2), \\ (1-L)^d &= \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} L^j, \end{aligned} \tag{6}$$

where $\Gamma(\cdot)$ is the Gamma function, L is the lag operator; $\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\Theta(L) = 1 + \theta_1 L + \cdots + \theta_r L^r$ are the autoregressive and moving-average polynomials with all their roots outside the unit circle and no common factors; and $(1-L)^d$ is the fractional differencing operator characterizing long memory (LM).
(ii) $d \in (0, 1/2)$.
(iii) There exists an auto-covariance function γ such that: (a) $\gamma(0) = 0$; (b) $||\gamma_k|| \leq \gamma(0)$ for all integers k; (c) $\gamma(\cdot)$ is an Hermitian function.

Assumption 3 part (i) allows us to focus on a well-known and feasible econometric framework where judgmental issues are nested. We restrict our attention to this family of models primarily for ease of statistical treatment (i.e., a known state-space representation) and for comparison with the economic literature. However, the ARFIMA(p, d, r) model is not the only family of models capable of efficiently handling long memory. Corsi (2009) introduces an alternative model that allows econometricians to capture long memory without using fractional operators and high numbers of lags. However, the mixed nature of this model would require a revision of the entire state-space representation and associated statistical inference, which is left for future research. Assumption 3 part (ii) ensures stationarity of the ARFIMA(p, d, r) model. Assumption 3 part (iii) is a standard assumption on the dynamics of the process that allows the Herglotz theorem to hold. Remarkably, if $d > 0$, the autocorrelations decay at a hyperbolic rate (and are therefore no longer absolutely summable), so the process is considered to exhibit long memory as opposed to the considerably quicker exponential rate in the case of weak dependency. If $d = 0$, the spectral density is bounded at the origin, and the process is an AR process with minimal dependency (short memory).

An alternative definition based on the frequency domain is provided by Beran et al. (2013); see the next section for details.

Then we are going to verify the following

Proposition 1. *Let Assumptions 1–3 hold and $y \sim ARFIMA(p,d,0)$–that is a pure (fractionally integrated) autoregression for ease of treatment and with no loss of generality. Then:*

(a) The LMJP can be represented as the following dynamic system, named (Long-Memory) Dynamic Judgmental Protocol ((LM)DJP, henceforth):

$$\begin{aligned} y_t &= Z_t\alpha + X_t\beta + G_t\epsilon_t, \qquad \epsilon_t \sim iid(0,\sigma_\epsilon^2) \\ \alpha_{t+1} &= T_t\alpha_t + W\beta + H_t\eta_t, \qquad \eta_t \sim iid(0,\sigma_\eta^2); \\ Z &= [1,\ \mathbf{0}_{m-1}]; \qquad X = [1,\ \mathbf{0}_{k-1}]; \qquad H = \boldsymbol{\phi}' \\ G &= \begin{bmatrix} \boldsymbol{I}_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix}; \qquad T = \begin{bmatrix} \boldsymbol{\phi}' \\ G \end{bmatrix}; \qquad W = [\boldsymbol{\phi}',\ G]', \end{aligned} \tag{7}$$

where Z_t is a $(m \times 1)$ vector of fixed effects, X_t a $(1 \times k)$ vector of covariates, β a $(k \times 1)$ vector of parameters, α_t an $(m \times 1)$ vector of states, T an $(m \times m)$ matrix of fixed coefficients, G an $(m \times g)$ matrix, η_t a $(g \times 1)$ vector of disturbances and $\boldsymbol{\phi}$ a $(p+1)$ vector of AR parameters. The initial conditions are:

$$\begin{aligned} &\alpha_0 = [\mathbf{0}_m]; \quad \beta_0 = [\mathbf{0}_k]; \quad \boldsymbol{I}_{m^2} = \boldsymbol{I} \otimes [T,\ T]; \\ &H^2 = HH'; \qquad vec(P) = \boldsymbol{I}_m^{-1} H^2. \end{aligned} \tag{8}$$

Moreover, (7) can be modeled via ARFIMA(p, d, r) models.

(b) The estimation by Deformed Kalman Filter (DKF, henceforth) has the following properties: (i) isomorphism: it is isomorphic to a linear state-space system; (ii)q-dependency: its recursions depend on q parameter.

Proof. See the Supplement. □

Proposition 2. *The judgmental parameter q affects the Power Spectrum close to the origin on top of d.*

Proof. See the Supplement. □

A direct consequence of part (i) in Proposition 2 is that the state-space system representing the LMJP does not require any nonlinear over-parametrization for estimation using the DKF.

Proposition 2 provides a testable hypothesis. To this end, we consider the following hypothesis test:

$$H_0 : d = 0 \quad \text{versus} \quad H_1 : d > 0, \tag{9}$$

which can be tested via non-standard inference—see the next Section.

4. Estimation and modeling strategy

This section is dedicated to the estimation methods. Subsections 4.1 and 4.2 outline the main properties of Long Memory (LM) and ML_q-Likelihood (MLqE) estimators, respectively, summarizing the works of Ferrari and Yang (2010) and Grassi and Santucci de Magistris (2014). Finally, subsection 4.3 describes the modeling strategy for the LMDJP previously introduced.

4.1. Estimating long memory

The spectral density of the process can be represented as

$$f_\Phi(\lambda) \sim G|\lambda|^{-2d} \quad \text{as} \quad \lambda \to 0, \tag{10}$$

where G corresponds to the spectral density (at the origin) of an AR(p) process. The following standard assumptions are necessary to ensure the effectiveness of the spectral representation theorem for any real-valued stationary process:

Assumption 4. (i) $f(\lambda) = f(-\lambda)$; (ii) $f(\lambda) \geq 0$; (iii) $\int_{-\pi}^{\pi} f(\lambda)\,d\lambda < \infty$; (iv) $G(\lambda)$ exists for all λ.

See Brockwell and Davis (1991, Ch 4) for details.

Such a function can be estimated using the Local Whittle Likelihood (LW):

$$\begin{aligned}
&\sum_{j=1}^{m} \log f_{\Phi}(\lambda_j) + \sum_{j=1}^{m} \frac{I_{\Phi}(\lambda_j)}{f_{\Phi}(\lambda_j)}, \\
&I(\lambda) = |w(\lambda)|^2, \\
&w(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{T} y_t e^{it\lambda}, \\
&\lambda_j = \frac{2\pi j}{n},
\end{aligned} \tag{11}$$

where $w(\lambda)$, $I(\lambda)$, and λ_j are the discrete Fourier transform, the periodogram, and the frequency of y_t, respectively, and m is the truncation parameter. The choice of m is a non-trivial issue. Several strategies can be employed, including information criteria or simulation. In our application, we used the Bayesian Information Criterion (BIC) for this purpose; however, further investigation is needed to better understand the influence of this parameter on q and the overall modeling strategy.

Fig. 5 provides a simple illustration of the form of the DJP specified as AR(1). We consider four different Maximum Likelihood estimators of d: the Local Whittle Likelihood estimator (LW), the Exact Whittle Likelihood (ELW), the Modified Whittle Likelihood (ELW2), and the Two-Step Feasible Whittle Likelihood (2SFEWL). For details, see Robinson (1995); Shimotsu and Phillips (2005, 2006); Shimotsu (2010).

4.2. Estimating judgment

Let $y_1, \dots, y_T$ be an i.i.d. sample from $p(y_i, \theta_0)$, where $\theta_0 \in \Theta$, and all elements of Θ are part of the LMDJP defined in Section 3. The maximum L_q-estimator (MLqE) of θ_0 is given by

$$\hat{\theta}_T \doteq \max_{\tilde{\theta}\in\Theta} \sum_{t=1}^{T} L_q[p(y_t;\theta)], \quad q > 0, \tag{12}$$

where $L_q(x)$ is a L_q-transform. If $p(\cdot)$ is Gaussian, its Deformed Likelihood is:

$$\mathcal{L}^q(\theta; y_t) = -0.5\left[q_0 T \log(2\pi) + (\log p(y_t) + u_t^2)^q\right], \tag{13}$$

where $\theta = [\mu, \sigma^2]$, $u_t = (y_t - \mu)/\sigma$, and π is the usual Archimedean constant. The estimated version uses $\hat{q}$ and $\hat{u}_t$ instead of q and u_t. Moreover, (12) results from the maximization of:

$$\sum_{t=1}^{T} w_t U(y_t, \theta) = 0, \tag{14}$$

which is a weighted version of the likelihood equation with $U(y_t, \theta) = \frac{\partial p(y_t;\theta)}{\partial\theta} / p(y_t; \theta)$ and weights $w_t = p(y_t; \theta)^{1-q}$. When $q < 1$, data points with high likelihoods are assigned larger weights. As q approaches 1, the MLqE converges to the standard Maximum Likelihood estimator. Typically, outliers are associated with very small weights. According to Fig. 2, a small deviation from 1 can drastically alter the log-likelihood. However, such contamination is not uniform. For instance, analytic results from Ferrari and Yang (2010) show that if $p(\cdot,\cdot)$ is Gaussian, the estimated mean $\hat{\theta}$ does not depend on q. The parameter q measures judgmental bias in the estimated model due to FP (or FU) *individually*.

Finally, it can be proven that the q-Entropy coincides with the H-function corresponding to the DJP.

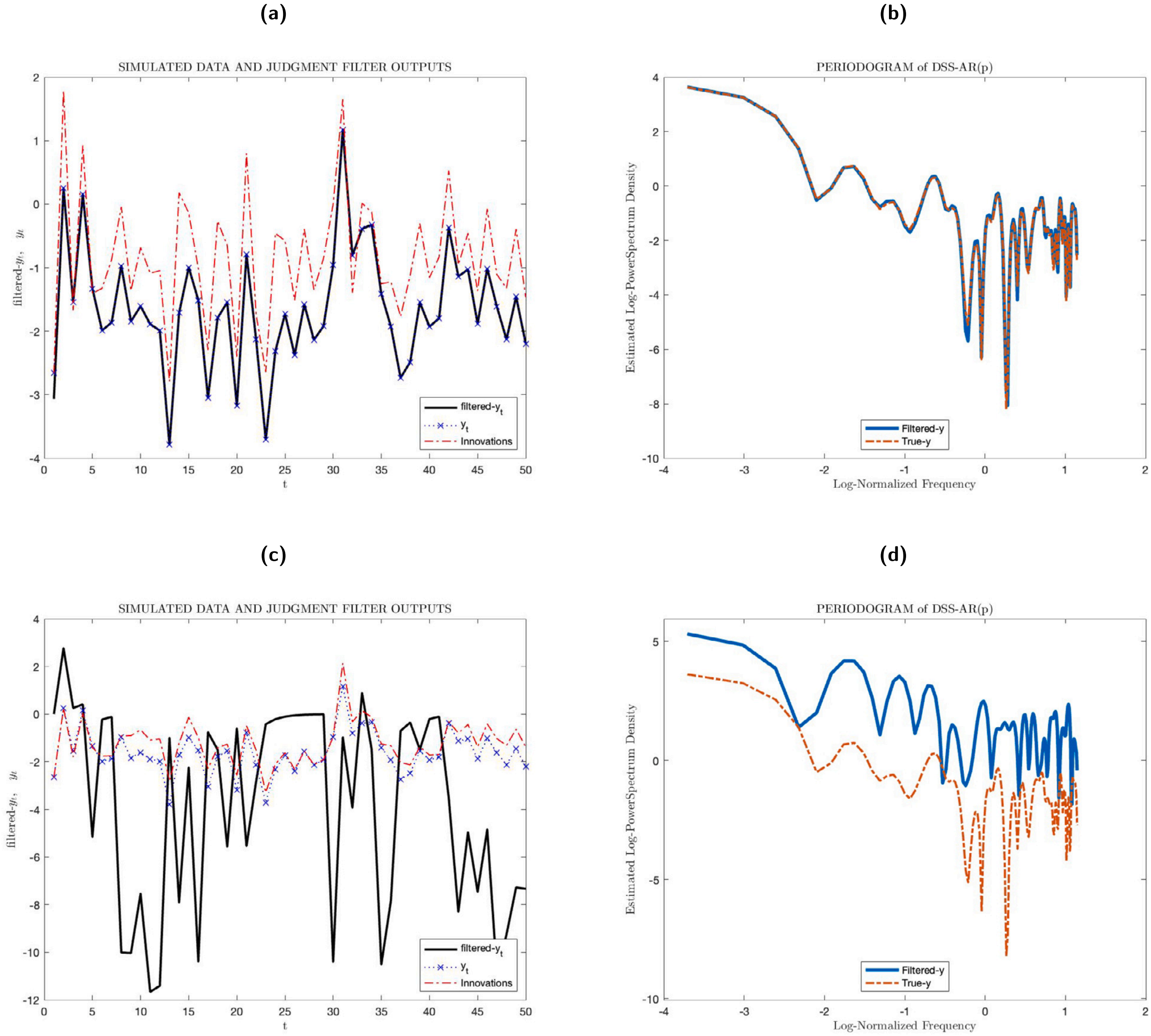
In this setting, q effectively affects the variance of the process, as confirmed by the behavior of the Power Spectrum, see Section 5. The following assumption ensures that the spectral representation theorem holds for Z_t (from which q is estimated) and that the existence of $q \neq 1$ does not impede inference on the power spectrum:

Assumption 5. Let a and b be two linear combination coefficients 3×1 vectors applied to Z_t and e^{it}. Then:

1. There exists an orthogonal increment process $\mathcal{Z}(\lambda)_t$, $\lambda \in [-\pi, \pi]$, associated with Z_t.
2. There exists a mapping $\mathcal{T}$ that provides an isomorphism between Z_t and $e^{it\lambda}$.
3. $\left\| \sum_{j=1}^{T} Z'_{tj} a_j - \sum_{k=1}^{k} Z'_{tk} b_k \right\| = 0$ for $j \neq k$.

Point (i) is necessary to identify an equivalent process for Z_t in the frequency domain. Point (ii) ensures that existing spectral methods can be applied to the LMDJP. Point (iii) is necessary to establish an equivalent finite second moment condition for linear combinations of Z_t and forms the basis for proving the isomorphism in point (ii). We can now ensure the feasibility of spectral estimation with the following:

Proposition 3. *Under Assumptions 2–5, the LMDJP is stationary and has spectral density* $f_{\Psi}(\lambda)$.



NOTE: This figure displays an example of simulated DJP-AR(1) process (left panels) and estimated periodograms (right panels) for two different values of $q_0 = 1$ (upper panels) and $q_0 = 0.1$ (lower panels). In all the cases the process is: $y_t = \phi = 0.5y_{t-1} + \epsilon_t$, $\epsilon_t \sim N(0, 1)$ and T = 50.

Fig. 5. Examples of DJP-AR(1) and associated power-spectra for several q_0 and d_0.

Proof. See Supplement. □

4.3. Modeling strategy

Two modeling strategies are possible: a 'Judgment-First' (J-First) approach consists in estimating the autoregressive order governing the DJP; alternatively, a 'LM-First' approach involves estimation of $\hat{q}_t$ via DKF explained in Supplement and using it as initial value to fit (6) via spectral density estimation exposed in Appendix. These two approaches are defined as follows:

Definition 3.

1. **LM-First**
   (i) Specify a linear ARFIMA(p, d, r) model for LMDJP; eventually, if some part of the structure is missed (for example, no Xs are observed), specify a linear ARFIMA(p, d, r) model for the existing parts of the LMDJP.
   (ii) Test the hypothesis of $d = 0$ in the specified ARFIMA(p, d, r) model.
   (iii) If the test is rejected, estimate the model $\hat{y}_t^{ARFIMA}$.

(iv) Run the DKF on $\hat{y}_t^{ARFIMA}$ and get $\hat{y}^f_{ARFIMA}$.
(v) Use the filtered $\hat{y}^f_{ARFIMA}$ to perform some diagnostic on $\hat{d}$ by computing the Power Spectrum of the estimated process.

2. **J-First**
(i*) Specify a linear ARFIMA(p, d, r) model for LMDJP; eventually, if some part of the structure is missed (for example, no X are observed), specify a linear ARFIMA(p, d, r) model for the existing parts of the LMDJP.
(ii*) Run the DKF on the LMDJP in step (i*) and get $\hat{y}_t^f$.
(iii*) Test the hypothesis of $d = 0$ in $\hat{y}_t^f$.
(iv*) If the test is rejected, estimate the model $\hat{y}_t^{ARFIMA}$ corresponding to the LMDJP.
(v*) Use the estimated $\hat{y}_t^{ARFIMA}$ to perform some diagnostic on $\hat{d}$ by computing the Power Spectrum of the estimated process.

5. Monte Carlo simulation

In this section, we apply the LMDJP to Monte Carlo simulated data to verify the effectiveness of Proposition 1. We consider two different data generating processes (DGPs) for ease of comparison with Zanetti Chini (2023):

$$y_{1,t} = 0.5y_{1,t-1} + \epsilon_{1,t}, \quad \epsilon_{1,t} \sim N(0,1) \tag{15}$$

and

$$y_{2,t} = 0.5y_{2,t-1} - 0.88x_{t,2} + \epsilon_{2,t}, \quad \epsilon_{2,t} \sim N(0,1), \tag{16}$$

where, in both (15) and (16), the total number of draws is 5,000. Both DGPs are run with a set of starting q-values ($q_0 = \{0.1, 0.5, 1.0\}$) corresponding to extreme, medium, and no amount of judgmental bias (JB), and d-values ($d_0 = \{0.1, 0.5, 1.0\}$) corresponding to long-memory, moderate memory, and differenced processes.

It is crucial to note the perspective of the analysis: (15) represents what the forecaster (FP) observes and analyzes, as the autoregressive model is her subjective choice. Specifically, $y_{1,t}^{(i)}$ is a linear autoregressive model with essentially stationary behavior, focusing on the impacts of the starting value of q_0. This could be seen as a macroeconomic indicator evaluated with a specific quantity of JB to mitigate the risk of loss in the next period. Conversely, $y_{2,t}^{(i)}$ presents a more comprehensive scenario: the autoregression includes an explanatory variable X_t alongside the initial judgment q_0. If y_t represents a time series of final announcements from a forecasting unit (FU), the entire DJP can be interpreted as a dynamic system where the FP's output (in this case, x_t) is an input that coexists with reality. The difference between them forms the basis for an ex-post evaluation of FU through a utility function, integrated into the estimation step via the L_q-Likelihood. Thus, q_0 reflects an a priori belief of FU concerning which FP adjusts her projections.

Tables 2 and 3 present the bias and empirical power of the test for spurious long-memory (LM) across all estimators mentioned in Section 4.1, using sample sizes $T = 100$ and $T = 1000$, respectively. Two notable observations arise: (i) the progressive reduction in bias with increasing d_0 and q_0; and (ii) the sensitivity of the test to the algorithmic design, evidenced by significant differences in empirical power between the J-First and LM-First approaches. Specifically, all estimators perform similarly, with a slight advantage for ELW and ELW2 over LW and 2SFELW, which are the most effective in both approaches. The LW estimator tends to be underpowered, particularly in the "LM-First" approach when the initial q_0 is low. Additionally, with a small sample size, variations in power due to different initial amounts of judgment are considerable within the same modeling strategy, while these differences decrease in longer samples. The variation in d_0 values shows minimal impact within the same q_0. The average $\hat{q}$ varies notably between estimation methodologies and sample sizes. However, this does not necessarily indicate a flaw in the estimation approach per se, as q_0 is not the true judgmental level but merely an initial value.

Fig. 5 illustrates the average Power Spectrum resulting from the DGP with $d_0 = 0$ fixed. Unlike the pure autoregressive case with no judgment, the spectra exhibit a shift whose magnitude depends on the presence of exogenous regressors (more than 100% for pure AR, around 60% in the ARX case).

Why do the two modeling strategies produce different power levels for the LM test? Figs. 6 and 7 offer an explanation: when q_0 is not 1 (i.e., when there is an initial judgment level), it significantly increases the power spectral densities of the filtered processes in most cases.

Another important result, detailed further in the Supplement, is that the relationship between judgment and LM is bidirectional: parameterizing long-memory leads to an underestimation of judgment effects. The weighting function w_t changes significantly if d is not correctly included in the model.

6. Empirical applications

This section applies the LMDJP to several real datasets relevant for macroeconomists and financial economists. Subsection 6.1 illustrates the application to Nelson and Plosser data and Subsection 6.2 reports the results for Federal Reserve Bank's survey data.

6.1. Nelson-Plosser data

The LMDJP is applied to three real-world case studies. The first case is a purely autoregressive SU-LMDJP on macroeconomic data using the Nelson and Plosser dataset Nelson and Plosser (1982), a standard reference in macroeconometrics and time series

Table 2
Simulation results for the Long Memory estimators and tests in LMDJP for T = 100.

LMDJP-AR(1), "Judgment First"

q_0	d_0	Mean $\hat{q}$	Mean $\hat{d}$ (bias)				Empirical Power (× 100)			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.20	0.22	0.21	0.21	0.19	30.5	31.2	36.1	31.6
0.50	0	0.56	0.17	0.15	0.16	0.18	40.8	42.4	40.4	44.1
1.00		0.79	0.10	0.09	0.10	0.14	51.9	62.7	61.4	65.4
0.10		0.30	0.22	0.15	0.15	0.18	42.2	43.2	40.5	40.1
0.50	0.5	0.67	0.19	0.17	0.12	0.15	64.8	66.4	63.4	63.1
1.00		0.70	0.13	0.15	0.14	0.10	55.4	52.7	52.2	50.8
0.10		0.31	0.14	0.15	0.13	0.15	54.5	54.5	47.3	42.9
0.50	1	0.67	0.11	0.10	0.08	0.11	55.1	54.9	49.2	53.0
1.00		0.78	0.12	0.13	0.05	0.08	70.4	70.3	63.4	63.5
LMDJP-ARX(1), "Judgment First",										
0.10		0.21	0.19	0.21	0.23	0.19	30.8	41.7	38.1	31.0
0.50	0	0.77	0.16	0.16	0.17	0.18	41.2	42.9	43.0	41.0
1.00		0.70	0.11	0.12	0.14	0.15	84.4	63.6	66.1	61.6
0.10		0.25	0.23	0.23	0.18	0.20	44.2	43.2	46.4	43.2
0.50	0.5	0.65	0.17	0.15	0.15	0.15	60.6	63.0	58.3	58.3
1.00		0.72	0.18	0.17	0.13	0.13	64.5	64.0	61.3	63.5
0.10		0.25	0.21	0.20	0.15	0.18	48.9	48.5	45.8	51.4
0.50	1	0.57	0.14	0.18	0.12	0.16	66.1	66.4	65.3	61.7
1.00		0.80	0.09	0.10	0.08	0.11	79.4	74.2	76.1	81.9

LMDJP-AR(1), "LM First"

q_0	d_0	Mean $\hat{q}$	Mean $\hat{d}$ (bias)				Empirical Power (× 100)			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.39	0.26	0.27	0.25	0.25	17.9	22.6	39.4	31.3
0.50	0	0.55	0.22	0.23	0.21	0.22	29.2	25.9	40.6	32.0
1.00		0.59	0.10	0.14	0.15	0.11	32.4	36.8	45.6	32.4
0.10		0.42	0.22	0.22	0.23	0.25	29.0	27.6	28.0	25.4
0.50	0.5	0.64	0.22	0.20	0.21	0.18	26.3	27.6	28.9	25.8
1.00		0.70	0.20	0.12	0.13	0.11	20.3	28.6	27.3	26.7
0.10		0.44	0.20	0.22	0.22	0.23	50.0	48.2	36.7	39.6
0.50	1	0.65	0.16	0.20	0.15	0.18	54.3	49.3	38.9	31.3
1.00		0.67	0.12	0.18	0.14	0.13	58.8	49.6	38.5	33.0
LMDJP-ARX(1) "LM First"										
0.10		0.45	0.27	0.28	0.23	0.36	20.2	25.3	31.8	24.9
0.50	0	0.66	0.35	0.24	0.26	0.18	19.5	25.1	38.1	25.8
1.00		0.69	0.22	0.16	0.16	0.15	23.9	35.5	29.0	26.9
0.10		0.45	0.35	0.23	0.28	0.33	20.5	21.2	21.1	15.9
0.50	0.5	0.60	0.33	0.26	0.26	0.38	20.9	33.8	39.3	26.7
1.00		0.70	0.18	0.13	0.24	0.31	22.6	34.6	30.7	36.5
0.10		0.46	0.28	0.31	0.25	0.25	60.0	33.4	58.5	21.0
0.50	1	0.61	0.28	0.27	0.27	0.22	31.0	31.5	29.9	30.2
1.00		0.70	0.20	0.22	0.21	0.20	28.8	44.4	42.5	40.2

NOTES: This table reports the MonteCarlo average estimates of the strategic judgmental parameter (column 3) and bias of the four estimators of long memory parameters, jointly with their empirical powers (columns from 4 to the end) for different initial values of strategic judgment and memory. The upper part displays the results where $\hat{d}$ is estimated before applying Judgmental Filter, while the lower half reports the results for the case that $\hat{d}$ is estimated after that Judgmental Filter is run on data. The DGP is explained in Section 6.

analysis. This dataset includes 14 macroeconomic variables for the U.S. economy, such as GNP and its deflator, employment, industrial production, and money velocity, collected annually up to 1970.

The model structure is simplified: the forecaster (FP) is represented by an AR(p) model, while the forecast updater (FU) is represented by an AR(p^*), with $p \neq p^*$—indicating that the two agents differ only in their autoregressive order. Any differences between FU and FP outputs should be attributed to dynamic misspecification. Thus, the judgmental parameter q is assumed to be approximately 1 everywhere, and the DKF is assumed to function as a standard (robust) Kalman Filter. Additionally, d is expected to vary across time series, independent of the estimation strategy (i.e., d should vary due to process characteristics).

Table 3
Simulation results for the Long-Memory estimators and tests in DJP for T = 1000.

LMDJP-AR(1), "Judgment First"

q_0	d_0	Mean $\hat{q}$	Mean $\hat{d}$				Empirical Power (× 100)			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.53	0.14	0.15	0.14	0.14	55.7	60.0	60.5	65.5
0.50	0	0.76	0.16	0.14	0.13	0.10	45.6	56.5	59.5	65.6
1.00		0.90	0.02	0.03	0.03	0.02	60.5	60.4	67.1	65.5
0.10		0.53	0.11	0.12	0.11	0.10	59.5	59.5	58.2	58.7
0.50	0.5	0.75	0.12	0.08	0.06	0.05	86.6	83.4	78.1	85.6
1.00		0.99	0.02	0.01	0.02	0.01	84.2	83.3	78.7	85.6
0.10		0.42	0.12	0.10	0.14	0.11	80.2	81.2	74.5	75.5
0.50	1	0.75	0.10	0.08	0.07	0.05	85.9	85.0	80.5	85.6
1.00		1.00	0.00	0.01	0.01	0.00	95.8	96.0	95.8	95.6
LMDJP-ARX(1), "Judgment First"										
0.10		0.53	0.05	0.13	0.17	0.11	99.5	99.6	64.5	65.8
0.50	0	0.78	0.04	0.09	0.14	0.12	100	100	62.1	65.8
1.00		0.90	0.06	0.06	0.12	0.03	100	100	63.5	65.8
0.10		0.65	0.13	0.13	0.12	0.15	56.1	65.7	48.1	49.8
0.50	0.5	0.78	0.08	0.09	0.09	0.08	66.7	65.4	56.7	57.7
1.00	.	0.90	0.03	0.04	0.03	0.02	95.3	99.6	94.6	58.7
0.10		0.66	0.11	0.08	0.06	0.08	66.3	65.1	68.6	65.6
0.50	1	0.78	0.09	0.09	0.10	0.09	66.2	70.4	70.6	65.6
1.00		0.90	0.01	0.02	0.02	0.01	99.4	99.3	89.3	75.7

LMDJP-AR(1), "LM First"

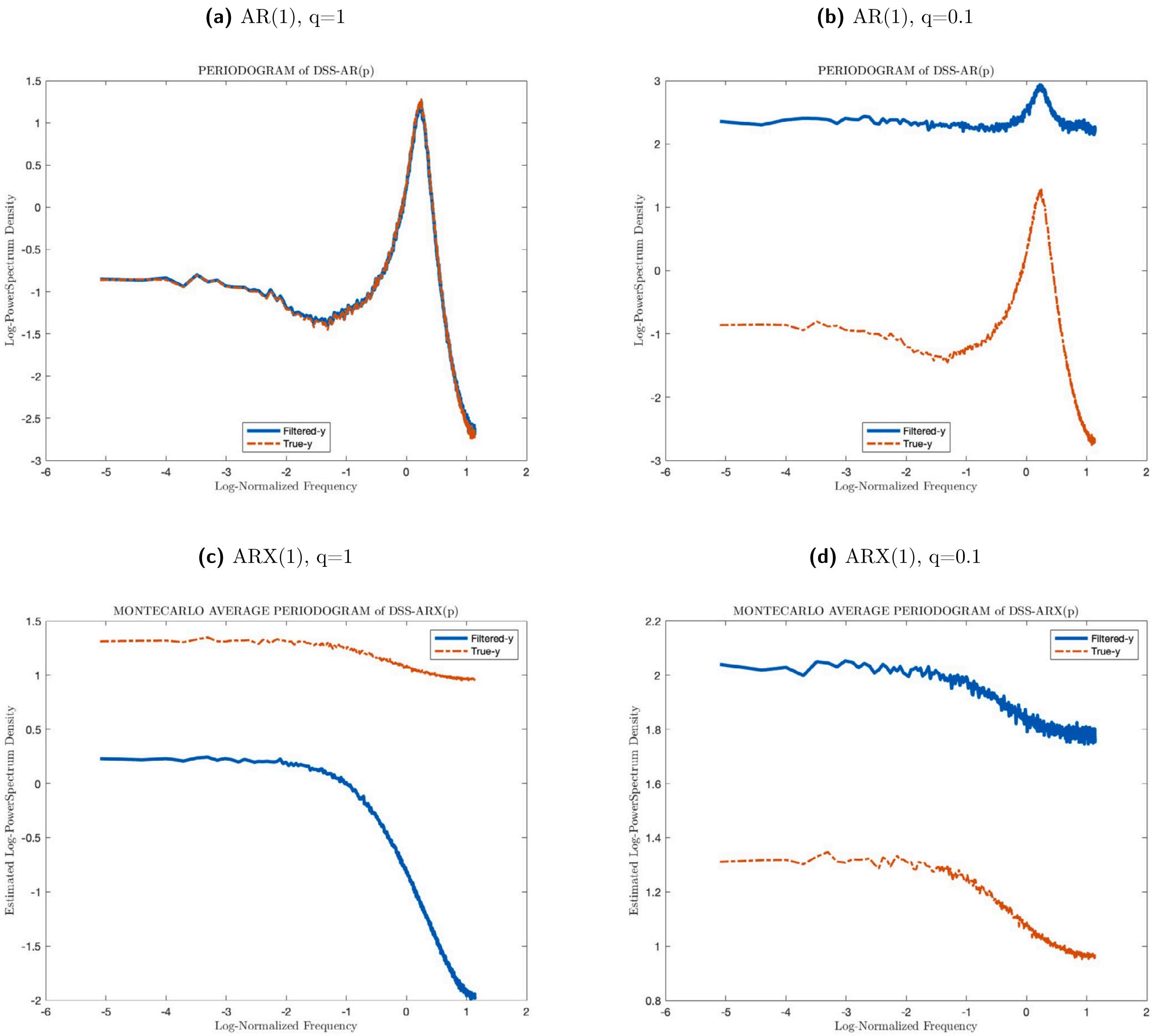
q_0	d_0	Mean $\hat{q}$	Mean $\hat{d}$				Empirical Power (× 100)			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.53	0.18	0.16	0.17	0.18	48.4	41.7	39.2	41.4
0.50	0	0.95	0.16	0.15	0.13	0.15	49.0	41.8	40.5	41.4
1.00		0.99	0.09	0.05	0.08	0.07	49.5	41.4	39.2	42.3
0.10		0.53	0.13	0.14	0.18	0.12	68.0	61.5	57.8	51.5
0.50	0.5	0.95	0.08	0.10	0.12	0.12	68.4	61.3	58.5	51.3
1.00		0.99	0.05	0.09	0.08	0.10	68.1	61.3	59.6	51.5
0.10		0.53	0.13	0.12	0.15	0.12	57.4	63.9	57.9	61.5
0.50	1	0.95	0.12	0.10	0.13	0.10	66.1	65.6	57.5	61.0
1.00		1.00	0.10	0.07	0.08	0.08	68.3	66.3	68.2	61.5
LMDJP-ARX(1) with $\hat{d}$, "LM First"										
0.10		0.65	0.15	0.15	0.13	0.16	39.5	38.5	40.9	45.7
0.50	0	0.98	0.12	0.14	0.15	0.14	42.1	45.2	49.5	45.8
1.00		1.00	0.10	0.12	0.11	0.09	56.3	57.2	51.3	55.8
0.10		0.65	0.17	0.14	0.15	0.16	37.8	41.3	57.8	52.3
0.50	0.5	0.98	0.12	0.11	0.12	0.11	50.4	51.9	52.4	53.0
1.00		1.00	0.10	0.07	0.07	0.09	52.4	53.9	52.4	54.0
0.10		0.65	0.12	0.14	0.13	0.13	68.5	68.4	68.8	64.4
0.50	1	0.98	0.10	0.10	0.10	0.10	68.3	68.6	64.8	62.2
1.00		1.00	0.05	0.07	0.05	0.18	69.2	69.5	67.6	63.0

NOTES: This table reports the MonteCarlo average estimates of the strategic judgmental parameter (column 3) and bias of the four estimators of long memory parameters (exposed in Appendix), jointly with their empirical powers (columns from 4 to the end) for different initial values of strategic judgment and memory. The upper part displays the results where $\hat{d}$ is estimated before applying Judgmental Filter, while the lower half reports the results for the case that $\hat{d}$ is estimated after that Judgmental Filter is run on data. The DGP is explained in Section 6.

According to Table 4, the average $\hat{q}$ is nearly zero for almost all series, and the estimated $\hat{d}$ varies significantly across time series and estimation strategies. Notably, nearly all series fail the test for no long-memory (LM) in the "LM-First" approach, whereas only a few do in the "Judgment-First" approach.

6.2. FED's survey data

The second case study examines the role of explanatory variables in a more general LMDJP with an AR-X(p) specification using survey data on Real GDP, Inflation, and Unemployment Rate for the U.S. economy from 1970:Q1 to 2020:Q4, provided by the



NOTE: This figure displays the average of a Monte Carlo simulated periodograms of the DJP-AR(1) and ARX(1) (in higher and lower panel, respectively) for $q_0 = 1$ (left panels) and $q_0 = 0.1$ (right panels) using the DGP illustrated in Section 6.

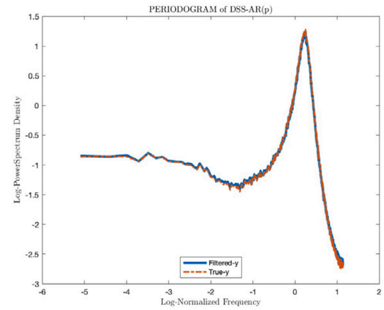
Fig. 6. MonteCarlo average of 1-order DJP-AR(X) Power spectra under different q_0 and assuming $d_0 = 0$.

Federal Reserve Bank of Philadelphia. We use all available measurements: real data, nowcasts, and forecasts up to one year ahead. The data source includes Reality (Real data) and FP outputs (nowcasts and forecasts up to one year ahead from the Survey of Professional Forecasters), while the Federal Open Market Committee (FOMC), which periodically publishes its projections, serves as FU. Explanatory variables include production, consumption, money velocity, and house prices—all available from FRED. Table 5 shows that the estimates and tests are qualitatively similar to those for Nelson-Plosser data.

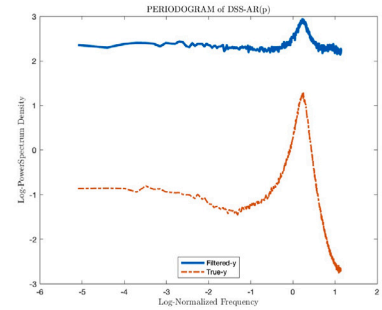
The erroneous over-rejection of the null hypothesis for the LM-First strategy is evident in two of the three case studies, while the Unemployment Rate case is largely immune to this issue. Two economic arguments can explain this observation. First, the judgmental protocol's design mechanism implies that while FP and FU interact, data are not free from judgmental bias. Instead, Reality is unaffected by who wins the game, meaning the winner is the agent whose utility gain does not increase significantly due to a large spread between FP forecasts and Reality outcomes—measured by $D(\cdot)$. This notion aligns with Vovk and Shafer (2005), who discuss modeling and testing such scenarios. Consequently, long-term strategic behavior based on judgment tends to exaggerate persistence, often mistaken for long-memory, though it results from recurrent games among agents.

The second argument relates to Bullard et al. (2008)'s concept of near-exuberance equilibria. These equilibria arise when (i) agents face rational expectations with limited information; (ii) each agent rationally includes judgment in their forecasting model, assuming others do the same; and (iii) the forecasting model remains stable and converges to rational expectations under recursive estimation. Our LMDJP framework meets these criteria.

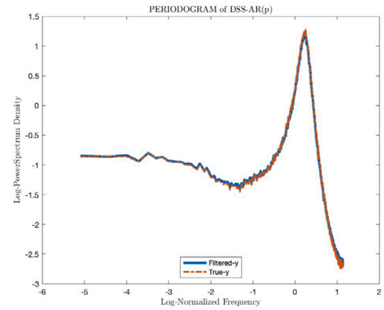
(a) AR(1), d=1, q=1



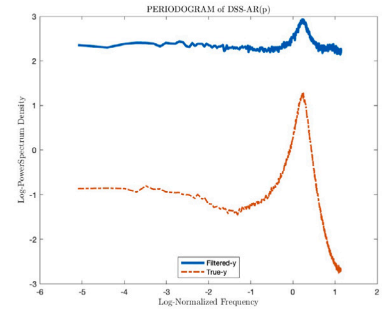
(b) AR(1), d=1, q=0.1



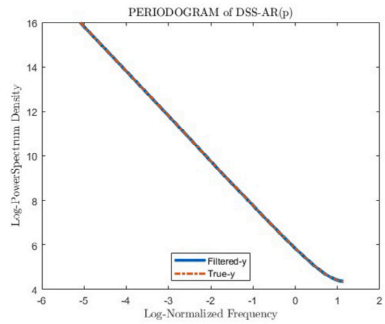
(c) AR(1), d=0.1, q=1



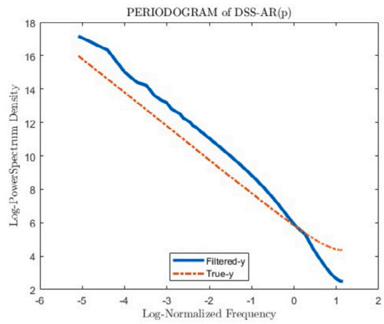
(d) AR(1), d=0.1, q=0.1



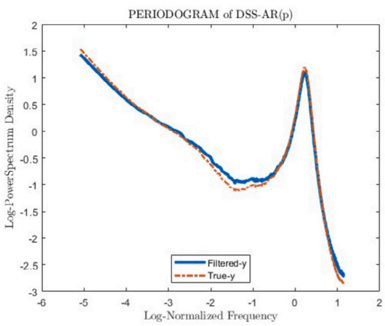
(e) ARX(1), d=1, q=1



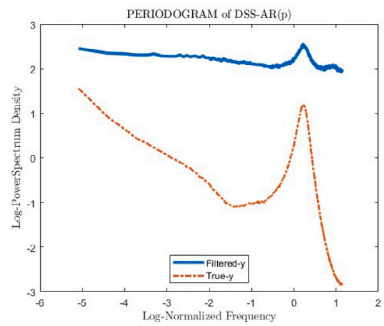
(f) ARX(1), d=1, q=0.1



(g) ARX(1), d=0.1, q=1



(h) ARX(1), d=0.1, q=0.1



NOTE: This figure displays the average of a Monte Carlo simulated periodograms of the LMDJP-AR(1) and LMDJP-ARX(1) several memory and judgmental parameters using the DGP illustrated in Section 6.

Fig. 7. MonteCarlo average of 1-order-DJP-AR(X) Power spectra under different q_0 and d_0.

Table 4
Application of the SULMDJP-AR model to Nelson-Plosser's data.

Series	T	$\hat{p}$	DJP-AR(p), "Judgment First"					DJP-AR(p), "LM First"				
			Mean $\hat{q}$	Mean $\hat{d}$				Mean $\hat{q}$	Mean $\hat{d}$			
				LW	ELW	ELW2	2SFELW		LW	ELW	ELW2	2SFELW
Real GNP	62	1	0.00	1.11	1.25	1.20	1.99	0.00	1.64	1.35	1.28	1.19
			(0.04)	(0.23)	(0.25)	(0.23)	(1.02)	(0.05)	(0.45)	(0.41)	(0.39)	(0.38)
Nominal GNP	62	1	0.00	1.19	1.68	1.70	1.70	0.00	1.00	1.78	1.81	1.70
			(0.05)	(0.14)	(0.35)	(0.25)	(0.41)	(0.06)	(0.16)	(0.21)	(0.23)	(0.22)
Real per capita GNP	62	1	0.00	0.84	0.94	0.92	0.92	0.00	0.64	1.03	0.99	0.92
			(0.04)	(0.19)	(0.09)	(0.09)	(0.58)	(0.05)	(0.12)	(0.17)	(0.15)	(0.14)
Industrial production	111	2	0.03	1.01	1.26	1.19	1.19	0.03	1.72	1.38	1.31	1.21
			(0.02)	(0.23)	(0.21)	(0.17)	(0.74)	(0.04)	(0.29)	(0.34)	(0.30)	(0.33)
Employment	81	2	0.00	0.57	0.66	0.66	0.67	0.00	0.67	0.76	0.77	0.77
			(0.04)	(0.22)	(0.21)	(0.18)	(0.52)	(0.06)	(0.08)	(0.12)	(0.12)	(0.09)
Unemployment rate	81	2	0.07	0.74	0.77	0.80	0.80	0.07	1.01	0.80	0.64	0.64
			(0.05)	(0.10)	(0.10)	(0.21)	(0.43)	(0.03)	(0.11)	(0.14)	(0.13)	(0.12)
GNP deflator	82	2	0.00	1.17	1.17	1.32	1.32	0.00	1.41	1.27	1.32	1.32
			(0.03)	(0.23)	(0.25)	(0.20)	(0.73)	(0.03)	(0.37)	(0.30)	(0.34)	(0.41)
CPI	111	2	0.00	0.97	1.32	1.24	1.23	0.00	1.27	1.42	1.33	1.23
			(0.03)	(0.23)	(0.21)	(0.24)	(0.71)	(0.06)	(0.14)	(0.39)	(0.40)	(0.23)
Nominal wage	71	2	0.00	1.27	1.38	1.54	1.54	0.00	1.68	1.48	1.64	1.54
			(0.02)	(0.26)	(0.27)	(0.27)	(0.98)	(0.05)	(0.41)	(0.45)	(0.39)	(0.42)
Real wage	71	2	0.00	0.98	1.02	1.04	1.04	0.00	1.50	1.12	1.11	1.04
			(0.03)	(0.07)	(0.11)	(0.24)	(0.53)	(0.04)	(0.23)	(0.24)	(0.28)	(0.27)
Money stock	71	2	0.00	1.20	1.71	1.78	1.78	0.00	1.46	1.80	1.92	1.78
			(0.04)	(0.22)	(0.25)	(0.41)	(1.01)	(0.05)	(0.28)	(0.31)	(0.44)	(0.45)
Velocity of money	102	2	0.00	0.99	1.02	1.06	1.06	0.00	0.07	1.12	1.12	1.06
			(0.04)	(0.13)	(0.18)	(0.14)	(0.46)	(0.05)	(0.05)	(0.20)	(0.18)	(0.20)
Bond yield	71	2	0.00	0.77	1.66	1.66	1.66	0.00	0.05	1.76	1.76	1.66
			(0.01)	(0.09)	(0.31)	(0.30)	(0.67)	(0.4)	(0.41)	(0.44)	(0.45)	(0.45)
Stock prices	100	2	0.03	1.15	1.22	1.12	1.12	0.03	1.47	1.35	1.25	1.43
			(0.02)	(0.17)	(0.20)	(0.21)	(0.38)	(0.03)	(0.28)	(0.29)	(0.18)	(0.28)
				Test for no Long Memory (p-values)								
Real GNP				0.058	0.067	0.064	0.012		0.065	0.078	0.069	0.069
Nominal GNP				0.063	0.093	0.094	0.012		0.062	0.099	0.100	0.094
Real per capita GNP				0.042	0.048	0.063	0.012		0.046	0.054	0.051	0.046
Industrial production				0.054	0.064	0.060	0.012		0.056	0.073	0.068	0.062
Employment				0.029	0.036	0.036	0.012		0.034	0.042	0.041	0.036
Unemployment rate				0.037	0.039	0.039	0.012		0.040	0.042	0.042	0.036
GNP deflator				0.061	0.064	0.070	0.012		0.065	0.068	0.071	0.076
CPI				0.048	0.071	0.063	0.012		0.051	0.077	0.072	0.066
Nominal wage				0.068	0.075	0.084	0.012		0.069	0.081	0.090	0.085
Real wage				0.050	0.052	0.054	0.012		0.056	0.059	0.058	0.054
Money stock				0.060	0.094	0.099	0.012		0.059	0.100	0.107	0.098
Velocity of money				0.050	0.052	0.055	0.012		0.056	0.059	0.059	0.055
Bond yield				0.038	0.092	0.092	0.012		0.041	0.097	0.092	0.098
Stock prices				0.061	0.064	0.058	0.012		0.061	0.072	0.066	0.060

NOTES: This table reports the result of the application of the LMDJP-AR model to Nelson-Plosser data on US economic variables. In the upper panel, the first three columns describe the variable names, the sample size and the estimated autoregressive orders; columns from fourth to eighth display the estimated q and the estimates of d for the "Judgment-First" modeling strategy for the four estimator here considered (and exposed in Appendix), while the equivalent estimates for "LM-First" are reported in the remaining columns. Standard errors are in parentheses below the parameter estimate. The lower panel displays the p-values of the test corresponding to (9), still for each of the four estimators. In both the modeling strategies we assume $q_0 = 0.1$.

The LMDJP is compared with the specification from Zanetti Chini (2023), which uses FRED survey data on y_t with four explanatory variables (industrial production, personal consumption, money velocity, and house prices)

The empirical comparison follows these steps:

(i*) We re-run the original DJP-AR(X) model from Zanetti Chini (2023), which corresponds to an AR(X)FIMA(2,0,0) model, with $d = 0$ indicating no long memory. Results are shown in the left column of Fig. 8.
(ii*) We then apply the d-operator to the same model, transforming it into an ARFIMA(2, d,0) model with $d = 0.5$. Results are shown in the right column of Fig. 8.
(iii*) We compare the estimated density of the LMDJP with an equivalent model estimated using traditional MLE (instead of ML_qE).
(iv*) We use the results to assess the forecasting properties of the LMDJP model.

The differences in dynamics and judgment are evident. The filtered time series and judgmental effects display increased noise and variations. Specifically, there is a noticeable inversion in the weighting function, with $w_t = 0$ (indicating no judgment) frequently

Table 5
Application of the ARX-LMDJP model to SPF data.

RGDP

Series	DJS-AR(p), "Judgment First"					DJS-AR(p), "LM First"			
	Mean $\hat{q}$	Mean $\hat{d}$				Mean $\hat{d}$			
		LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
Real data	. 0.87	0.81	0.83	0.86	0.86	1.20	0.05	0.05	0.45
	(0.09)	(0.21)	(0.17)	(0.22)	(0.22)	(0.34)	(0.02)	(0.01)	(0.08)
Nowcasts	0.08	0.17	0.21	0.21	0.21	0.22	0.22	0.12	0.22
	(0.02)	(0.04)	(0.06)	(0.06)	(0.07)	(0.07)	(0.02)	(0.07)	(0.05)
1-q-ahead	0.75	1.11	1.11	1.14	1.14	0.45	0.49	0.39	0.50
	(0.06)	(0.10)	(0.12)	(0.14)	(0.11)	(0.15)	(0.16)	(0.16)	(0.15)
2-q-ahead	0.72	1.12	1.16	1.16	1.16	0.45	0.53	0.54	0.43
	(0.08)	(0.14)	(0.12)	(0.10)	(0.11)	(0.11)	(0.19)	(0.18)	(0.17)
3-q-ahead	0.70	1.16	1.17	1.18	1.18	0.48	0.57	0.61	0.57
	(0.10)	(0.14)	(0.13)	(0.14)	(0.27)	(0.12)	(0.14)	(0.10)	(0.11)
1-y-ahead	0.57	1.22	1.04	1.08	1.09	0.29	0.58	0.75	0.70
	(0.05)	(0.13)	(0.14)	(0.20)	(0.13)	(0.12)	(0.13)	(0.13)	(0.11)
		Test for no Long Memory (p-values)							
Real data	.	0.04	0.04	0.04	0.04	0.00	0.00	0.00	0.00
Nowcasts		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
1-q-ahead		0.06	0.06	0.06	0.06	0.02	0.02	0.02	0.02
2-q-ahead		0.06	0.06	0.06	0.06	0.02	0.02	0.02	0.02
3-q-ahead		0.06	0.06	0.06	0.01	0.02	0.03	0.03	0.02
1-y-ahead		0.06	0.05	0.06	0.06	0.04	0.03	0.04	0.04

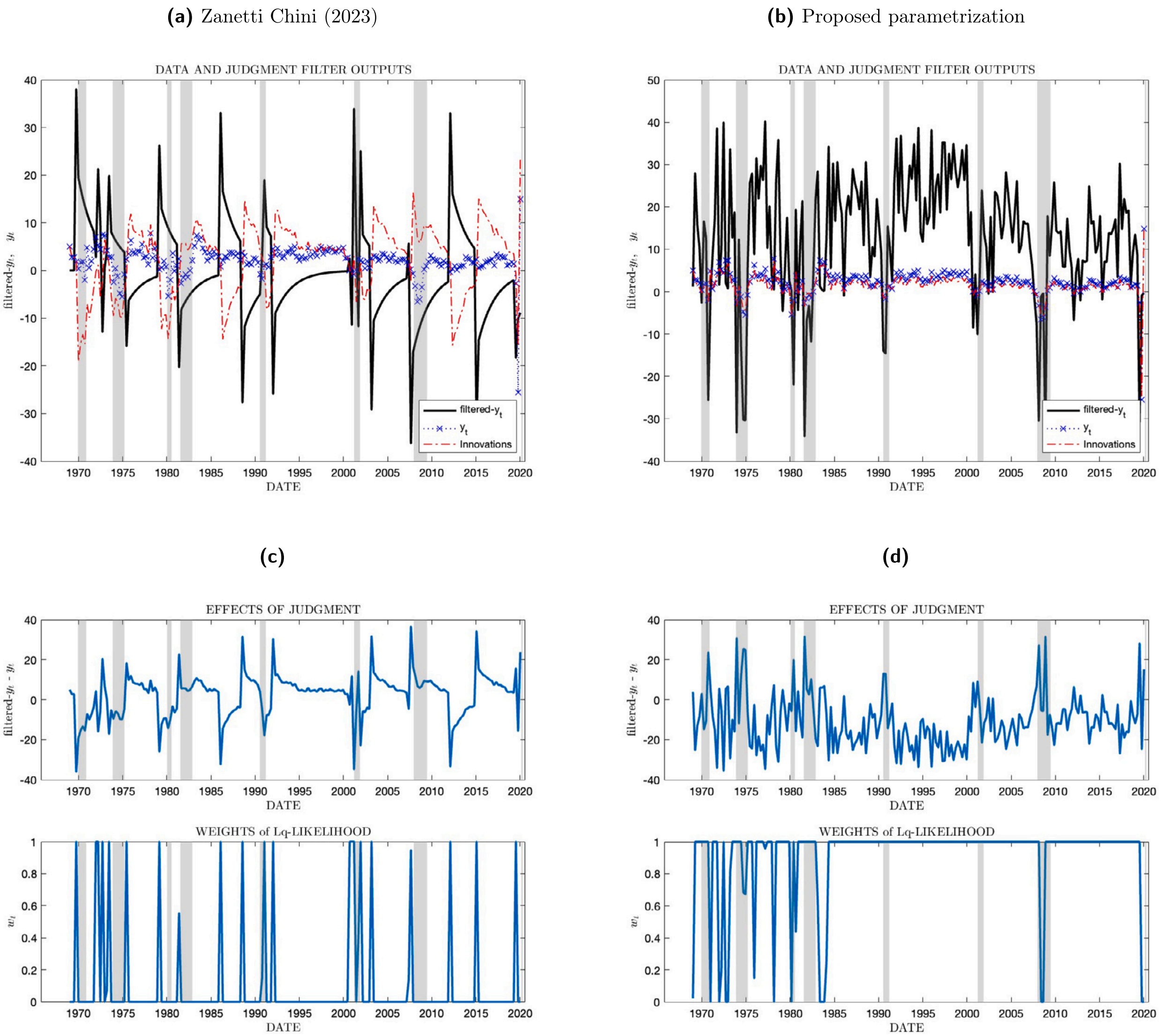
Inflation

Series	DJS-AR(p), "Judgment First"					DJS-AR(p), "LM First"			
	Mean $\hat{q}$	Mean $\hat{d}$				Mean $\hat{d}$			
		LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
Real data	0.91	0.85	0.89	0.91	0.92	1.08	0.69	0.69	0.65
	(0.08)	(0.08)	(0.07)	(0.09)	(0.09)	(0.10)	(0.00)	(0.10)	(0.09)
Nowcasts	0.12	0.86	0.89	0.96	0.95	1.04	0.38	0.46	0.45
	(0.02)	(0.10)	(0.09)	(0.11)	(0.12)	(0.14)	(0.21)	(0.23)	(0.22)
1-q-ahead	0.64	1.35	1.34	1.37	1.38	0.56	0.45	0.51	0.52
	(0.05)	(0.18)	(0.21)	(0.20)	(0.08)	(0.11)	(0.20)	(0.31)	(0.32)
2-q-ahead	0.58	1.26	1.26	1.30	1.29	0.97	0.67	0.67	0.65
	(0.07)	(0.30)	(0.26)	(0.24)	(0.26)	(0.10)	(0.14)	(0.13)	(0.11)
3-q-ahead	0.55	1.30	1.30	1.31	1.32	1.06	0.73	0.72	0.72
	(0.08)	(0.24)	(0.21)	(0.20)	(0.21)	(0.25)	(0.25)	(0.30)	(0.31)
1-y-ahead	0.49	1.38	1.39	1.39	1.40	0.67	0.70	0.72	0.69
	(0.07)	(0.31)	(0.24)	(0.23)	(0.24)	(0.20)	(0.25)	(0.28)	(0.22)
		Test for no Long Memory (p-values)							
Real data		0.04	0.05	0.05	0.04	0.00	0.01	0.01	0.01
Nowcasts		0.04	0.05	0.05	0.05	0.01	0.01	0.01	0.01
1-q-ahead		0.05	0.05	0.05	0.05	0.01	0.02	0.02	0.02
2-q-ahead		0.05	0.05	0.05	0.05	0.02	0.02	0.02	0.02
3-q-ahead		0.05	0.05	0.05	0.05	0.02	0.02	0.02	0.02
1-y-ahead		0.05	0.05	0.05	0.05	0.01	0.02	0.03	0.02

Unemployment

Series	DJS-AR(p), "Judgment First"					DJS-AR(p), "LM First"			
	Mean $\hat{q}$	Mean $\hat{d}$				Mean $\hat{d}$			
		LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
Real data	0.93	0.94	0.96	0.97	0.97	1.13	1.04	1.03	1.20
	(0.09)	(0.11)	(0.12)	(0.10)	(0.14)	(0.31)	(0.24)	(0.30)	(0.21)
Nowcasts	0.17	0.95	0.97	0.98	0.98	1.24	1.24	1.24	1.25
	(0.04)	(0.13)	(0.11)	(0.12)	(0.13)	(0.28)	(0.33)	(0.35)	(0.32)
1-q-ahead	0.73	1.05	1.08	1.08	1.08	1.02	1.02	1.09	1.01
	(0.09)	(0.10)	(0.13)	(0.16)	(0.13)	(0.35)	(0.15)	(0.16)	(0.12)
2-q-ahead	0.67	1.18	1.15	1.14	1.15	0.86	0.85	0.86	0.87
	(0.07)	(0.13)	(0.12)	(0.13)	(0.12)	(0.20)	(0.17)	(0.16)	(0.15)
3-q-ahead	0.52	1.23	1.22	1.24	1.18	0.90	0.94	0.90	0.83
	(0.06)	(0.13)	(0.20)	(0.14)	(0.09)	(0.25)	(0.21)	(0.20)	(0.16)
1-y-ahead	0.39	1.19	1.21	1.22	1.21	0.85	0.78	0.84	0.82
	(0.10)	(0.15)	(0.14)	(0.20)	(0.12)	(0.12)	(0.08)	(0.11)	(0.10)
		Test for no Long Memory (p-values)							
Real data		0.06	0.06	0.07	0.06	0.04	0.05	0.05	0.05
Nowcasts		0.07	0.07	0.07	0.07	0.05	0.05	0.05	0.05
1-q-ahead		0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.05
2-q-ahead		0.07	0.07	0.07	0.07	0.05	0.05	0.06	0.06
3-q-ahead		0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.06
1-y-ahead		0.07	0.07	0.07	0.07	0.06	0.06	0.65	0.06

NOTES: This table reports the result of the application of the LMDJP-ARX model to SPF data on US Real GDP, PCE Inflation and Unemployment rates for several forecasting horizons; see citezc:2023 for exact labeling and information on data informations. In the upper panel, the first three columns describe the variable names; columns from second to sixth display the estimated q and the estimates of d for the "Judgment-First" modeling strategy for the four estimator here considered (exposed in Appendix), while the equivalent estimates for "LM-First" are reported in the remaining columns. Standard errors are in parentheses below the parameter estimate. The lower panel displays the p-values of the test corresponding to (9), still for each of the four estimators. In both the modeling strategies we assume $q_0 = 0.1$ and an autoregressive order $p = 2$.



NOTE: This figure displays a comparison of LM-DJP-ARX(2) and DJP-ARX(2) using one-quarter-ahead forecasts of US-SPF. Left column corresponds to Zanetti Chini (2023) parametrization, while right column illustrates the behavior of the LMDJP proposed in this paper.

Fig. 8. Comparison of LM-DJP model with equivalent model by Zanetti Chini (2023) in U.S. RGDP forecasts.

observed in the bottom sub-panel (d) and $w_t = 1$ (indicating pure judgment) in the bottom sub-panel (c). Other forecasts not shown here confirm these findings, emphasizing that long memory can lead to an underestimation of judgment in economic forecasts.

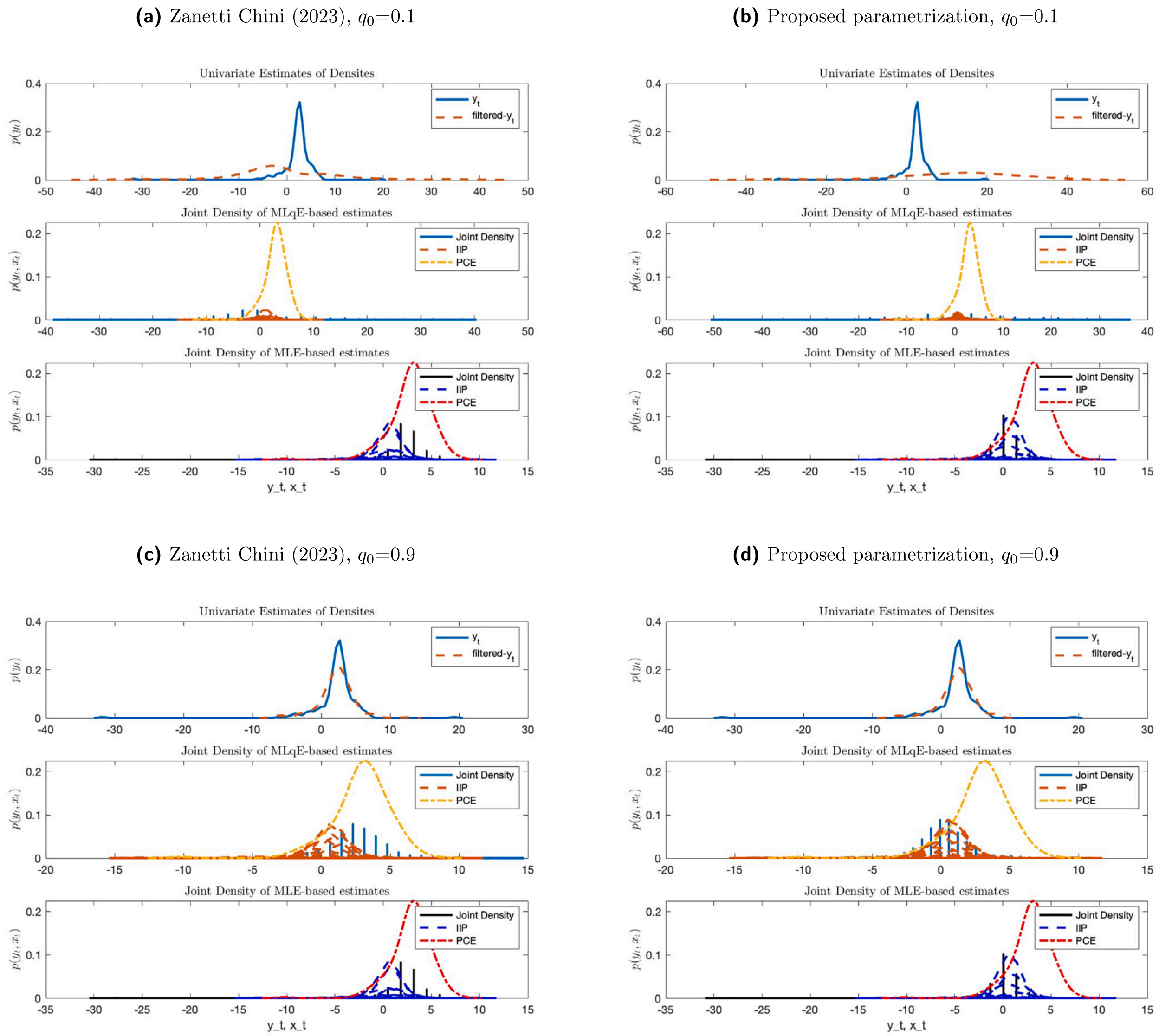
Fig. 9 compares the two models' density estimations using the example from Zanetti Chini (2023, Fig. 5). Although both models' estimated densities are similar, filtered estimates in cases of high initial judgment ($q_0 = 0.1$) tend to be more skewed in the LMDJP model.

In summary, while there are similarities in the estimated densities, the LMDJP model demonstrates a moderate improvement in forecasting capability over traditional ARFIMA models, as shown in Table 6. This indicates a trade-off between the out-of-sample performance of long-memory modeling and the in-sample estimation of judgment.

7. Conclusions

Long memory and judgment are systematically interrelated. Efficiently modeling this relationship can be achieved using the novel and flexible approach known as LMDJP. This method integrates various established econometric techniques—including spectral analysis, signal extraction, and time series modeling—re-evaluated through a robust estimation framework. Such an approach enables econometricians to replicate and build upon numerous novel findings in the macro-financial literature.

We consistently observe that environments with a significant degree of judgment tend to erroneously reject the null hypothesis of spurious long memory. This issue is particularly pronounced when the long-memory parameter is estimated in advance. Con-



NOTE: This figure displays a comparison of LMDJP-ARX(2) and DJP-ARX(2) using one-quarter-ahead forecasts of US-SPF. Left panel corresponds to Zanetti Chini (2023) parametrization, while right panel illustrates the behavior of the LMDJP proposed in this paper. Upper panels show that estimated model assuming low initial q (that is high judgment), while lower panels display the same model under high initial q (low judgment). Inside each panel, the upper sub-panel display univariate probability density functions of data and the DKF-estimates of system with one-year-ahed RGDP forecasts as dependent variables and IIP and PCE indexes as explanatory variables; in the central sub-panel, the univariate density functions of IIP and PCE as well as their joint density function by MLqE; and, in the lower sub-panel, the univariate density functions of IIP and PCE as well as their joint density function by MLE.

Fig. 9. Comparison of LM-DJP model with equivalent Zanetti Chini (2023) in density estimation using real data.

versely, including the long-memory parameter in judgment models often leads to an underestimation of the influence of non-sample information. Thus, we recommend identifying judgmental bias before assessing long-run dependence. The underlying microeconomic mechanisms responsible for these outcomes require further investigation. One possible explanation for this spuriousness could be the strategic application of judgment by economic agents.

Further research is needed to accurately detect, estimate, and evaluate the effects of agents' strategic behavior using empirical data. The structural treatment of judgment remains an open question, particularly regarding the causal links between economic variables and judgment, as well as the impact of shocks on judgment within dynamic systems.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jedc.2024.105005.

Table 6
In-sample and out-of-sample goodness-of-fit of LMDJS.

Model	Maximum Likelihood Estimation					
	Reality	Nowcasts	1-q-ahead	2-q-ahead	3-q-ahead	1-yr-ahead
RMSE	2.16	3.37	0.16	0.20	0.26	0.66
"In-sample RMSE"	1.22	1.13	0.08	0.12	0.13	0.34
"Out-of-sample RMSE"	1.23	1.70	0.06	0.11	0.12	0.33
	Maximum L_qLikelihood Estimation					
RMSE	1.90	2.76	0.15	0.21	0.25	0.64
"In-sample" RMSE	1.18	1.02	0.05	0.08	0.13	0.33
"Out-of-sample RMSE"	1.20	1.67	0.05	0.10	0.12	0.33

NOTE: This table reports the RMSE of the AR(2)-LMDJS estimated using standard Kalman Filter (upper part) and the DKF (lower part) for the case of the whole sample as well as for the case of pseudo-out-of-sample exercise with fixed window; in this last case, the estimation part ("in-sample") is 1970:Q2–1990:Q1, while the evaluation part ("out-of-sample") is 1990:Q2–2020:Q1.

References

Abadir, K.M., Distaso, W., Giraitis, L., 2007. Nonstationarity-extended local Whittle estimation. J. Econom. 141 (2), 1353–1384.
Afrouzi, H., Kwon, S.Y., Landier, A., Ma, Y., Thesmar, D., 2023. Overreaction in expectations: evidence and theory. Q. J. Econ. 138 (3), 1713–1764.
Arteche, J., 2020. Exact local Whittle estimation in long memory time series with multiple poles. Econom. Theory 36 (6), 1064–1098.
Asai, M., McAleer, M., Medeiros, M.C., 2012. Asymmetry and long memory in volatility modeling. J. Financ. Econom. 10 (3), 495–512.
Atkinson, A.C., Riani, M., Corbellini, A., et al., 2021. The Box-Cox transformation: review and extensions. Stat. Sci. 36 (2), 239–255.
Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. J. Econom. 73 (1), 5–59.
Baillie, R.T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. J. Econom. 74 (1), 3–30.
Banerjee, A., Urga, G., 2005. Modelling structural breaks, long memory and stock market volatility: an overview. J. Econom. 129 (1–2), 1–34.
Beran, J., Feng, Y., Ghosh, S., Kulik, R., 2013. Statistics for Long-Memory Processes: Probabilistic Properties and Statistical Methods. Springer.
Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2020. Overreaction in macroeconomic expectations. Am. Econ. Rev. 110 (9), 2748–2782.
Breiman, L., Friedman, J.H., 1985. Estimating optimal transformations for multiple regression and correlation. J. Am. Stat. Assoc. 80 (391), 580–598.
Brenna, F., Budrysb, Z., 2024. The term structure of judgement: interpreting survey disagreement.
Brockwell, P., Davis, R., 1991. Time Series: Theory and Methods. Springer Science & Business Media.
Bullard, J., Evans, G., Honkapohja, S., 2008. Monetary policy, judgment, and near-rational exuberance. Am. Econ. Rev. 98 (3), 1163–1177.
Chevillon, G., Mavroeidis, S., 2017. Learning can generate long memory. J. Econom. 198 (1), 1–9.
Christensen, B.J., Nielsen, M.Ø., 2007. The effect of long memory in volatility on stock market fluctuations. Rev. Econ. Stat. 89 (4), 684–700.
Cipriani, M., Guarino, A., 2014. Estimating a structural model of herd behavior in financial markets. Am. Econ. Rev. 104 (1), 224–251.
Clark, J., Friesen, L., 2009. Overconfidence in forecasts of own performance: an experimental study. Econ. J. 119 (534), 229–251.
Clements, M.P., 2018. Do macroforecasters herd? J. Money Credit Bank. 50 (2–3), 265–292.
Coibion, O., Gorodnichenko, Y., 2012. What can survey forecasts tell us about information rigidities? J. Polit. Econ. 120 (1), 116–159.
Coibion, O., Gorodnichenko, Y., 2015. Information rigidity and the expectations formation process: a simple framework and new facts. Am. Econ. Rev. 112 (7), 2644–2678.
Corsi, F., 2009. A simple approximate long-memory model of realized volatility. J. Financ. Econom. 7 (2), 174–196.
Dahlhaus, R., 2006. Efficient parameter estimation for self-similar processes. Ann. Stat. 34 (2), 1045–1047.
Davidson, J., Sibbertsen, P., 2005. Generating schemes for long memory processes: regimes, aggregation and linearity. J. Econom. 128 (2), 253–282.
DeGroot, M.H., Mortera, J., 1991. Optimal linear opinion pools. Manag. Sci. 37 (5), 546–558.
DellaVigna, S., 2009. Psychology and economics: evidence from the field. J. Econ. Lit. 47 (2), 315–372.
Diebold, F.X., Inoue, A., 2001. Long memory and regime switching. J. Econom. 105 (1), 131–159.
Diebold, F.X., Rudebusch, G.D., 1989. Long memory and persistence in aggregate output. J. Monet. Econ. 24 (2), 189–209.
Ehrbeck, T., Waldmann, R., 1996. Why are professional forecasters biased? Agency versus behavioral explanations. Q. J. Econ. 111 (1), 21–40.
Ergemen, Y.E., 2023. Parametric estimation of long memory in factor models. J. Econom.
Ferrari, D., Yang, Y., 2010. Maximum Lq-likelihood. Ann. Stat. 38 (2), 753–783.
Gallo, G.M., Granger, C.W., Jeon, Y., 2002. Copycats and common swings: the impact of the use of forecasts in information sets. IMF Staff Pap. 49 (1), 4–21.
Garratt, A., Henckel, T., Vahey, S.P., 2023. Empirically-transformed linear opinion pools. Int. J. Forecast. 39 (2), 736–753.
Gelain, P., Manganelli, S., 2020. Monetary policy with judgment.
Genest, C., Schervish, M.J., 1985. Modeling expert judgments for Bayesian updating. Ann. Stat., 1198–1212.
Geweke, J., Amisano, G., 2011. Optimal prediction pools. J. Econom. 164 (1), 130–141.
Goodwin, P., Fildes, R., 2022. Forecasting with judgment. In: The Palgrave Handbook of Operations Research. Springer, pp. 541–572.
Granger, C.W., 1966. The typical spectral shape of an economic variable. Econometrica, 150–161.
Granger, C.W., 1980. Long memory relationships and the aggregation of dynamic models. J. Econom. 14 (2), 227–238.
Granger, C.W., Ding, Z., 1996. Varieties of long memory models. J. Econom. 73 (1), 61–77.
Grassi, S., Santucci de Magistris, P., 2014. When long memory meets the Kalman filter: a comparative study. Comput. Stat. Data Anal. 76, 301–319.
Hall, S.G., Mitchell, J., 2007. Combining density forecasts. Int. J. Forecast. 23 (1), 1–13.
Hassler, U., 2018. Time Series Analysis with Long Memory in View. John Wiley & Sons.
Hosking, J., 1981. Fractional differencing. Biometrika 68 (1), 165–176.
Hurst, H.E., 1951. Long-term storage capacity of reservoirs. Trans. Am. Soc. Civ. Eng. 116 (1), 770–799.
Hurvich, C.M., Moulines, E., Soulier, P., 2005. Estimating long memory in volatility. Econometrica 73 (4), 1283–1328.
Ilut, C., Valchev, R., 2023. Economic agents as imperfect problem solvers. Q. J. Econ. 138 (1), 313–362.
Kapetanios, G., Mitchell, J., Price, S., Fawcett, N., 2015. Generalised density forecast combinations. J. Econom. 188 (1), 150–165.
Lamont, O.A., 2002. Macroeconomic forecasts and microeconomic forecasters. J. Econ. Behav. Organ. 48 (3), 265–280.
Lawrence, M., Goodwin, P., O'Connor, M., Önkal, D., 2006. Judgmental forecasting: a review of progress over the last 25 years. Int. J. Forecast. 22 (3), 493–518.
Li, D., Ling, S., Tong, H., 2012. On moving-average models with feedback. Bernoulli, 735–745.

Mandelbrot, B.B., Van Ness, J.W., 1968. Fractional Brownian motions, fractional noises and applications. SIAM Rev. 10 (4), 422–437.
Manganelli, S., 2009. Forecasting with judgment. J. Bus. Econ. Stat. 27 (4), 553–563.
Mankiw, N., Reis, R., 2002. Sticky information versus sticky prices: a proposal to replace the new Keynesian Phillips curve. Q. J. Econ. 117 (4), 1295–1328.
Monti, F., 2010. Combining judgment and models. J. Money Credit Bank. 42 (8), 1641–1662.
Nelson, C.R., Plosser, C.R., 1982. Trends and random walks in macroeconmic time series: some evidence and implications. J. Monet. Econ. 10 (2), 139–162.
Olszewski, W., 2015. Calibration and expert testing. In: Young, H., Zamir, S. (Eds.), Handbook of Game Theory with Economic Applications. North Holland.
Phillips, J.K., Klein, G., Sieck, W.R., 2004. Expertise in judgment and decision making: a case for training intuitive decision skills. In: Blackwell Handbook of Judgment and Decision Making, vol. 297, p. 315.
Robinson, P., 1995. Gaussian semiparametric estimation of long range dependence. Ann. Stat., 1630–1661.
Robinson, P.M., 1994. Efficient tests of nonstationary hypotheses. J. Am. Stat. Assoc., 1420–1437.
Robinson, P.M., Hidalgo, F., 1997. Time series regression with long-range dependence. Ann. Stat. 25 (1), 77–104.
Schennach, S.M., 2018. Long memory via networking. Econometrica 86 (6), 2221–2248.
Shimotsu, K., 2010. Exact local Whittle estimation of fractional integration with unknown mean and time trend. Econom. Theory 26 (2), 501–540.
Shimotsu, K., Phillips, P.C., 2005. Exact local Whittle estimation of fractional integration. Ann. Stat. 33 (4), 1890–1933.
Shimotsu, K., Phillips, P.C., 2006. Local Whittle estimation of fractional integration and some of its variants. J. Econom. 130 (2), 209–233.
Sims, C.A., 2003. Implications of rational inattention. J. Monet. Econ. 50 (3), 665–690.
Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. J. Econom. 53 (1–3), 165–188.
Svensson, L.E., 2005. Monetary policy with judgment: forecast targeting. Int. J. Cent. Bank. 1 (1), 1–54.
Tibshirani, R., 1988. Estimating transformations for regression via additivity and variance stabilization. J. Am. Stat. Assoc. 83 (402), 394–405.
Vovk, V., Shafer, G., 2005. Good randomized sequential probability forecasting is always possible. J. R. Stat. Soc., Ser. B, Stat. Methodol. 67 (5), 747–763.
Zaffaroni, P., 2004. Contemporaneous aggregation of linear dynamic models in large economies. J. Econom. 120 (1), 75–102.
Zanetti Chini, E., 2023. Can we estimate macroforecasters' misbehavior? J. Econ. Dyn. Control 149, 104632.