



Can we estimate macroforecasters' mis-behavior?

Emilio Zanetti Chini¹

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



ARTICLE INFO

Article history:

Received 20 August 2022

Revised 21 December 2022

Accepted 3 March 2023

Available online 10 March 2023

JEL classification:

C1

C2

C5

E3

E7

Keywords:

Deformed likelihood

Dynamic systems

Judgment

Repeated games

Robust filtering

ABSTRACT

We answer positively to this question by using Maximum Lq-Likelihood (or Deformed Likelihood) estimator. This is based on a parameter which measures the aggregate quote of judgment in the forecasting (game-based) system formed by three players—Forecaster, Policy Maker and Reality. For the first time in econometric literature, we apply this estimator to a dynamic system and derive a robust version of the Kalman Filter—the Deformed Kalman Filter (DKF). The evidence from U.S. data suggests that the judgmental dynamics exists and is correlated (but not coincident) with the phases of the Business Cycle. Furthermore its knowledge improves in-sample as well as out-of-sample estimation.

© 2023 The Author. Published by Elsevier B.V.

This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>)

"It's easier to invent than to observe" (Gioacchino Rossini)

"If you torture data long enough, it will confess to anything" (Ronald H. Coase).

1. Introduction

The difficulty in observing and correctly interpreting a phenomenon is a non-trivial aspect of any scientific knowledge. When the Italian composer formulated his statement, these capabilities—and their understanding—were a privilege for a restricted élite. Thus, despite its humor, the statement should be considered as a (partial) truth. The growing pervasiveness of information technology and the rapid acceleration and development of statistical tools for data processing since the mid 20th Century made the incentive for exploiting the capability of data science superior to the incentive to think of new approaches to Economics, for the first time in Social Science history. Thus the Nobel Laureate's statement might be considered less paroxysmal than suggested by a superficial interpretation.

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

¹ The Author is particularly grateful to Giovanni Urga and Salvatore Piccolo for their support. He also thanks Barbara Annicchiarico, Marcelle Chauvet, Raffaella Giacomini, Ana Maria Herrera, Alessandra Luati, Giuseppe Ragusa, Francesco Ravazzolo and Tom Stark for their feedback, as all the participants to the 2nd Ventotene Workshop in Macroeconomics, the 41st International Symposium on Forecasting 2021, the internal Seminar Series in Mathematics and Statistics of the University of Bergamo, the IWEEE2022 in Rimini and the IAAE2022 in London. In some of that occasion the paper had been presented with the title "Measuring Unobserved Strategic Judgment". The co-editor, Juan Francisco Rubio-Ramirez, and a referee are grateful for their contributions to the paper's readability and scientific correctness. Finally, the Author is in debt with doctors, nurses and workers of Policlinico "S. Matteo" of Pavia, without whose uncompensated cares this paper would not have been written.

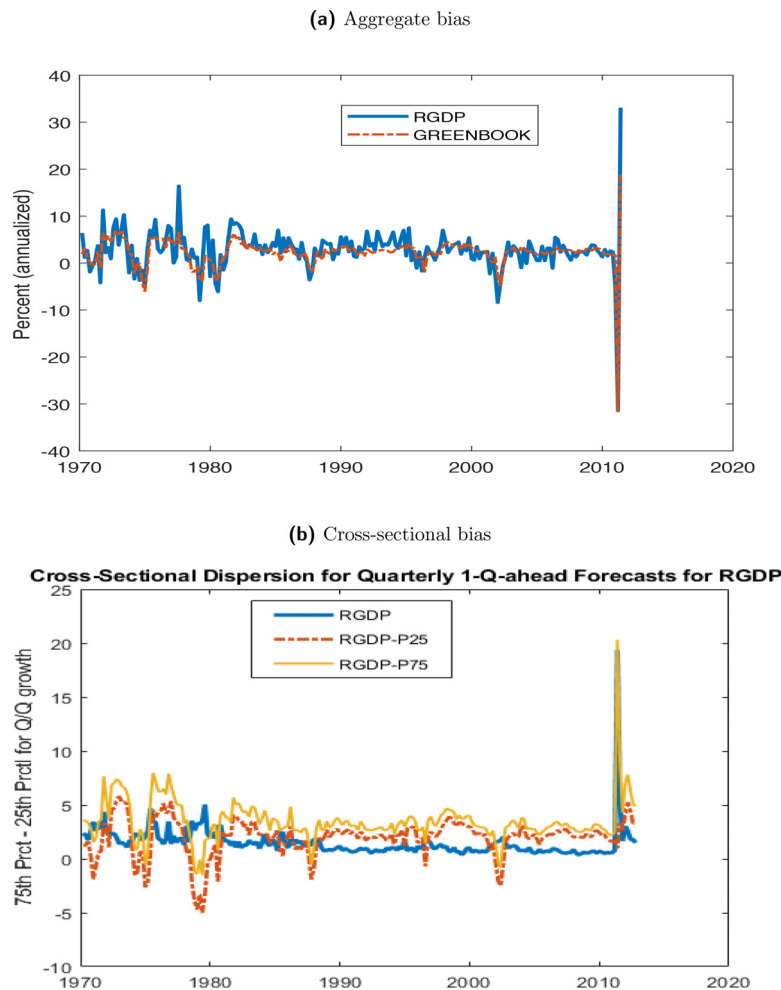


Fig. 1. The FED-SPF forecasting bias. *NOTE:* This figure plots the evidence of time-varying forecasting bias (top) and cross-sectional dispersion among professional forecasters (bottom). Namely, panel (a) considers the bias defined as $y_t - \hat{y}_{t+1|t}$, where y_t are realizations and $\hat{y}_{t+1|t}$ the Greenbook forecasts; panel (b) displays the $y_t - \hat{y}_{t+1|t}^{(i)}$, for $i = 1, \dots, N$, $t = 1 \dots, T$ with t denoting time period and i denoting the i -esim individual respondent to the survey for the 75th and 25th percentiles of the distribution of survey respondents. The sample spans from 1969Q4 to 2020Q1. All panels concern 1-quarter-ahead forecasts. The white space among 2012 and 2020 in top panels is a Matlab bug that compresses time dimension of the lines no matter of the dating labels in X-axis.

The bias in the survey of the U.S. economy's professional forecasters is an example of mismatching among observation and interpretation in contemporary Macroeconomics. Figure 1 illustrates the case of 1-quarter-ahead forecasts of the U.S. Real GDP (RGDP, henceforth): namely, panel (a) considers the bias across time—the spread among realizations and Greenbook estimates²—while panel (b) investigates the same bias in a cross-sectional perspective—that is, the spread among the realizations and the 1-quarter ahead estimates for individuals corresponding to the 75th and 25th percentiles of the distribution of survey respondents (denoted RGDP-P75 and RGDP-P25, respectively). The spread among data and Greenbook forecasts is pervasive both in time and in cross-sectional dimension, even if considering the difference among the 75th and 25th percentile' survey respondents as a centered measure to make our investigation robust to extremely optimistic/pessimistic forecasters.

The existence of the experts' forecasting bias may not constitute a challenge if an optimal prediction would convey estimates 'near enough' to consensus forecasts. In fact, the problem of the optimal prediction in a dynamic regression framework is well-known since the introduction of Kalman (1960)'s recursive algorithm known as 'the Kalman filter' (KF, henceforth). Nevertheless, the evidence displayed in Fig. 2 also contradicts this hypothesis: the Greenbook forecasts are structurally over the optimum and below the realizations (in absolute values). In other words, the optimal forecasts lie in a very strict band around zero, below (above) both the data for positive (negative) values and consensus estimate. Moreover, the

² Greenbook estimates is a label for what is commonly known with "consensus" forecasts.

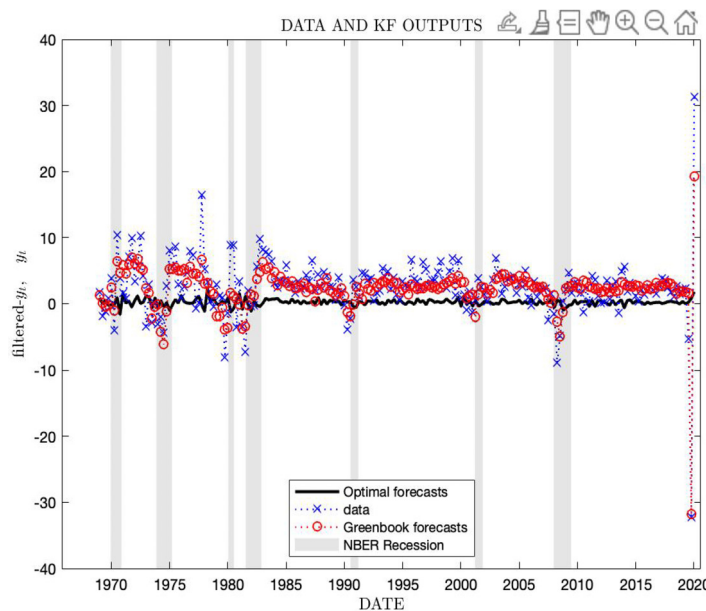


Fig. 2. A comparison of optimal dynamic forecasts, Greenbook forecasts and data. *NOTE:* This figure compares the Greenbook 1-quarter-ahead forecasts (red) of RGDP for U.S. with optimal predictions based on the application of the Kalman filter (black) on the data (in blue). Grey areas are the NBER Recession dates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

existence of large outliers is ignored and, in some cases, poorly addressed: for example, the 2020 Pandemic shock in RGDP growth is interpreted as a (dramatic) increase from 0 to 5% approximately, while the corresponding observation lies at a -30% level. Thus, the U.S. macroforecasters were effectively more accurate than an agnostic—and uninformed—econometrician.

One may be tempted to cover this lack of standard signal extraction methods by complicating the model assumed as a data generating process around which optimal forecasting exercise is applied. For example, the data may be assumed to follow a nonlinear and/or time-varying parameter evolution in several of its moments; or to be driven by some unobserved factors to be estimated using large datasets; or a combination of these features. The econometric literature is considerable and in a mature state, see [Durbin and Koopman \(2012\)](#) for an exhaustive introduction. However, using more complex econometric modelling, while improving our forecasting exercise, would *not* explain why U.S. professional forecasters diverge from realization of the RGDP measurements. Explaining this bias and using this information to improve our forecasting capability is the aim of this paper.

According to the literature, economic agents (and thus, professional forecasters) face some sort of limitation in their data processing capabilities and other types of costs arising from interaction with other agents (for example, reputation). These (unobserved) costs, in turn, may easily influence the agents' utility maximization and their expectations—and consequently, their forecasts based on standard, rational expectation-based econometric models. In particular, limited attention and imperfect utility maximization justify the existence of outliers—that considerably complicate the problem of signal extraction—in forecasting. In this regard, a number of (apparent) expectation-formation biases have been documented, see [Coibion and Gorodnichenko \(2012, 2015\)](#) and [Bordalo et al. \(2020\)](#) *inter alia*.

The identification of the ultimate source of these biases is an open issue. Its exploration requires (i) a theory that explains the genesis of the macroforecasters' misbehavior and its effect on the business cycle; (ii) an estimation method that allows economists to capture the dynamics of such misbehavior. While there is a growing literature on the former issue, the latter is still at an early stage. In other words, we know that economic agents face with behavioral aspects that justify the forecasting bias but there is still not a recognized measure of aggregate distortion caused by agents' behavior. In this paper we consider macroforecaster's mis-behavior an epiphenomenon of the *judgment*, defined by [Svensson \(2005, p. 2\)](#) as “information, knowledge, and views outside the scope of a particular model” and represented, by the same Author, as a Central Bank's estimate of arbitrary stochastic factor added to the econometric model equations.

Hence our research question: *How to elicit judgment directly in phase of estimation?* Our contribution is represented by the definition of a novel robust methodology to extrapolate judgment empirically. Namely, we propose the dynamically judgmental system (DJS) as a general characterization of the judgmental dynamics via state-space modelling and introduce a direct estimation method of the amount of judgment via signal extraction techniques. In our framework the (aggregate) amount of judgment in the system is parametrized by a functional of the likelihood—named “*Lq-Likelihood*”, or “*Deformed Likelihood*”—of the model to be estimated. The Deformed Likelihood estimator is characterized by a parameter that defines the degree of deformation of the logarithmic transform of the likelihood function due to the presence of additive outliers, hence providing a direct, robust measure of the degree of deformation due to judgment in the whole forecasting system.

The DJS allow us to estimate the dynamics of this deformation via a peculiar version of KF, named “Deformed Kalman Filter” (DKF).

Our simulation exercise demonstrates that the DJS-DKF has good general properties in terms of accuracy in small samples and that the distribution of the deformation parameter is very well approximated by a Gaussian distribution. Finally, we apply our methodology to the Survey of Professional Forecasters of the Federal Reserve Bank (FED-SPF) with focus on the forecast of RGDP. The empirical evidence leads us to the conclusion that the judgmental dynamics varies considerably according to the ownership of the dataset and is not perfectly coincident with recession dating. These findings open new perspectives on beliefs formation in macroeconomic forecasting and related economic theory.

Section 2 allocates our contribution in the scientific debate; Section 3 describes the theoretical issues to set-up the DJS; Section 4 illustrates the robust statistical methodology to be adopted in the new framework; the results of the application on real U.S. data are illustrated in Sections 5; 6 concludes; finally, an Appendix provides mathematical proofs, while a separate Supplement reports the results of some Monte Carlo simulations and additional results.

2. Literature

This paper contributes to both the economic and statistical literature. Concerning the economic contribution, it relates to the increasing strand of macroeconomic literature (Mankiw and Reis, 2002; Sims, 2003; Woodford, 2003) that relaxes the ‘Rational Expectation Hypothesis’ (REH). It also contributes to the econometric literature that relates the measurement of uncertainty and rationality testing to professional forecasting; see Jo and Sekkel (2019) and therein mentioned literature. In particular, we are inspired by two ideas originally formalized by Townsend (1978, 1983): (i) macroeconomic forecasting is a complex activity that requires to include suitably collected experts’ opinions in the classical econometric modelling of general equilibrium systems; (ii) as a consequence, it acts as a sort of clearing-market condition that, under standard REH corresponds to a Bayes-Nash-equilibrium. As a logical consequence, deviations from REH should correspond to deviations from the Bayesian updating of beliefs. This is (partially) documented by a recent strand of literature; see, *inter alia*, Manzan (2011), Ortaleva (2012), Manzan (2021), Giacomini et al. (2020), Gaglianone et al. (2022). Moreover, this paper is near to the field that reconsiders the literature in Business Cycle (and related research fields) in light of very well established results in behavioral sciences; see De Long et al. (1989), Ehrbeck and Waldmann (1996) and Ottaviani and Sørensen (2001) *inter alia*.

Secondly, we nest the literature on professional forecasters’ disagreement and their learning mechanisms that complicate the definition of suitable measures of uncertainty.³ The judgmental dynamics’ effective measurement in survey data and its econometric treatment are still open issues: the literature that focuses on judgment is represented only by Manganelli (2009) and Kocięcki et al. (2012) in non-Bayesian and Bayesian frameworks, respectively, while the only reference focused on theory-based modelling is Monti (2010).

This is our methodological contribution: we provide a novel framework that allows judgment in forecasting and includes the *prequential approach* to forecast evaluation introduced by Dawid (1984)⁴ and the literature on signal extraction and robust filtering: namely, we use the approach based on influence function by Masreliez and Martin (1977) to estimate the parameter that controls the degree of deformation of the likelihood function for any observation in the sample. The static version of this estimator has been introduced by Ferrari and Yang (2010). In the case of no judgment, our DKF coincides with the robust Kalman filter (RKF) by Marczak et al. (2018) and is similar, in the economic meaning, to the one derived by Bordalo et al. (2019).⁵ Finally, this paper can be viewed as a methodological counterpart of the generalized noisy information theory by Broer and Kohlhas (2022), and as a simpler alternative to the Bianchi et al. (2022)’s machine-learning technique recently introduced in the literature.

3. Theoretical framework

This section outlines the theory of DJS. Namely, Section 3.1 describes the formation process of judgment in a prototypical forecasting environment; Section 3.2 introduces the notation, while the main definitions for the set-up of the new tool are exposed in Section 3.3.

3.1. The judgmental forecasting protocol

Definition 1. (Judgmental Protocol) We assume that the forecasting bias due to judgment is the output of a repeated game (named *Judgmental Protocol*, henceforth) among three players: Policy Maker and Forecaster (henceforth PM and F, respectively) and Reality. Then, for $t = 1 \dots T$,

³ See Boero et al. (2008), Capistrán and Timmermann (2009), Patton and Timmermann (2010), Lahiri and Sheng (2010), Dovern et al. (2012), Andrade and Le Bihan (2013), Clements (2014), Rossi and Sekhposyan (2015), Andrade et al. (2016), Abel et al. (2016).

⁴ Namely, our DKF is the empirical counterpart that has never developed to our best knowledge despite the Author’s suggestions; see Dawid (1984), p. 289.

⁵ The DKF and the one by Bordalo et al. (2019), however, do not coincide from a statistical point of view, as will be clarified in the course of the treatment. Concerning the economic foundation, since the aim of this paper is purely empirical we do not assume any particular behavioral model. The occurrence of a judgmental action is only postulated when the amount of judgment is found statistically significant.

1. PM and F individually collect, process and analyze data, consisting of a set of explanatory variables, to make out-of-sample inference on a target variable; at the same time both F and PM form their own utility functions;
2. F produces (potentially biased) projections of the objective variable under alternative assumptions about the explanatory variables and/or under different scenarios;
3. PM receives the projection sets, adds judgment in each of them and, on the base of his/her objectives and information, decides which projection looks better;
4. PM announces next period's current official forecast;
5. Reality draws the realization.

Remark 1. The Judgmental Protocol above stated is a simplified version of the policy decision process illustrated by Sims (2002) and Svensson (2005). In particular, our simplifying assumptions are the following: (i) PM and F are single, homogeneous players. Thus, F may well represent both external experts and Central Bank's staff and no role is given to the Central Bank staff during the final decision process. (ii) There is no feedback effect in PM's decision—that is, Reality is not influenced by PM when the process repeats in a subsequent period, nor F replies to PM in the course of the process. In turn, this holds because we do not assume the agents' strategic behavior. (iii) Reality is neutral with respect to the decisions taken by PM and F. While this may seem counterintuitive in a purely macroeconomic policy perspective, in a microeconomic perspective it is a necessary condition to isolate the role of F and MP in the production of q time by time.

Step 1 explicitly allows the existence of the utility function and its possible divergence among PM and F. This is necessary to allow an economic interpretation of the macroforecasters' misbehavior—since it can naturally emerge directly from asymmetry in utility functions without adding any judgmental component, see Patton and Timmermann (2007) *inter alia*. The existence of (potentially different utility) functions among the two Players is definitional for setting the statistical tool that we are going to discuss in the course of this Section. Noticeably, Reality *does not* have a utility function.

Step 2 explicits two key characteristics of the judgmental process: (i) F is independent from PM in producing the projections; (ii) *projection* and *forecasting* are different objects: the former is a product of a single Player, the latter a product of a (simplistic) interaction among two Players. In Step 3 this interaction takes the form of a decision by the other Player.

Step 4 is where the judgment becomes empirically measurable. The judgmental bias is incorporated in forecasts when PM makes the decision on the best projection, which is not of public domain by definition. However, this step allows to test for the effectiveness of judgmental components only, not for the optimality of the whole forecast. The optimality testing becomes feasible only in Step 5. At this point, it is possible to compare all the phases of the forecasting process.

3.2. Notation and set-up

We are interested in the stochastic process $[Y, X, \xi] \doteq \{[Y_t, X_t, Z_t] : \Omega \rightarrow \mathbb{R}^{k+1}, k \in \mathbb{N}, t = 1, \dots, T\}$ where $Y_t = \{y_1, \dots, y_T\}$ is the vector of observed data, and $X_t = \{x_1, \dots, x_T\}'$ is a vector of explanatory variables, ξ is the vector of judgmental values added to the system and the set of all possible values taken by $[Y, X, \xi]$ is $[\mathcal{Y}, \mathcal{X}, \Xi]$. Moreover, it is defined on a complete probability space $\{\Omega, \mathcal{F}, \mathcal{P}\}$, where Ω is the sample space and $\mathcal{F}_t$ the event space; $\mathcal{P} \doteq \{\mathbf{p} \in \mathcal{A} : \sum_x p_x = 1\}$ defines the set of all distributions on $[\mathcal{Y}, \mathcal{X}]$ that are absolutely continuous respect to a σ -finite measure μ , $\mathcal{A}$ an algebraic subset of $[Y, X, \xi]$ representing the set of F's actions, in turn denoted as $q \in \mathbb{R}$.

The Log-likelihood function of $\hat{y}_t$ is denoted as $\mathcal{L}(\Theta, q)$; the equivalent for PM and F are $\mathcal{L}(\Theta^{PM}, q)$ and $\mathcal{L}(\Theta^F, q)$, respectively; the density of Y is denoted $P(Y) \doteq \int p_Y(y) dy$, where $p(\cdot)$ is a continuous density function defined on $\mathcal{L}(\Omega)$ and t is omitted to ease the notation; the (one-step-ahead) distributional and density forecasts of Y_t are denoted as $P(Y_{t+1})$ and $p(Y_{t+1})$, respectively. The utility function corresponding to the true assessment of event Y in $t + 1$ is denoted as $U \in \mathbb{R}$ and labeled as Scoring Rule (SR). Let $\overline{\mathbb{R}} = [-\infty, +\infty]$ denote the extended real line and the functions $H(Y, X) : \mathcal{P} \rightarrow \overline{\mathbb{R}}$ and $D(X, Y) : \mathcal{P} \times \mathcal{P} \rightarrow \overline{\mathbb{R}}$ be associated with any $U(Y, X)$.

At time t , we denote $\mathcal{H}(X) = U(X, X) \equiv \sup_{X \in \mathcal{X}} U(Y, X)$ the maximum utility (or Entropy function) that a forecaster gains when X (truly) realizes and $D(X, Y) = H(X) - S(X, Y)$ the Divergence function between the predictive densities of X and Y . Notice that $\mathcal{H}(X) \equiv D(X, Y)$ in case of perfect evaluation of both X and Y . Finally, a hat denotes estimates while a tilde the judgment-biased objects.

3.3. Characterizing judgmental dynamics

To make the previously notation operational we need the following

Assumption 1. $\mathcal{P}$ is a strictly convex probability measure.

Strict convexity of $\mathcal{P}$ implies that $\mathcal{U}$ is strictly proper. In turn, this requires the use of a more general distance measure compared to the one generally assumed in forecast evaluation exercises; see, among others Gneiting and Raftery (2007). The next assumption characterizes the statistical treatment of this paper:

Assumption 2. (i) $\mathcal{U}(\cdot, \cdot, \cdot)$ and q exist; (ii) they are not observed.

Assumption 2 (i) ensures that data are affected by bias. This assumption is supported by the literature mentioned in Section 2. **Assumption 2** (ii) implies that there are at least two sources of uncertainty: the (subjective) utility function

and the (objective) degree of judgment. Whereas the latter is known, it would still be possible to recover $\hat{\mathcal{U}}$ via state-space representation of the ARX(p) and to assume a quadratic loss function, so that OLS and Kalman filtering work. When the judgment term is also unknown, the classical state-space modelling is still possible, but the Kalman filter is not the minimizer of the mean square criterion because q_i is treated as a non-stochastic process that can only be inferred. The next Subsection deals with this issue.

The following definition unifies all the objects previously introduced in a common environment:

Definition 2 (Dynamically judgmental system). We define *Dynamically Judgmental (Generating) System*—DJS, henceforth—the 7-ple

$$\mathcal{DJS} := \{\mathcal{Y}_t, \mathcal{X}_t, \mathcal{F}_t, \mathcal{P}, S(\cdot, \cdot), H(\cdot), D(\cdot, \cdot)\} \text{ where } [Y, X] \text{ is measured by a T-dimensional dynamic system.}$$

Example 1. (i) A p-order autoregression with explanatory variables, ARX(p), producing density forecasts $\tilde{p}(\cdot)_{t+1}$ using a T-dimensional dynamic system is a DJS, where professional forecasters' utility corresponds to $\tilde{\mathcal{U}}(\mathcal{L}, \hat{\theta}; \cdot)$ with $\hat{\theta} \in \Theta$ is a DJS, respectively. (ii) A p-order autoregression with explanatory variables, ARX(p), producing density forecasts $\tilde{p}(\cdot)_{t+1}$ using a static regression framework is a Judgmental System (JS) but not a DJS.

4. Statistical methodology

This section illustrates the statistical treatment of DJS. Namely, [Section 4.1](#) summarizes the main results by [Ferrari and Yang \(2010\)](#) on Deformed Likelihood estimation; [Section 4.2](#) discusses the representation of the DSS in state-space form; finally, [Section 4.3](#) introduces the DKF which represents the main methodological innovation and the intersection of all topics presented in previous Subsections.

4.1. Robust estimation

The Deformed Likelihood is based on the following

Definition 3. (i) Consider any $x \in [-\infty, 0) \cup (0, +\infty]$; then its Lq-transform (or Box-Cox transform) is

$$L_q(x) = \begin{cases} \log(x) & \text{if } q = 1. \\ x^{1-q}/(1-q) & \text{otherwise.} \end{cases} \quad (1)$$

(ii) Let $y_1, \dots, y_T$ be an i.i.d. sample from $p(y_i, \theta_0)$, $\theta_0 \in \Theta$, where all the elements are defined in [Section 3.2](#). Then the Maximum Lq-Estimator (MLqE) of θ_0 is

$$\hat{\theta}_T \doteq \max_{\hat{\theta} \in \Theta} \sum_{t=1}^T L_q[p(y_t; \theta)], 0 < q < 1. \quad (2)$$

(iii) Moreover, if $p(\cdot)$ is Gaussian, its Deformed Likelihood is:

$$\mathcal{L}^q(\theta; y_t) = -0.5 * [q_0 T \log(2\pi) + (\log p(y_t) + u_t^2)^q], \quad (3)$$

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

Remark 2. The requirement that $0 < q < 1$ is an identifying condition. When $q < 0$, the form of the curve changes and becomes “U”-shaped with some points coinciding with the $q > 1$ case. When $q > 1$ the function (i) tends to explode and (ii) is discontinuous around zero. This avoids the inversion of the Lq-function and its derivatives.

Remark 3. Equation (2) is the result of the maximization of the Lq-likelihood equation, which, at the t-observation, takes the form:

$$\sum_{t=1}^T w_t U(y_t, \theta) = 0 \quad (4)$$

which is a weighted version of the likelihood equation with $U(y_t, \theta) = p(y_t; \theta)' / p(y_t; \theta)$ and weights $w_t = p_t(y_t; \theta)^{1-q}$. When $q < 1$, data points with high likelihoods are assigned large weights. As q tends to 1, the MLqE coincides to standard Maximum Likelihood Estimator (MLE, henceforth). Typically, outliers are associated to very small weights. The estimation of (4) in a time series framework is the ultimate aim of this paper.

A small deviation from 1 is sufficient to noticeably deform the log-Likelihood; see the Supplement for a graphical view of Deformed Logarithm and likelihood functions. Thus, it is fundamental to localize the moments in which this deformation from the LogLikelihood is most problematic. The following result goes in this direction:

Lemma 1. If $p(\cdot, \cdot)$ is Gaussian, then the estimated mean of $\hat{\theta}$ does not depend on q .

Proof. The surrogate parameter for the normal equation in a Gaussian regression is $\theta_t = (\mu', \sqrt{q} \text{vech}(\Sigma))'$. See Example 3.2 in Ferrari and Yang. $\square$

The parameter q is a point measure of judgment in the estimated model. Analyzing the effects of this deformation parameter requires to measure the discrepancy between U obtained (a) when event realizes using the same $p(\cdot, \cdot)$ or one of its moments; and (b) the maximal utility obtained by F when no bias is assumed. When this discrepancy is computable, it is possible to make comparison and hypothesis testing to verify the effectiveness of the bias, hence validate the claim on the predictive density $p_{t+1}(\cdot)$ using the empirical equivalent of (a) and (b). By Lemma 1, these are the matrices of V_t , (the optimal, non-biased variance) and $\tilde{V}_t$ (the matrices of observed, potentially biased variance). We suggest the use of the Brègman distance:

$$D(\tilde{V}_t, V_t) = \Lambda(\tilde{V}) - \Lambda(V) + \Lambda'(\tilde{V}) \text{vech}(\tilde{V} - V_t), \quad (5)$$

where $\Lambda : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}$, with $\mathbb{R}^{N \times N}$ being the space of positive semi-definite matrices, is a scalar function three times continuously differentiable with $\Lambda(V_t) = \nabla(V_t)$ and $\Lambda'(V_t) = \nabla^2 \Lambda(V_t)$, denoting the gradient and the Hessian of Λ with respect to the $K = N(N+1)/2$ unique elements of V_t and $\Lambda'(V)$ is negative semidefinite.⁶ An important question is why adopt the MLqE for our problem, being this just one of the many possible families of estimators available.⁷ We answer this question by using some arguments of Information Theory. Namely, the microeconomic foundation and set-up of Judgmental Protocol implies that DJS is a system characterized by Shannon Entropy, the most standard way to measure the uncertainty in complex systems.⁸ The use of this kind of entropy, among the many nowadays available, allows us to simplify our mathematical structure and to rely on standard maximization *at least* in the ideal case of no judgment. This implies that we can still rely on several standard results to make estimation, provided that we can represent a forecasting system according to a standard Entropy-based measure. This problem is nontrivial and can be solved in two steps: (i) prove that any optimal estimation in DJS leads to entropy-based methods; (ii) conversely, prove that any entropic estimation requires the DJS representation.

In the first step need to show that (i) there is a one-to-one relation between Shannon Entropy and the maximum possible utility that agents get in the forecasting system and, at the same time, between (5) and the Kullback-Liebler divergence – the most classical and simple measure of the distance among two density functions; (ii) and, as a consequence, we need to show that any minimization of (5) coincides with the minimization of the Shannon Entropy. These results are obtained by the following known preliminary results:

Lemma 2. Consider the Deformed Logarithm in (1). Then, (i) if q approaches to 1, $H_q(\cdot, \cdot)$ nests the Shannon Entropy and $D_q(\cdot, \cdot)$ the Kullback-Liebler Divergence. (ii) The minimizer over θ of $D_q(\theta, \theta^0)$, where θ^0 is the true parameter, is the same as the minimizer $H_q(\theta_0, \theta)$ where $q = 1/r$ and $r > 0$.

Proof. See Ferrari and Yang, pp. 755–756. $\square$

In the second step, we need to prove that our DJS is effectively an object that nests the q -Entropy and H -function previously discussed. This is possible by the following.

Lemma 3. The q -Entropy is the H -function corresponding to the DJS.

Proof. See the Appendix. $\square$

4.2. State-space representation

Until now we have introduced a set of elements for economic and statistical treatment of judgment in a forecasting system without any consideration on the dynamics. Characterizing the law of motion of such a system is the next definition's goal.

Definition 4 (State-space representation). Consider the n -dimensional univariate time series y_t of forecasting outputs. Then the state-space form for y_t is the following system of equations

$$\begin{aligned} y_t &= Z_t \alpha_t + X_t \beta + G_t \epsilon_t, \epsilon_t \sim iid(0, \sigma_\epsilon^2); \\ \alpha_{t+1} &= T_t \alpha_t + W \beta + H_t \eta_t, \eta_t \sim iid(0, \sigma_\eta^2), \\ Z &= [1, \mathbf{0}_{m-1}]; X = [1, \mathbf{0}_{k-1}]; H = \phi' \\ G &= \begin{bmatrix} I_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix}; T = \begin{bmatrix} \phi' \\ G \end{bmatrix}; W = [\phi', G], \end{aligned} \quad (6)$$

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 ϕ a

⁶ See Laurent et al. (2013) for theoretical properties of this general family of distance and its use in multivariate forecasting evaluation exercise.

⁷ Among the most used in statistical literature, we mention the Functional Likelihood, the Local Likelihood and Empirical Likelihood approaches; see, Solari (1969), Tibshirani and Hastie (1987), Chen and Qin (1993) *inter alia*. Notice that these methods can also be combined.

⁸ See Maasoumi (1993) for a compendium of Entropy-based methods for econometric applications.

$(p + 1)$ vector of AR parameters. The initial conditions are:

$$\begin{aligned} \alpha_0 &= [\mathbf{0}_m]; \beta_0 = [\mathbf{0}_k]; \mathbf{I}_{m^2} = \mathbf{I} \otimes [T, T]; \\ H^2 &= HH'; \text{vec}(P) = \mathbf{I}_m^{-1} H^2. \end{aligned} \quad (7)$$

The first equation of the system (6), known as *observation equation*, has a linear regression structure where the time-varying coefficient vector α_t characterizes the strength of the process' oscillations; the second equation, known as *state equation*, is a first order (vector) autoregression and propagates the properties of Markov-processes to the system. The initial conditions are necessary to make the estimation by Kalman Filter operational under the general assumption that some elements in the initial state $\alpha_1 \sim N(a_1, P_1)$, which distribution is independent on $\epsilon_1, \dots, \epsilon_n$ and $\eta_1, \dots, \eta_n$, are unknown. Then, we call on the following assumptions:

Assumption 3. (i) $E(\epsilon_t, \epsilon_s) = 0$ for all $t \neq s$; (ii) $E(\eta_t, \eta_s) = 0$ for all $t \neq s$; (iii) $E(\epsilon_t, \eta_t) = 0$; (iv) $E(\alpha_0, \epsilon_t) = 0$ for all $t = 1 \dots n$.

A3 (i)–(iv) are standard in the literature and set to simplify the notation and treatment. The next result provides the mathematical foundations to link (6) to the DJS:

Lemma 4. (i) *The DJS is never isomorphic to JS.*

(ii) *The DJS is isomorphic to (6).*

Proof. See the Appendix. $\square$

This concludes the representation of the DJS via state-space modelling.

4.3. The deformed Kalman filter

Once the representation of the DJS via state-space modelling has been demonstrated, it is possible to study its estimation via robust signal extraction. The corresponding DKF is derived by the following

Proposition 1. *Consider the system (6). Under A1 (i)–(iv) and the MLqE in (2)–(4), the DKF recursive equations are:*

(i) For $1, \dots, t, \dots, n$,

$$v = y_t - Z\alpha_t - X\beta; F = ZPZ' + GG'; \quad (8)$$

$$C = PZ'/F; t = v/\sqrt{q_0 F^{(q_0-1)}}; \quad (9)$$

$$v_t = \alpha + C\sqrt{F}h; P_t = \alpha_t + Cq_0 F C'(h/t); Q = HG'/F; \quad (10)$$

$$P_{t+1|t} = TV_t T' + HH' - (QFQ' + QFC'T' + TCFQ')(h/t); \quad (11)$$

$$\alpha_{t+1|t} = Tv_t + W\beta + Q\sqrt{(q_0 F)}(h/t); \quad (12)$$

where, for an arbitrary small number of time periods t^* ,

$$h = \begin{cases} t & \text{if } (t < t^*) \\ h(t, a, b) & \text{otherwise,} \end{cases} \quad (13)$$

and

$$h(t, a, b) = \begin{cases} t & \text{if } |t| \leq a, \\ \frac{a}{b-a}(b-t) & \text{if } a < t \leq b, \\ \frac{a}{b-a}(b+t) & \text{if } -b < t \leq t-a, \\ 0 & \text{if } |t| \geq b. \end{cases} \quad (14)$$

(ii) Then, by setting $I = v$ and $\sigma_I^2 = F$, we get

$$L = \log(F) + \log(q_0); \quad S = v^2/F; \quad (15)$$

$$y_f = Zv_t + GG'\sqrt{q_0 F}h/(q_0 F); \quad h_t = h/t; \quad (16)$$

$$\alpha_{t+1} = a; \quad \Sigma_{t+1} = \text{diag}(P) \quad (17)$$

Proof. See the Appendix. $\square$

The function h in (13) is a selector among time units that exceed a minimal value constituted by the two piecewise Hampel function $h(t, a, b)$ in (14). The latter is part of a family of functions—called *influence functions*—originally proposed in Robust Statistics to measure the dependence of an estimator on the value of any one of the points in the sample. Influence functions pre-select the path of the optimal trajectory of the (mean) estimator in a sample characterized by outliers, acting as an algorithm for automatic detection, weighting or removing outliers from a sample. This is particularly useful in long-span time series or in large databases where the identification of outliers is not trivial; see Hampel et al. (1986) among others. In our framework, the role of the Hampel function is not interesting *per se* because it serves as a “link” among the deformation parameter and the time recursions. Such a link is applied primarily via the deformed standard deviation t in (9) and the estimated deformed state-dependent residual v_t and the deformed variance matrix P_t in Eq. (10); subsequently, via the conditional forecasts of the variance $P_{t+1|t}$ and the forecasted conditional states $\alpha_{t+1|t}$. The formal link between the deformation parameter and time is provided by the following

Corollary 1. From the above recursions we get the Deformed Likelihood's weights:

$$w_t = 1/(L)^{q_0} \mathcal{G}_t \quad (18)$$

and the averaged measures:

$$\mathcal{L} = -0.5[T \log(2\pi) + L + S]; \quad \mathcal{L}c = -0.5[T(\log(2\pi S) + 1) + L]; \quad (19)$$

$$\mathcal{L}q = 0.5((q_0 T \log(2\pi) + (L + S)^{q_0}); \quad \hat{\mathcal{L}}q = 0.5((\hat{q} T \log(2\pi) + (L + S)^{\hat{q}}). \quad (20)$$

Proof. Trivial if defining $S = S/T$ and $\hat{q} = \sum_{t=1}^T h_t / T$ and $\mathcal{G}_t = -0.5((q_0 T \log(S^2) S)$.
□

Equation (18) is the core result of this paper because it defines the weight of deformation parameter q associated to each period via the gradient vector $\mathcal{G}_t$.

Finally, since Lemma 4 in the previous Subsection ensures that DJS can be represented by state-space model (6), we need to extend such a strict relationship also to their respective statistical treatments. This is possible by the following

Corollary 2. The DKF is isomorphic with (R)KF.

Proof. Let DJS_1 denote the DJS under Lq-transform, DJS_2 an equivalent without Lq-transformation and operator $\mathcal{T}$ the state-space system (6). By Lemma 3 and Lemma 4, it suffices to note that $\mathcal{T}$ is the same in either DJS_1 and DJS_2 , since DJS_2 is a DJS_1 with $q=1$. □

Remark 4. The role of explanatory variables is important for the economic interpretation of the DJS-DKF. In fact, as illustrated in the next Section, they represent the public information used by FED to make their forecasts. Thus, they are necessary to consider the parameter q as measure of PM's judgment—that is, private information as defined by Svensson (2005) and reported in the Introduction. Without X_t the model becomes a simple AR(p) model for a univariate time series and the state-space form and DKF works as well. However, q cannot be interpreted as measure of judgment but just as a tuning parameter of the Log-likelihood that summarizes the amount of deformation due to outliers—that is, something not related to private information.

5. Application

This section applies the DJS-DKF to real macroeconomic data. Section 5.1 describes the data; the results and their discussion are in Sections 5.2 and 5.3, respectively.

5.1. Data

The SPF-FED, which data collection starts in 1968, is the eldest dataset on the professional forecasting activity on a macroeconomic dataset. The survey was originally administrated jointly by the National Bureau of Economic Research (NBER) and the American Statistical Association. Since the second quarter of 1990, these two institution have been replaced by the Real Time Data Research Center of the Federal Reserve of Philadelphia. In more than 50 years, the SPF-FED has considerably changed and increased, both in its amplitude (from 10 variables to forecast in the first survey in 1968, to several dozens in the current release)⁹ and in its complexity (with increasing numbers and technicalities of therein items); see Croushore and Stark (2019).¹⁰

In particular, our research effort is stimulated by the fact that SPF-FED respondents are given a considerably short time to answer the survey.¹¹ Albeit assuming that the survey has been conducted by the best human and computational capital

⁹ We refer to the Release 2020 Q3.

¹⁰ See also the SPF website at <https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/survey-of-professional-forecasters>.

¹¹ “Currently, the forecasters are given just over a week to send in their forecasts” (Croushore and Stark, 2019, p. 1).

Table 1
In-samples and out-of-sample goodness-of-fit of DJS-DKF.

Model	Maximum Likelihood Estimation					
	Reality	Nowcasts	1-q-ahead	2-q-ahead	3-q-ahead	1-yr-ahead
RMSE	2.1703	3.3934	0.1702	0.2208	0.2746	0.6805
"In-sample RMSE"	1.2400	1.1449	0.0902	0.1044	0.1303	0.3502
"Out-of-sample RMSE"	1.2551	1.7220	0.0643	0.1205	0.1299	0.3730
			Maximum LqLikelihood Estimation			
RMSE	1.9257	2.7796	0.1681	0.2208	0.2746	0.6800
"In-sample" RMSE	1.1745	1.0043	0.0640	0.1035	0.1300	0.3500
"Out-of-sample RMSE"	1.2250	1.6881	0.0533	0.1192	0.1295	0.3765

NOTE: This table reports the RMSE of the AR(2)-DJS 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.

in the market—so that publicity-seeking behavior (Laster et al., 1999) and noise-information (Orphanides, 2003), among the most popular arguments for criticism of the use of survey data in Macroeconomics, may be annihilated—the use of non-sample estimates in forecasters' output cannot be neglected completely.

This paper focuses on Real GDP growth as a case study variable due to its universal use for monitoring and addressing economic policies by deputed institutions. A set of 4 additional variables are also included to verify the behavior of the DJS-DKF with exogenous variables. These last are: production, consumption, money velocity and house prices; see the Supplement for details.¹² We apply our DSS-DKF over 10 systems of variables where all combinations are considered. All the real data—apart the SPF forecasts—has been downloaded by FRED in quarterly frequency and, whenever possible, seasonally adjusted.

5.2. Results

The estimates of the aggregated quote of judgment and the main (deformed) functionals and variance parameters are reported in the Supplement. In summary, the forecasts of the U.S. RGDP are characterized by, approximately, 20% of judgment. The first issue to address is the credibility of these results. For this purpose, we set-up a MonteCarlo experiment—also in the Supplement. According to our results, there is no evidence of a systematic over evaluation of the q-parameter. Additionally, Fig. 3 plots the estimates of the corresponding dynamic system. In this exercise we make two trials, each corresponding to a different initial value of q. In both cases the estimated parameter is always 1 (that is, zero effect in judgment). Since in this case we assumed no judgment (because we only used real data) this result ensures credibility to our method.

The second issue to address is the time-varying evidence of the judgment. When is it more evident? The outputs of DJS-DKF for univariate time series, plotted in Fig. 4, panel (a), are coincident with Fig. 3, panels (c) and (d), if considering only historical data. Differently, other univariate time series corresponding to SPF nowcasts and forecasts vary considerably—and in a similar fashion to Tables 4 and 5 in the Supplement. Things change when looking at nowcasts (Fig. 4, panel b), where several negative peaks can be noticed at regular intervals in the first half of the sample, while in the second half, they tend to appear only in correspondence of periods of economic crisis. Instead, the one-quarter-ahead forecasts are characterized by a weighting function generally near-zero with important jumps, the frequency of which varies considerably according to the model.

The two/three/four-quarters-ahead forecasts—reported in Fig. 4—are characterized by weighting function lying in 1 with a small number of downturns, generally in the first half of the sample with the only exceptions of the SARS-COV2 Pandemics where the evidence of judgment is considerable. However, such extreme changes in weighting function are characterized by a timing not consistent with what an agnostic analyst may expect: in most of the models, the evidence in support of judgment begins with a small deviation from 1 in 2001 and arrives to zero in 2005–2006 and (only in a few models for one quarter-ahead forecasts) in 2012–13. Interestingly, the official NBER recessions dates for the Great Recession are never involved.

A final issue is the effectiveness of the AR(2) estimated by MLqE with respect to the same model estimated via standard MLE. Figure 5 displays a comparisons of density estimates for the univariate time series as well as for the joint density of the system M5 (see the Supplement for details): the MLE-based estimates convey a joint density highly overlapped on one of the two explanatory variables (the industrial production in this case); on the contrary, the joint density estimated by MLqE is more centered among them. Not surprisingly, the pseudo-out-of-sample exercise exposed in Table 1 confirms that that DJS-DKF is, in general, superior to the standard autoregression based on MLE in terms of point forecasting.

¹² We remark that the aim of this illustration is not an assessment of the monetary policy, but only an experiment of a new methodology. Evaluating judgmental issue in monetary policy would require the use of theory-based model; see Monti (2010).

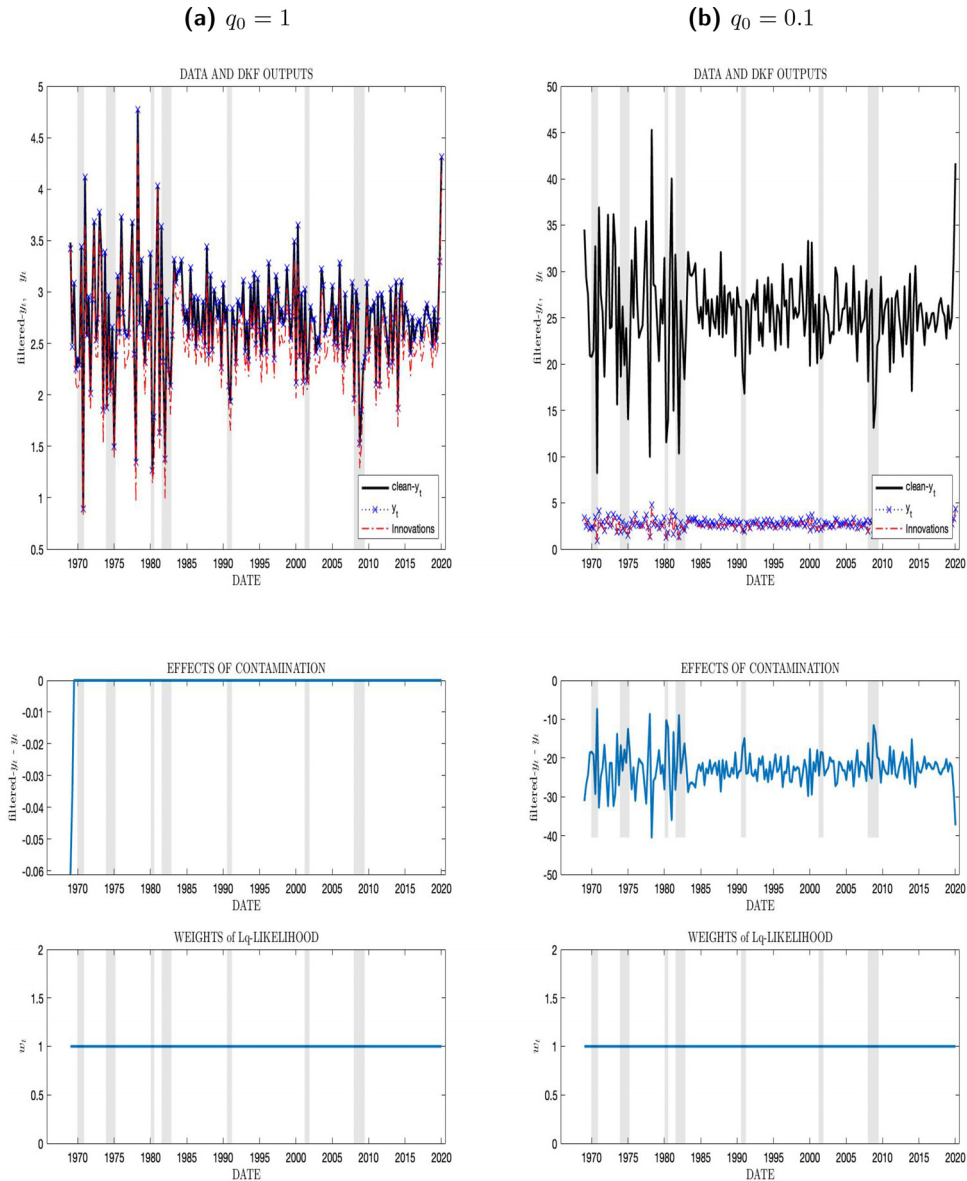


Fig. 3. Estimation of unobserved judgment on U.S. Real GDP using historical data only. *NOTE:* This figures plots the results of the application of the DKF on the data of U.S. Real GDP by the Federal Reserve Bank of St. Louis. The left panels deal with $q_0 = 1$ assumed in the DJS, while the right panels display the results of a DJS with $q = 0.1$. The top panels display the time series of the observations, the series of estimates as well as the vector of innovations; the bottom panels explicit the effect of the DKF: namely, central sub-panels display the filtered series and the bottom sub-panels plot the weighting function (18).

5.3. Discussion

Several considerations can be made from the above empirical investigation: first, there is a dynamics in the aggregate quote of judgment during the span of the sample, and it varies considerably if considering the two subsamples 1968–1990 and 1991–2020. Namely, the first half of the sample is characterized by an unclear prevalence of the data on non-data estimation. We interpret this finding as a consequence of three main factors: (a) the poor quality of the first decades of observations, due to the experimental state of the forecasting science and survey administration; (b) the evolution of econometric and computing techniques gradually made the use of statistical tools more affordable for economists who were asked to monitor the U.S. economy; (c) the changes in forecast making and in data collection, occurred during the two sub-samples (like, for example, the change in the GDP deflator). These factors, *per se* not sufficient to explain the judgmental dynamics' evolution, influence the utility function of the Forecaster. A change in utility function is very well-known to explain the persistence of the (non-judgmental) bias according to the standard theory, see [Patton and Timmermann \(2007\)](#).

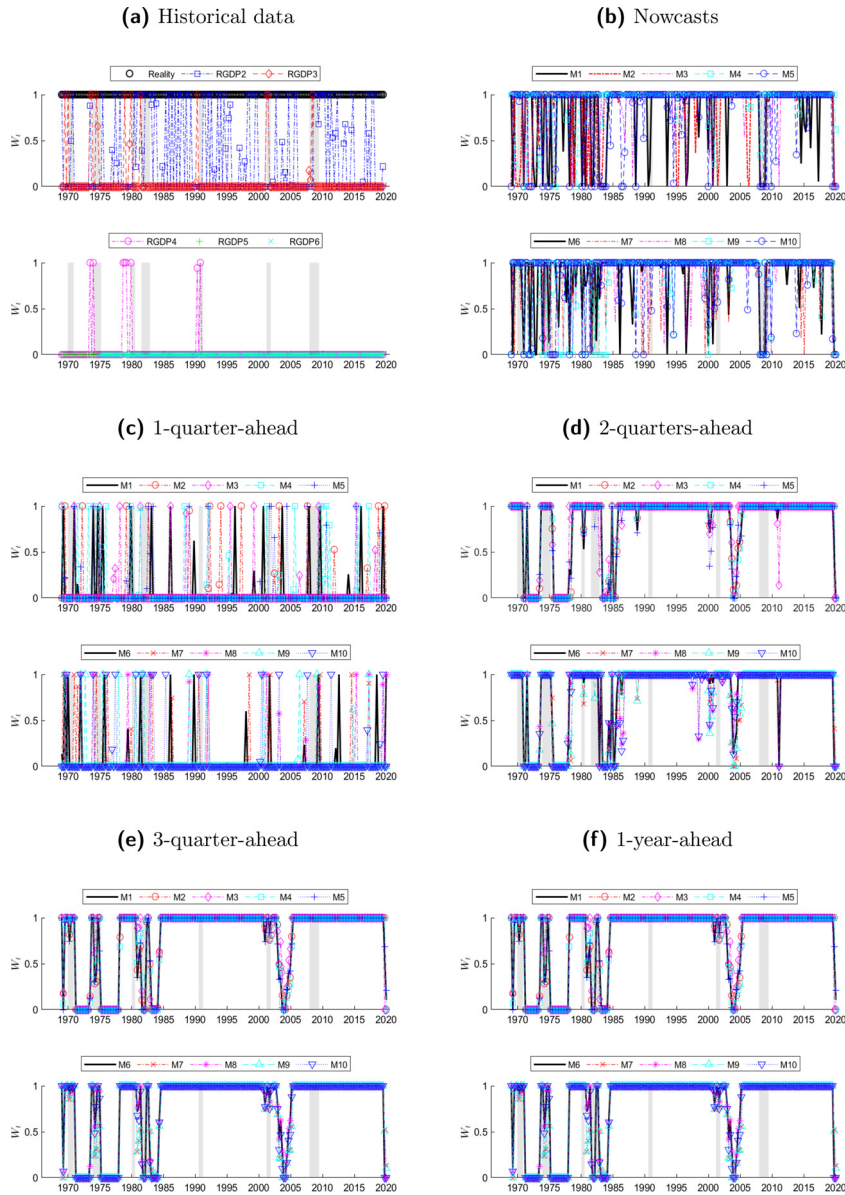


Fig. 4. Estimation of unobserved judgment on U.S. Real GDP forecasts using several DSS specifications. *NOTE:* This figure plots the results of the application of the DJS on U.S. Real GDP forecast by SPF using several data and model specifications.

Thus, the problem becomes how to relate utility function to judgment and how to do testing and inference in a judgmental forecasting system (that is, how to relate the F change in utility to PMs decision). This means that the judgment becomes an indicator of a strategic behavior among the agents involved in the forecasting system. We are currently focusing on this and other related issues.

Second, the non coinciding dynamics of the switches, in the weighting function, of the majority of models with NBER recession dates implies that professional macroforecasters do not change their attitudes mechanically or according to a simple algorithm like “*use-judgment-when-recession-arrives*”. Instead, several different, more complex reasoning schemes have to be assumed to justify their output. This seems consistent with [Coibion et al. \(2018\)](#)’s claim for a novel rethinking of the expectation formation. In this sense, the mechanics of endogenous formation of misvaluation that can be inferred by [Ilut and Valchev \(2022\)](#) is confirmed by our evidence.

This finding questions the effectiveness of data revisions that FED makes periodically: if a certain amount of judgment is proved, its role in a forecast revision is not. One may argue that neutralizing the professional forecasters’ judgment is exactly one of the roles of forecast revision. This consideration implies that revisions should coincide, or be in the neighborhood, with the degree of judgment found in our analysis. Instead, the revisions certified by the Bureau of Economic Analysis (BEA)

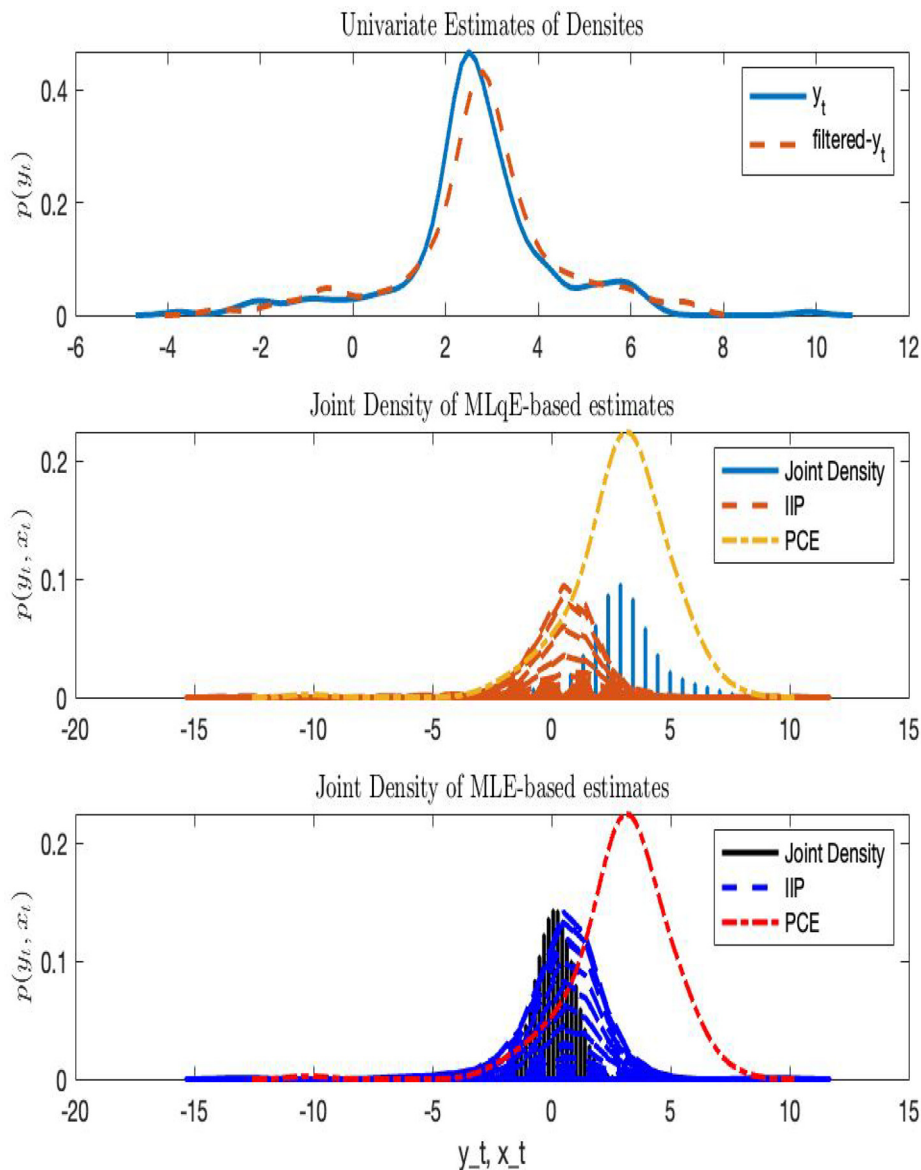


Fig. 5. Comparing MLE and MLqE in Density Estimation. *NOTE:* This figure plots (in the upper panel) the estimates of univariate probability density functions of data and the DKF-estimates of M5 (a system with one-year-ahead RGDP forecasts as dependent variables and IIP and PCE as exogenous variables; in the central panel, the univariate density functions of IIP and PCE as well as their joint density function by MLqE; and, in the lower panel, the univariate density functions of IIP and PCE as well as their joint density function by MLE. See the Supplement for further data explanation and labels.

in the period 1993–2019 reject this hypothesis: the role of revisions has been considerably small, only between 0.5 and 1.2 percentage points of the estimates (on average), while a simple graphical inspection of Fig. 3 and Table 4 in the Supplement suggests the bias is higher.¹³ A future development of our DJS methodology that also takes into account the effect of FED's staff revisions is highly recommended, specially in light of the results by Broer and Kohlihas (2022), according to which the forecasts revisions can be seen as a form of strategic behavior.

The finding that the amount of judgment is not negligible—especially, in some periods—should not necessarily be interpreted as a claim that FED staff does not use all the available information but as the evidence that the survey respondents diversify their effort among computational mechanics and human deliberation. This interpretation is coherent with Casey (2020), according to which the three main surveys of professional forecasts are strongly driven by macroeconomic theoretical relations, so that the judgment may be seen as a link among theory and final forecasts—albeit, the recent liter-

¹³ See the BEA website: <https://www.bea.gov/gdp-revision-information>.

ature suggests that professional forecasters are generally over-confident. In turn, this last finding makes us to question the nature of the link between judgment and strategic behavior. Further theoretical research will be necessary.

Finally, we are aware about the limits of our methodology. The DJS is a very general framework that applies a (time series) regression framework on data coming from two different agents with (possibly) different utility functions. Thus, the current parametrization does not allow to discriminate the contribution of each single individual survey respondent, unless repeating the same DJS on individual data. Moreover, the Deformed Likelihood estimator, due to its logarithmic structure, may not be the best option in presence of repeated, large outliers. In this last case some extreme-value method like, for example, the one proposed by [Burridge and Taylor \(2006\)](#) may be preferable. Our simulations and empirical illustrations do not solve the issue of causality among judgment and the use of selected indicators. This issue requires an analytical derivation of the LqLikelihood-based estimator from a fully micro-founded forecasting scenario. While a game-theoretic foundation of forecasting has been proposed by [Vovk and Shafer \(2005\)](#), and is currently under development for econometric application in our current research, the causality nexus of judgment with macroeconomic variables is still an open issue.

6. Conclusions

The professional survey-based forecasts for the U.S. RGDP are characterized by a significant, persistent bias. The economic literature has explained this puzzling result by using several arguments—professional forecasters' asymmetry in utility function and the lack of agent's rationality, are among the most successful. This paper contributes to the macroeconomic literature on macroforecasters' bias by introducing a novel judgmental approach to belief formation. The judgment, defined as everything outside the scope of the econometric model generating the RGDP forecasts, is the output of a forecasting game among three Players. The econometric counterpart of this repeated game is called Dynamic Judgmental System. For the first time in the literature, we estimate the dynamics of unobserved judgment in survey data owned by Federal Reserve Bank. This is possible by combining the properties of the Deformed Likelihood estimation and (robust) signal extraction in a unique. The corresponding Deformed Kalman Filter allows econometricians to extrapolate the judgment from a time series of forecasting outputs using a set of explanatory variables.

Our simulation experiments reveal that the DJS-DKF does not tend to over-evaluate the judgmental quote in forecasting activity. The empirical evidence on a case study on the U.S. survey data on RGDP supports the hypothesis that professional macroforecasters have a non negligible amount of judgment which is not distributed uniformly during the span of the sample. Moreover, the knowledge of time-varying judgment is useful in applied forecasting exercises. These results confirm the recent claims by several authors that the rational-expectation framework, despite the recent refinements, should be severely modified or substituted by a more realistic hypothesis.

The proposed methodology is still in early-stages of development. Further research is required to understand the inner motivations of such a dynamics and the capability of DJS in more complex forecasting environments.

Appendix A

A1. Proof of Lemma 1

To prove an isomorphism among two sets it suffices to show that these two differ at most by the nature of their elements; see [Kreyszig \(1991\)](#), among others. DJS and JS do not differ by the nature of the observations enclosed in the 7-ple $\{\mathcal{Y}_t, \mathcal{X}_t, \mathcal{F}_t, \mathcal{P}, S(\cdot, \cdot), H(\cdot), D(\cdot, \cdot)\}$, but in the way in which the collection of $[\mathcal{Y}, \mathcal{X}]_t$ is estimated—by dynamic system, the former and by static regression, the latter. On the contrary, the DSS and (6) may differ at most by the nature of the observations in $\mathcal{Z}_t$.

A2. Proof of Lemma 4

It descends directly from (a) the part (ii) of Lemma 1 and (b) Lemma 2. Let's consider each of them separately. (a) the part (ii) of Lemma 1 is necessary but not sufficient. If DJS were not isomorphic to the state-space form in equation 6, there could exist an H -function different from H_q that maximizes $U(X, Y)$ given q . On the opposite side, assume there exists a H -function different from H_q that maximizes $U(X, Y)$ given q ; in this case, there is no implication of uniqueness of q nor $H \neq H_q$. (b) Lemma 2 is a sufficient, non-necessary condition for Corollary 1 to hold. Part (i) ensures a one-to-one relation among q -Entropy, a known D -function and the DJS; part (ii) ensures the one-to-one relationship among the true parameter θ_0 and the amount of the aggregate quote of judgment represented by q .

A3. Proof of Proposition 1

(i) CASE 1: $q=1$. By Lemma 3, the Lq-Likelihood function coincides with the standard Likelihood; by Corollary 2, the DKF coincides with the RKF. Thus, the proof is delegated to [Kalman \(1960\)](#), [Duncan and Horn \(1972\)](#), [Harvey and Phillips \(1979\)](#), [Marczak et al. \(2018\)](#). CASE 2: $0 < q < 1$. Since $q \neq 1$ implies a bias to $\hat{\alpha}$ and $\hat{\beta}$, these are no more the minimum mean square error of α and β . Thus, we only need to justify the introduction of (a) $t = \nu/\sqrt{q_0 F^{q_0-1}}$ in eq. (10)–(12); (b) q_0 in the second addend of (11).

Lets start from (b). By Lemma 2, we only need to notice that, in the [Harvey and Phillips \(1979\)](#) notation here adopted, the variance is parametrized by F via P_t , and that F is a scalar. Thus, there is no need of half-vectorizing P . The same argument holds for equation (13).

To prove (a), let us remark that, in exponential family, $\theta^* = \theta_0/q$, where $\theta = [\mu, \sigma^2]$. By Lemma 2, μ is not influenced by q , thus it suffices to consider σ^2/q , where σ^2 is known. Thus, the only unknown variable is $1/q$. According to the normal [Eq. \(4\)](#), the optimal solution is the first derivative of $\sqrt{F}q$.

(ii) Direct consequence of (i).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.jedc.2023.104632](https://doi.org/10.1016/j.jedc.2023.104632).

References

- Abel, J., Rich, R., Song, J., Tracy, J., 2016. The measurement and behavior of uncertainty: evidence from the ECB survey of professional forecasters. *J. Appl. Econom.* 31 (3), 533–550.
- Andrade, P., Crump, R., Eusepi, S., Moench, E., 2016. Fundamental disagreement. *J. Monet. Econ.* 83, 106–128.
- Andrade, P., Le Bihan, H., 2013. Inattentive professional forecasters. *J. Monet. Econ.* 60 (8), 967–982.
- Bianchi, F., Ludvigson, S.C., Ma, S., 2022. Belief distortions and macroeconomic fluctuations. *Am. Econ. Rev.* 105 (8), 2269–2315.
- Boero, G., Smith, J., Wallis, K., 2008. Uncertainty and disagreement in economic prediction: the Bank of England survey of external forecasters. *Econ. J.* 118 (530), 1107–1127.
- Bordalo, P., Gennaioli, N., La Porta, R., Shleifer, A., 2019. Diagnostic expectations and stock returns. *J. Finance* 74 (6), 2839–2874.
- Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2020. Overreaction in macroeconomic expectations. *Am. Econ. Rev.* 110 (9), 2748–2782.
- Broer, T., Kohlhas, A., 2022. Forecasters mis-behavior. *Rev. Econ. Stat.* 1–45 [10.1162/rest_a.01210](https://doi.org/10.1162/rest_a.01210).
- Burrage, P., Taylor, R., 2006. Additive outlier detection via extreme-value theory. *J. Time Ser. Anal.* 27 (5), 685–701.
- Capistrán, C., Timmermann, A., 2009. Disagreement and biases in inflation expectations. *J. Money Credit Bank.* 41 (2–3), 365–396.
- Casey, E., 2020. Do macroeconomic forecasters use macroeconomics to forecast? *Int. J. Forecast.* 36 (4), 1439–1453.
- Chen, J., Qin, J., 1993. Empirical likelihood estimation for finite populations and the effective usage of auxiliary information. *Biometrika* 80 (1), 107–116.
- Clements, M., 2014. Forecast uncertainty ex ante and ex post: US inflation and output growth. *J. Bus. Econ. Stat.* 32 (2), 206–216.
- Coibion, O., Gorodnichenko, Y., 2012. What can survey forecasts tell us about information rigidities? *J. Polit. Economy* 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.
- Coibion, O., Gorodnichenko, Y., Kamdar, R., 2018. The formation of expectations, inflation, and the Phillips curve. *J. Econ. Lit.* 56 (4), 1447–1491.
- Croushore, D., Stark, T., 2019. Fifty years of the survey of professional forecasters. *Econ. Insights* 4 (4), 1–11.
- Dawid, A., 1984. Present position and potential developments: some personal views statistical theory. The prequential approach (with Discussion). *J. R. Stat. Soc. Ser. A (General)* 147 (2), 278–290.
- De Long, J., Shleifer, A., Summers, L., Waldmann, R., 1989. The size and incidence of the losses from noise trading. *J. Finance* 44 (3), 681–696.
- Dovern, J., Fritsche, U., Slacalek, J., 2012. Disagreement among forecasters in G7 countries. *Rev. Econ. Stat.* 94 (4), 1081–1096.
- Duncan, D., Horn, S., 1972. Linear dynamic recursive estimation from the viewpoint of regression analysis. *J. Am. Stat. Assoc.* 67 (340), 815–821.
- Durbin, J., Koopman, S.J., 2012. *Time Series Analysis by State Space Methods*, Vol. 38. Oxford.
- Ehrbeck, T., Waldmann, R., 1996. Why are professional forecasters biased? Agency versus behavioral explanations. *Q. J. Econ.* 111 (1), 21–40.
- Ferrari, D., Yang, Y., 2010. Maximum Lq-likelihood. *Ann. Stat.* 38 (2), 753–783.
- Gaglianone, W.P., Giacomini, R., Issler, J.V., Skreta, V., 2022. Incentive-driven inattention. *J. Econom.* 231 (1), 188–212.
- Giacomini, R., Skreta, V., Turen, J., 2020. Heterogeneity, inattention, and Bayesian updates. *Am. Econ. J. Macroecon.* 12 (1), 282–309.
- Gneiting, T., Raftery, A., 2007. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* 102 (477), 359–378.
- Hampel, F., Ronchetti, E., Rousseeuw, P., Stahel, W., 1986. *Robust Statistics: The Approach Based on Influence Functions*. Wiley-Interscience, New York.
- Harvey, A., Phillips, G.D., 1979. Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika* 66 (1), 49–58.
- Ilut, C., Valchev, R., 2023. Economic agents as imperfect problem solvers. *Q. J. Econ.* 138 (1), 313–362.
- Jo, S., Sekkel, R., 2019. Macroeconomic uncertainty through the lens of professional forecasters. *J. Bus. Econ. Stat.* 37 (3), 436–446.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82 (1), 35–45.
- Kocięcki, A., Kolasa, M., Rubaszek, M., 2012. A Bayesian method of combining judgmental and model-based density forecasts. *Econ. Model.* 29 (4), 1349–1355.
- Kreyszig, E., 1991. *Introductory Functional Analysis with Applications*, Vol. 17. John Wiley & Sons.
- Lahiri, K., Sheng, X., 2010. Measuring forecast uncertainty by disagreement: the missing link. *J. Appl. Econom.* 25 (4), 514–538.
- Laster, D., Bennett, P., Geom, I.S., 1999. Rational bias in macroeconomic forecasts. *Q. J. Econ.* 114 (1), 293–318.
- Laurent, S., Rombouts, J., Violante, F., 2013. On loss functions and ranking forecasting performances of multivariate volatility models. *J. Econom.* 173 (1), 1–10.
- Maasoumi, E., 1993. A compendium to information theory in economics and econometrics. *Econom. Rev.* 12 (2), 137–181.
- Manganelli, S., 2009. Forecasting with judgment. *J. Bus. Econ. Stat.* 27 (4), 553–563.
- Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: a proposal to replace the new Keynesian Phillips curve. *Q. J. Econ.* 117 (4), 1295–1328.
- Manzan, S., 2011. Differential interpretation in the survey of professional forecasters. *J. Money Credit Bank.* 43 (5), 993–1017.
- Manzan, S., 2021. Are professional forecasters Bayesian? *J. Econ. Dyn. Control* 123, 104045.
- Marczak, M., Proietti, T., Grassi, S., 2018. A data-cleaning augmented Kalman filter for robust estimation of state space models. *Econom. Stat.* 5, 107–123.
- Masreliez, C., Martin, R., 1977. Robust Bayesian estimation for the linear model and robustifying the Kalman filter. *IEEE Trans. Automat. Contr.* 22 (3), 361–371.
- Monti, F., 2010. Combining judgment and models. *J. Money Credit Bank.* 42 (8), 1641–1662.
- Orphanides, A., 2003. Monetary policy evaluation with noisy information. *J. Monet. Econ.* 50 (3), 605–631.
- Ortoleva, P., 2012. Modeling the change of paradigm: non-Bayesian reactions to unexpected news. *Am. Econ. Rev.* 102, 2410–2436.
- Ottaviani, M., Sørensen, P., 2001. Information aggregation in debate: who should speak first? *J. Public Econ.* 81 (3), 393–421.
- Patton, A., Timmermann, A., 2007. Testing forecast optimality under unknown loss. *J. Am. Stat. Assoc.* 102 (480), 1172–1184.
- Patton, A., Timmermann, A., 2010. Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *J. Monet. Econ.* 57 (7), 803–820.
- Rossi, B., Sekhposyan, T., 2015. Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *Am. Econ. Rev.* 105 (5), 650–655.
- Sims, C., 2003. Implications of rational inattention. *J. Monet. Econ.* 50 (3), 665–690.
- Sims, C.A., 2002. The role of models and probabilities in the monetary policy process. *Brookings Pap. Econ. Act.* 2002 (2), 1–40.

- Solari, M.E., 1969. The maximum likelihood solution of the problem of estimating a linear functional relationship. *J. R. Stat. Soc. Ser. B (Methodological)* 31 (2), 372–375.
- Svensson, L., 2005. Monetary Policy with Judgment: Forecast Targeting.
- Tibshirani, R., Hastie, T., 1987. Local likelihood estimation. *J. Am. Stat. Assoc.* 82 (398), 559–567.
- Townsend, R., 1978. Market anticipations, rational expectations, and Bayesian analysis. *Int. Econ. Rev. (Philadelphia)* 481–494.
- Townsend, R.M., 1983. Forecasting the forecasts of others. *J. Polit. Economy* 91 (4), 546–588.
- Vovk, V., Shafer, G., 2005. Good randomized sequential probability forecasting is always possible. *J. R. Stat. Soc. Ser. B (Statistical Methodology)* 67 (5), 747–763.
- Woodford, M., 2003. Imperfect common knowledge and the effects of monetary policy. National Bureau of Economic Research Cambridge, Mass, USA.