UNIVERSITÀ DEGLI STUDI DI BERGAMO

Department of Economics

WORKING PAPERS

Judgment can spur long-memory

Emilio Zanetti Chini

July 2023 - WP N. 23 Year 2023

Working papers – Department of Economics
n. 23

Judgment can spur long-memory

UNIVERSITÀ DEGLI STUDI DI BERGAMO
Department of Economics

Emilio Zanetti Chini

Università degli Studi di Bergamo
2023

Judgement can spur long memory / Emilio Zanetti Chini – Bergamo: Università degli Studi di Bergamo, 2023.
Working papers of Deparment of Economics, n. 23
ISSN: 2974-5586
DOI: 10.13122/WPEconomics_23

Progetto grafico: Servizi Editoriali – Università degli Studi di Bergamo
Università degli Studi di Bergamo
via Salvecchio, 19
24129 Bergamo
Cod. Fiscale 80004350163
P. IVA 01612800167

https://aisberg.unibg.it/handle/10446/248469

Judgment can spur long-memory

EMILIO ZANETTI CHINI*

University of Bergamo
Department of Economics
Via dei Caniana, 2 - 24127, Bergamo (ITALY)
e-mail: emilio.zanettichini@unibg.it

May 2023

Abstract

We arrive at this conclusion by using a new family of models—the Long-Memory Judgmental Protocol (LMJP)—where robust filtering and fractionally integrated autoregressions are combined in an environment characterized by several players—namely, Forecast Producer, Forecast User and Reality. Our simulated and empirical evidence reveals that (i) knowledge of the LM parameter matters for the p-values of tests for spurious long memory; (ii) secondly that the role of LM in belief formation is ambiguous.

Keywords: Belief Formation, Dynamic Systems, Power Spectrum, Filtering.

JEL: C12, C22, C44, C53.

*The Author would like to thank Salvatore Piccolo and Rosario Crinò for their support. On multiple instances, he praises Barbara Annicchiarico, Alessandra Bonfiglioli and Francesco Ravazzolo for their remarks. Moreover, he thanks Katsumi Shimotsu for sharing his MatLab routines, which showed vital information for the enhancement of this note; of course, the author accepts 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.

1 Introduction

Numerous economic time series processes are characterized by a delayed degradation of sample autocovariance or by power spectrum stability. This quality is known as *long-memory* (LM, henceforth). See, for example, Robinson (2003) and Hassler (2018) for a comprehensive overview of the literature on this particular topic. Recent evidence suggests that (i) the assumption that economic agents learn from their macroeconomic system influences the long-run dynamics of economic time series, thereby generating long-memory (Chevillon and Mavroeidis, 2017); and (ii) the network structure of an economic system is, *per se*, sufficient to cause LM without using other explanations such as temporal aggregation, non-stationarity, or nonlinearity—and frequently all of these features coexist and conflate (Schennach, 2018).

In Section 2, we introduce a new, micro-funded explanation for the rise of LM: the *judgment*, defined as everything outside the scope of an economic or statistical model (Svensson, 2005)—or, in econometric terms, the non-sample information introduced in the estimation phase by a forecaster or a decision maker (Manganelli, 2009). Therefore, anytime reality does not confirm the prediction, a judgmental bias (JB) occurs. This does not, however, correlate with the 'traditional' one. In reality, as stated and shown by Zanetti Chini (2023a), it may be seen as the outcome of a recurrent game between two players, a forecast producer (FP) and a forecast consumer (FU, say a Central Banker). Each of them has a subjective utility function depending on the future realization and available information that causes the other to think that the genuine forecasts/decisions have not been accurately communicated. As a result, FU will misjudge FP, causing FP to get used to being misjudged and to compensate for this loss by adding another JB in the subsequent period, etc. This judgmental dynamics can effectively caught by a robust filtering.

By combining these two bodies of research, Section 3 presents evidence that (i) the presence of JB is linked with significant changes in the spectral density of the process and hence annihilates the presence of LM and (ii) the null hypothesis of false

LM for certain traditional macroeconomic data is not rejected. Section 4 concludes.

2 Modelling Judgment and Long-Memory

We move on the idea that the JB is something that distorts any function based on data available to the investigator.

Definition 1. Let $x \in \mathbb{R}$ and q the tuning parameter of a Lq-transform (or Box-Cox transform) defined as follows:

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

We then define deformation a realization of Lq with $q \neq 1$.

The way in which q distort the "true" values of x can be verified in Figure 1. One of the most universally adopted functions linking data (that is, x) to unknown parameter is the Likelihood. Hence, the Lq-Likelihood (or Deformed Likelihood) estimator (MLqE) arises as a natural application of (1). Such an estimator, in its original formulation assumes q as given. The estimation of q is explained in Appendix as well as in the aforementioned reference.

After having defined the notion of deformation in an operational way, we need to characterize its relation to the JB and its formation according to a simple microeconomic model. This can be done via the following set of assumptions:

Assumption 1. *(i) The economy is formed by three agents: FP, FU and Reality. (ii) FU's choice does not affect Reality when the process repeats, nor does FP respond to FU throughout the process. (iii) Reality is impartial about FU and FP judgments.*

In Assumption 1 (i) FP may represent both external experts or Central Bank's internal officers, but these latter have no participation in the ultimate decision. Assumption 1 (ii) follows because we don't presume agents' strategic conduct. Assumption 1 (iii) may appear contradictory from a macroeconomic policy standpoint, but from a microeconomic perspective, it is crucial to isolate the involvement of the two players in judgment creation across time. In this work, the judgment is denoted by $q \in (0, 1)$, where 0 represents pure judgment and 1 is no judgment.

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

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 gets the projection sets, adds judgment $\tilde{\hat{y}}_t$ to each, and selects which projection appears better based on her objectives and information;

4. FU releases next period's official prediction $\hat{y}_{t+1|t}$;

5. Reality reveals and is characterized by a long-memory parameter d.

Remark 1. The deformation parameter q can be seen as the aggregate quote of judgment in the estimation process, that is $q = \sum_{t=1}^{T} (\tilde{\hat{y}}_t / \hat{y}_t)$.

Remark 2. d arises naturally as the iterations of LMJP occurs. In the course of this paper we treat it via fractionally integrated autoregressive moving average (ARFIMA) model for y_t defined as

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

where $\Gamma()$ is the Gamma distribution, L is the lag operator and $\Phi(L) = 1-\phi_1 L-, \ldots, \phi_p L^p$ and $\Theta(L) = 1+\theta_1 L-, \ldots, \theta_r L^p$ the autoregressive and moving-average polynomials with all their roots outside the unit circle with no common factors. The term $(1-L)^d$, which is the fractional difference operator, induces the long memory feature. The parameter d defines the process's long-memory degree: namely, if $d > 1/2$, the process is invertible and has a linear description; if $d = 1/2$, the covariance is stationary. Remarkably, if $d > 0$, the process is considered to have a LM because the autocorrelations decay at a hyperbolic pace (and therefore no longer absolutely summable), as opposed to the considerably quicker exponential rate in the case of weak dependency. If $d = 0$, the spectral density is bound at the origin, and the process is an AR with just a little dependency (short memory). An alternative definition (Beran *et al.*, 2013) is based on the frequency domain, see the Appendix for details. For easy of treatment and economy of space we will assume $y \sim ARFIMA(p, d, 0)$–that is a pure (fractionally integrated) autoregression.

Then we are going to verify the following

Proposition 1. (a) The LMJP can be represented as dynamic system modelled via ARFIMA models. (b) Moreover, it can be estimated by Deformed Kalman Filter.

Proof. See the Appendix. □

Proposition 2. The judgmental parameter q deforms the LM parameter d in the Power Spectrum.

Proof. See the Appendix. □

The Proposition 2 is a testable hypothesis to be verified by classical testing. Namely, we are interested in testing the hypothesis system:

$$H_0 : d = 0 \qquad vs \qquad H_1 : d > 0, \tag{3}$$

Two modelling strategies are possible: a 'Judgment-First' approach consists in estimating the autoregressive order governing the DJP; alternatively, a 'LM-First' approach involves estimation of $\hat{q}_t$ via LMDJP in Supplement and using it as initial value to fit (2) via spectral density estimation exposed in Appendix.

3 Results

3.1 MonteCarlo Simulation

First of all, we apply the LMDJP to MonteCarlo simulated data to verify the effectiveness of Proposition 1. We consider two different DGPs:

$$y^{(i)}_{1,t} = 0.5y^{(i)}_{1,t-1} + \epsilon^{(i}_{1,t} \quad \epsilon^{(i)}_{1,t} \sim N(0, 1) \tag{4}$$

and

$$y^{(i)}_{2,t} = 0.5y^{(i)}_{2,t-1} - 0.88x^{(i)}_{t,2} + \epsilon^{(i)}_{2,t}, \quad \epsilon^{(i)}_{2,t} \sim N(0, 1), \tag{5}$$

where, in both (4) and (5), $i = \{1, \ldots, I\}$ denotes the i-th draw of the process $\{y_t\}^T_{t=1}$ with a total number of draws $I = 5,000$. Both the DGPs run on a set of starting q-values ($q_0 = \{0.1, 0.5, 1.0\}$ corresponding to extreme, medium and no amount of JB) and $d-$values ($d_0 = \{0.1, 0.5, 1.0\}$ corresponding to no memory, moderate long-memory and differenced process).

We propose paying specific attention to who looks at the analysis: in this example, (4) reflects what FP sees and analyzes (the autoregressive model is her subjective choice). More specifically, $y^{(i)}_{1,t}$ is a linear autoregressive model with substantially stationary behavior, allowing us to concentrate on the impacts of the starting value of q_0. The latter might be represented by a macroeconomic indicator evaluated with a particular quantity of JB in order to reduce the chance of loss in the next period.

$y^{(i)}_{2,t}$, on the other hand, depicts a more comprehensive scenario: the autoregression is triggered by an explanatory variable x_t, in addition to the initial judgment q_0. If y_t is assumed to be a time series of FU's final announcement, the entire DJP can be interpreted as a full dynamic system in which the FP's output (in this case, x_t) is an input that coexists with reality, and the difference between them is the basis for an ex-post evaluation of FU via utility function. This is integrated into the estimation step via the Lq-Likelihood. As a result, q_0 reflects a kind of a-priori of FU in relation to which FP changes her projections.

There are two design options for estimating $\hat{d}$: (a) '*LM-First*': estimate first (2) and only after that $\hat{d}$ is known, estimate the model LMDJP. In this scenario, $\hat{d}$ serves as a second a priori in addition to q_0. (b) '*Judgment-First*': the LMDJP is applied to data, and $\hat{d}$ is projected based on filtered y_t. The LM is a fully endogenous product of judgment in this case. As a result, the LMDJP is comparable to the framework of Chevillon and Mavroeidis (without the necessity for explicit expectation modeling) and to the theory of Schennach—though the network structure in our concept is minimal.

The estimate of $\hat{d}$ is extremely sensitive to algorithm design, according to the findings of Tables 1 and 2, which pertain to sample sizes of T=100 and T=500, respectively, since the differences in terms of power of the hypothesis test (3) are large. Several facts emerge in more detail: first, not all estimators perform equally well, as evidenced by the fact that LW and ELW estimators are the most performant in both algorithmic techniques, but ELW2 is a close second and 2SFELW performs badly, particularly in the "Long-Memory First" approach. Second, the first two estimators show significant variations in power between "Judgment-First" and "LM-First" procedures, whilst the others are insensitive to the two modeling strategies. Finally, the average $\hat{q}$ varies significantly between estimation methodologies and sample sizes; nevertheless, this fact does not indicate a flaw in the estimate approach *per se*, since q_0 is not the genuine level but only the beginning value. A first result of this experiment can be appreciated in Figure 2, where the average Power Spectrum resulting from the DGP is displayed taking the LM parameter $d_0 = 0$ fixed: unless the pure autoregressive case under no judgment, the spectra displays a shift which length depends on the presence of exogenous regressors (more than 100% for pure AR, around 60% in ARX case). Why the two modelling strategies produces a difference in power of the LM test? Figure 4 gives a possible explanation: when q_0 is not 1 (that is, when an initial judgment level exists), it raises the power spectral densities of the filtered processes significantly in the majority of cases.

3.2 Three Empirical Examples

The LMDJP is applied to three separate real-world case studies. The first one is to run a purely autoregressive LMDJP-AR(p) model on macroeconomic data using Nelson and Plosser (1982)—a standard for applied Macroeconometrics and time series analysis literature. This is composed of 14 macroeconomic variables on US economy (GNP and its deflator, Employment, Industrial Production, Money Velocity and so on) collected at yearly frequency and ranging from various starting dates up to 1970. According to Table 3, the predicted level of $\hat{d}$ varies significantly across time series and estimation strategies. Noticeably almost all of the series fail the test for no LM in the "LM-first" approach, while only a minority does in the "Judgment-First" strategy. The 2SFELW estimator in the last approach is insensitive and fixed to 0.01—that is, at 99% of judgment.

In the second case study, we investigate the role of explanatory variables in LMDJP-ARX by using Federal Reserve Bank of Philadelphia survey data on Real GDP for the US economy from 1970:Q1 to 2020:Q4; specifically, we consider all available measurements (real data, nowcasts, and forecasts up to one year ahead);

as explanatory variables, we consider production, consumption, money velocity, and house prices—-all of which are downloadable by FRED. Table 4 shows that the generated estimates and tests are qualitatively quite comparable to Nelson-Plosser data.

The third case study investigate the performances of the LMDJP-AR in macro-finance using the Robert Shiller's online dataset at `http://www.econ.yale.edu/~shiller/data.htm`. This is composed of 19 variables on US economy (prices, dividends, interest rates, inflation and related transformations) collected at yearly frequency and ranging from various starting dates up to 1970, see Shiller (1992).[1]. Also in this case, that over acceptation of the null is still more evident, as illustrated in Table 5.

The erroneous over-rejection of the null hypothesis for the LM-first strategy is clear in all the three illustrations. Two orders of economic arguments can explain this fact. The first relates to the mechanism design of the JP: namely, the reality's neutrality in the game between FP and FU does not imply that revealed data are unaffected by JB, but rather that the reality is unconcerned about who wins the game—or, in other words, that it lacks its own utility function. The dominance of FP over FU or vice versa can be viewed as a test of FU's eventual utility gain: if she has a capital and this capital does not increase as a function of a (significant) spread between FP quotation and reality announcement—measured by a divergence function between the utility of each agent having it—then FP effectively wins this game. Vovk and Shafer (2005) contains the theory for such a strategic perspective (and testing) of the JP. According to Zanetti Chini (2023b), professional forecasters at some of the most influential central banks and the official forecasts of these last exhibit strategic judgment. Thus, we may conclude that, in the long run, strategic conduct based on judgment tends to exaggerate the process' persistence; this final trait is frequently confused with LM, although it is really the natural outcome of a recurrent game. The second argument corresponds to the Bullard *et al.* (2008)'s

[1]A monthly equivalent dataset is also available and the results of the application on it are available under request

notion of near-exuberance equilibria. According to these authors, these exist when (i) agents face rational expectations equilibrium with limited information; (ii) each agent makes a rational individual decision to include the judgment in her forecasting model, given that all other agents are using it; and (iii) the forecasting model is stable and converges to rational expectations under recursive estimation schemes. Our LMDJP architecture fulfills all three of these assumptions.

4 Conclusions

Our simulated and actual data demonstrate that judgment and long memory are interconnected. The Long-Memory Dynamic Judgmental Protocol provided here is the first in the literature to parameterize this relationship. This is adaptable and may replicate a variety of contemporary features in this strand of literature. However, the evidence of spurious long memory with judgment is strong. More in detail, when building an econometric model, *ex-ante* estimate of long memory parameter is associated to an over-rejection of the null of spurious long memory. Thus, we advocate identifying judgmental bias prior to measuring long memory. Further research on the topic is necessary, in particular to design a feasible and correct statistical detection of long memory with strategic judgment.

References

Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Statistics for long-memory processes: Probabilistic properties and statistical methods*. Springer.

Bullard, J., Evans, G. and Honkapohja, S. (2008). Monetary policy, judgment, and near-rational exuberance. *American Economic Review*, **98** (3), 1163–77.

Chan, N. and Palma, W. (1998). State space modeling of long-memory processes. *The Annals of Statistics*, **26** (2), 719–740.

Chevillon, G. and Mavroeidis, S. (2017). Learning can generate long memory. *Journal of econometrics*, **198** (1), 1–9.

Ferrari, D. and Yang, Y. (2010). Maximum Lq-likelihood. *The Annals of Statistics*, **38** (2), 753–783.

Harvey, A. and Phillips, G. D. (1979). Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika*, **66** (1), 49–58.

Hassler, U. (2018). *Time series analysis with long memory in view*. John Wiley & Sons.

Manganelli, S. (2009). Forecasting with judgment. *Journal of Business & Economic Statistics*, **27** (4), 553–563.

Nelson, C. R. and Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of monetary economics*, **10** (2), 139–162.

Robinson, P. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, pp. 1630–1661.

Robinson, P. M. (2003). *Time series with long memory*. Advanced Texts in Econometrics.

Schennach, S. M. (2018). Long memory via networking. *Econometrica*, **86** (6), 2221–2248.

Shiller, R. J. (1992). *Market volatility*. MIT press.

Shimotsu, K. (2010). Exact local whittle estimation of fractional integration with unknown mean and time trend. *Econometric Theory*, **26** (2), 501–540.

— and Phillips, P. C. (2005). Exact local whittle estimation of fractional integration. *The Annals of statistics*, **33** (4), 1890–1933.

— and — (2006). Local whittle estimation of fractional integration and some of its variants. *Journal of Econometrics*, **130** (2), 209–233.

Svensson, L. E. (2005). Monetary policy with judgment: Forecast targeting. *International Journal of Central Banking*, **1** (1), 1–54.

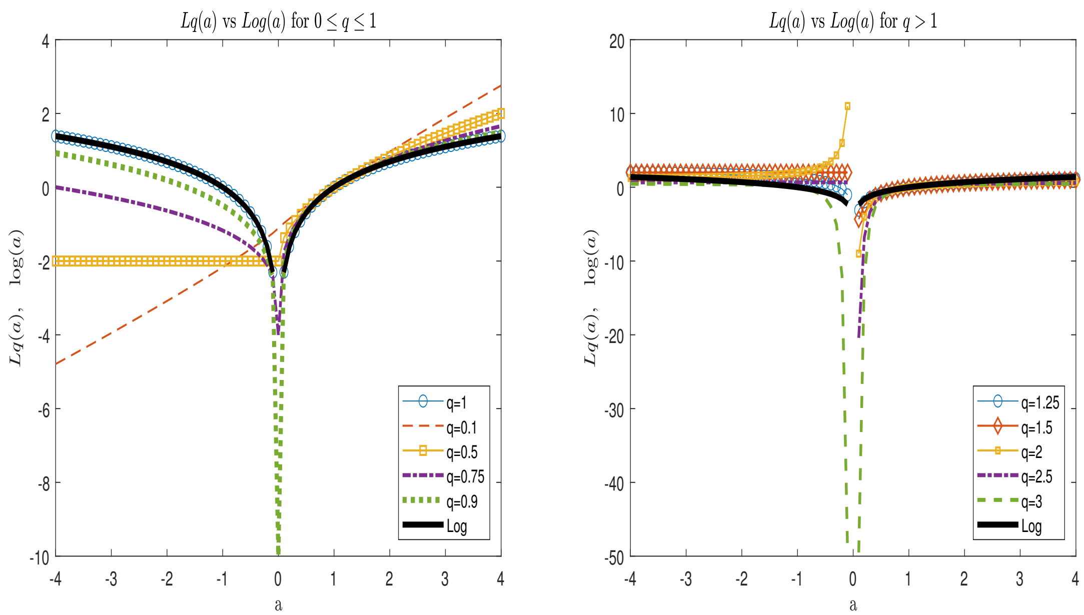
Vovk, V. and Shafer, G. (2005). Good randomized sequential probability forecasting is always possible. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67** (5), 747–763.

Zanetti Chini, E. (2023a). Can we estimate macroforecasters' misbehavior? *Journal of Economic Dynamics and Control*, **accepted**.

— (2023b). The strategy of professional forecasters. Department of Economics WP. 19, University of Bergamo.

Figure 1: The Deformed Logarithm function

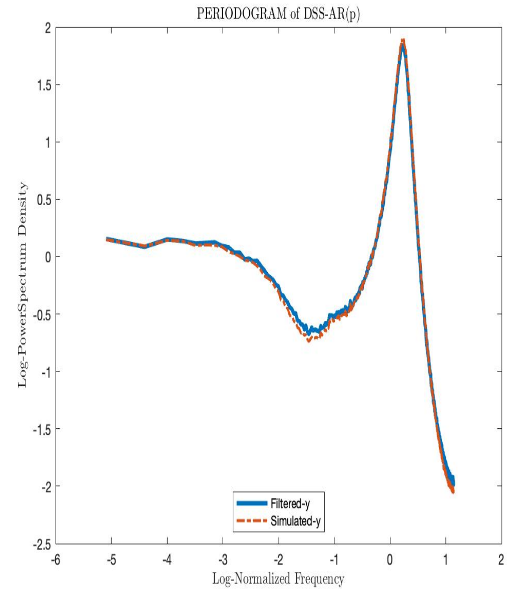
(a) Case 1: 0<q<1 **(b)** Case 2: q>1



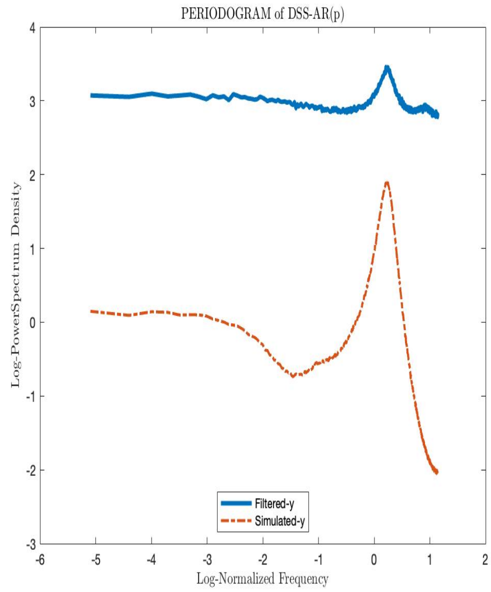
NOTE: This figure displays the Lq-function applied to the numerical sequence $a = [-4; 4]$ for different values of q and compares it with the natural logarithm function.

Figure 2: MonteCarlo average of 1-order DJP-AR(X) Power spectra under different q_0 and assuming d_0=0

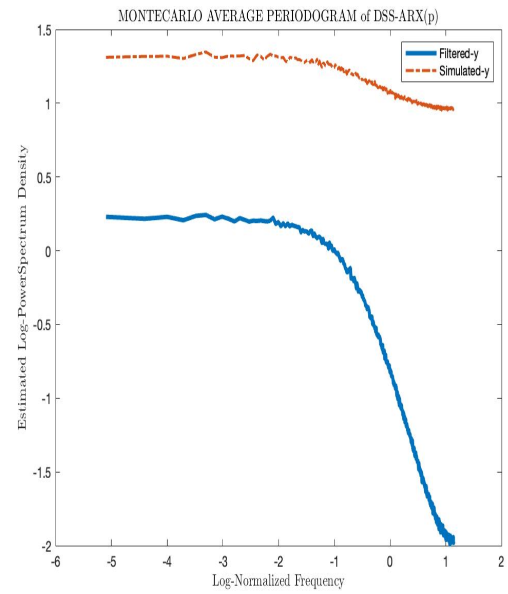
(a) AR(1), q=1



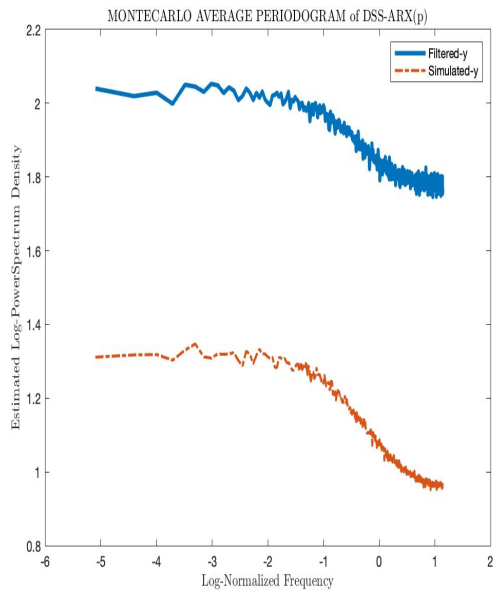
(b) AR(1), q=0.1



(c) ARX(1), q=1



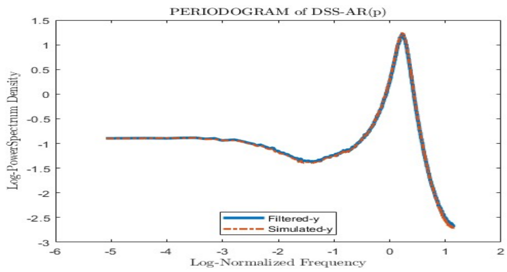
(d) ARX(1), q=0.1



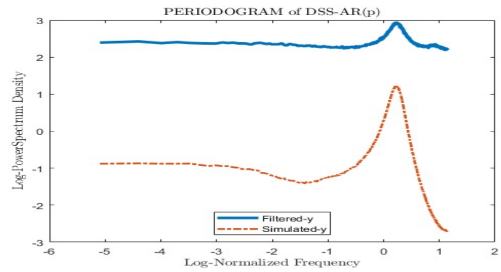
NOTE: This figure displays the average of a Monte Carlo simulated periodograms of the DJP-AR(1) and ARX(1) (in higher and lower panel, respectively) for $q_0 = 1$ (left panels) and $q_0 = 0.1$ (right panels) using the DGP illustrated in Section 3.

Figure 3: MonteCarlo average of DJP-AR(X)(1) Power spectra under different q_0 and d_0.

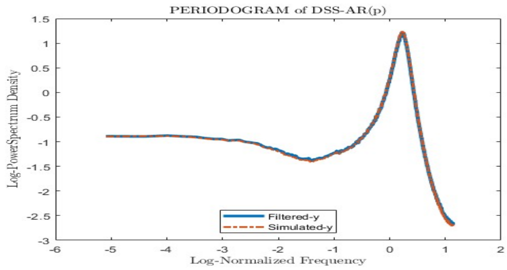
(a) AR(1), d=1, q=1



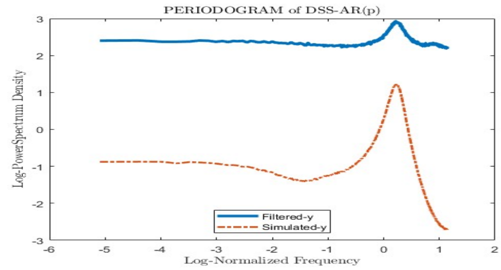
(b) AR(1), d=1, q=0.1



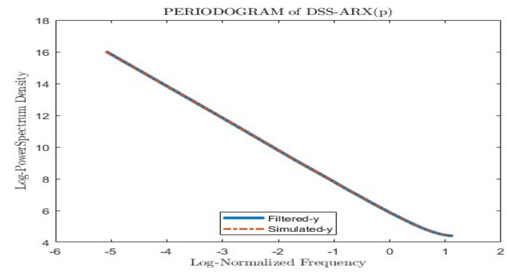
(c) AR(1), d=0.1, q=1



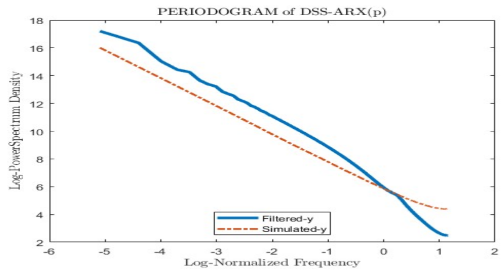
(d) AR(1), d=0.1, q=0.1



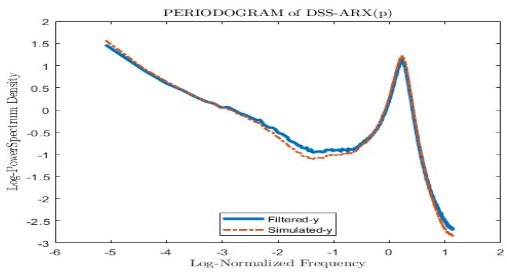
(e) ARX(1), d=1, q=1



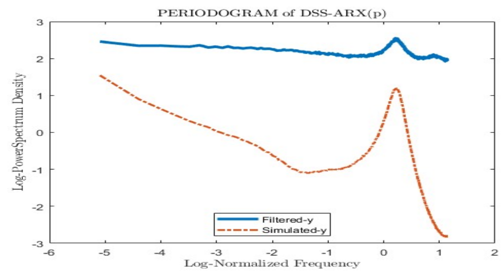
(f) ARX(1), d=1, q=0.1



(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 3.

Table 1: 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 $\hat{d}$				Empirical Power			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.2067	0.7428	0.7536	0.7742	0.2918	0.9050	0.9120	0.1613	0.1155
0.50	0	0.5678	1.0843	1.0912	1.0912	0.2918	0.9085	0.9240	0.2041	0.1341
1.00		0.7981	1.1043	1.1373	1.1392	0.2918	0.9192	0.9273	0.2144	0.1354
0.10		0.3054	0.5531	0.5538	0.5543	0.2918	0.9220	0.9320	0.2052	0.1318
0.50	0.5	0.6682	1.0866	1.1110	1.1115	0.2918	0.9481	0.9640	0.2342	0.1318
1.00		0.6980	1.1012	1.1240	1.1223	0.2918	1.0000	1.0000	0.2452	0.1377
0.10		0.3054	0.7495	0.7803	0.7808	0.2918	1.0000	1.0000	0.1733	0.1294
0.50	1	0.6681	0.9172	0.9150	0.9151	0.2918	1.0000	1.0000	0.1929	0.1304
1.00		0.7819	1.9170	1.9359	1.9361	0.2918	1.0000	1.0000	0.2344	0.1354
LMDJP-ARX(1), "Judgment First",										
0.10		0.2051	0.7594	0.7801	0.7847	0.2918	. 0.9082	0.9173	0.1813	0.1106
0.50	0	0.7678	1.0107	1.0137	1.0142	0.9126	0.9121	0.9288	0.1998	0.1109
1.00		0.6979	0.9464	0.9440	0.9441	0.2918	1.0000	1.0000	0.1613	0.1155
0.10		0.2501	0.7531	0.7820	0.7822	0.2918	0.9420	0.9320	0.1644	0.1324
0.50	0.5	0.6520	0.9310	0.9546	0.9370	0.2918	1.0000	1.0000	0.1834	0.1334
1.00		0.7197	1.1055	1.1241	1.1230	0.2918	1.0000	1.0000	0.2129	0.1354
0.10		0.2458	0.7531	0.7820	0.7822	0.2918	1.0000	1.0000	0.1582	0.1139
0.50	1	0.5702	1.0972	1.1130	1.1163	0.2918	1.0000	1.0000	0.1531	0.1177
1.00		0.7928	1.1121	1.1375	1.1377	0.2918	1.0000	1.0000	0.1608	0.1192
LMDJP-AR(1), "LM First"										
q_0	d_0	Mean $\hat{q}$	Mean $\hat{d}$				Empirical Power			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.3945	0.1128	0.1236	0.1342	0.1309	0.1788	0.2258	0.3944	0.2133
0.50	0	0.5489	0.5522	0.9121	0.9121	0.5539	0.1920	0.2595	0.4063	0.2195
1.00		0.5924	0.6640	0.7360	0.7433	0.6600	0.2244	0.2681	0.4566	0.2240
0.10		0.4222	0.6030	0.6329	0.6425	0.6427	0.2899	0.2756	0.2803	0.1541
0.50	0.5	0.6382	0.6058	0.6374	0.6466	0.6404	0.2634	0.2763	0.2889	0.1582
1.00		0.6980	0.5931	0.6221	0.6328	0.6321	0.2031	0.2860	0.2733	0.1677
0.10		0.4353	1.0954	1.1204	1.1209	0.4209	0.4995	0.4820	0.3672	0.1959
0.50	1	0.6455	1.1133	1.1321	1.1330	0.1448	0.5431	0.4933	0.3892	0.2130
1.00		0.6702	1.1055	1.1245	1.1203	0.1218	0.5875	0.4964	0.3850	0.2304
LMDJP-ARX(1) "LM First"										
0.10		0.4541	1.1244	1.1360	1.1378	0.3379	0.2024	0.2530	0.3181	0.1494
0.50	0	0.6564	1.1347	1.1427	1.1426	1.1451	0.1951	0.2509	0.3809	0.2583
1.00		0.6869	0.1264	0.1240	0.1241	0.1278	0.1990	0.2550	0.3903	0.2689
0.10		0.4534	0.6140	0.6422	0.6480	0.6542	0.4053	0.4120	0.5105	0.1590
0.50	0.5	0.6045	0.6212	0.6340	0.6355	0.6439	0.2093	0.4383	0.5927	0.1666
1.00		0.6990	0.6030	0.6351	0.6439	0.6480	0.2260	0.4455	0.6066	0.1646
0.10		0.4553	1.1024	1.1215	1.1237	1.1242	0.5995	0.3335	0.5845	0.2095
0.50	1	0.6088	1.1031	1.1232	1.1261	1.1262	0.6100	0.4154	0.5892	0.2021
1.00		0.6952	1.1024	1.1240	1.1252	1.1255	0.5875	0.4640	0.6250	0.2020

NOTES: This table reports the MonteCarlo average estimates of the strategic judgmental parameter (column 3) and 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 3.

Table 2: Simulation results for the Long-Memory estimators and tests in DJP for T=500.

LMDJP-AR(1), "Judgment First"										
q_0	d_0	Mean $\hat{q}$	Mean $\hat{d}$				Empirical Power			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.5327	0.0363	0.0356	0.3561	0.2918	0.9956	0.9995	0.3054	0.2546
0.50	0	0.7589	0.0353	0.0345	0.0345	0.2918	1.0000	1.0000	0.2946	0.2556
1.00		0.8959	0.0344	0.0338	0.0338	0.2918	1.0000	1.0000	0.2709	0.2556
0.10		0.5334	0.5284	0.5352	0.5385	0.2918	0.9953	0.9945	0.2824	0.2566
0.50	0.5	0.7508	0.0527	0.0534	0.0537	0.2918	1.0000	1.0000	0.2813	0.2555
1.00		0.9963	0.5296	0.5363	0.5397	0.2918	1.0000	1.0000	0.2870	0.2556
0.10		0.4228	1.0334	1.0356	1.0357	0.2918	1.0000	1.0000	0.3445	0.2554
0.50	1	0.7508	1.0320	1.0345	1.0344	0.2918	1.0000	1.0000	0.3045	0.2556
1.00		0.9969	1.0310	1.0337	1.0336	0.2918	1.0000	1.0000	0.3082	0.2556
LMDJP-ARX(1), "Judgment First"										
0.10		0.6463	0.6773	0.6867	0.6867	0.2918	0.9952	0.9963	0.3454	0.2576
0.50	0	0.7801	1.0125	1.0195	1.0195	0.2918	1.0000	1.0000	0.3214	0.2578
1.00		0.8991	1.0366	1.0366	1.0393	0.2918	1.0000	1.0000	0.3345	0.2577
0.10		0.6502	0.6754	0.6818	0.6837	0.2918	1.0000	1.0000	0.3810	0.2579
0.50	0.5	0.7806	1.0164	1.0193	1.0198	0.2918	1.0000	1.0000	0.4069	0.2767
1.00	.	0.8991	1.0296	1.0345	1.0377	0.2918	1.0000	1.0000	0.3456	0.2871
0.10		0.6553	0.6846	0.6932	0.6955	0.2918	1.0000	1.0000	0.3856	0.2556
0.50	1	0.7810	1.0123	1.0158	1.0156	0.2918	1.0000	1.0000	0.4056	0.2556
1.00		0.8991	1.0276	1.0304	1.0303	0.2918	1.0000	1.0000	0.3934	0.2567

LMDJP-AR(1), "LM First"										
q_0	d_0	Mean $\hat{q}$	Mean $\hat{d}$				Empirical Power			
			LW	ELW	ELW2	2SFELW	LW	ELW	ELW2	2SFELW
0.10		0.5329	0.0350	0.0344	0.0344	0.0344	0.3838	0.2173	0.2917	0.1135
0.50	0	0.9509	0.0369	0.0364	0.0364	0.0364	0.3897	0.4178	0.2918	0.1136
1.00		0.9962	0.0358	0.0350	0.0350	0.0350	0.3945	0.4114	0.2923	0.1233
0.10		0.5336	0.5295	0.5295	0.5362	0.5393	0.3801	0.3145	0.2777	0.1145
0.50	0.5	0.9502	0.5273	0.5341	0.5374	0.5374	0.3836	0.4134	0.3845	0.1129
1.00		0.9962	0.5284	0.5352	0.5386	0.5387	0.3809	0.4125	0.3956	0.1159
0.10		0.5334	1.0330	1.0356	1.0356	1.0356	0.4737	0.3091	0.2788	0.1145
0.50	1	0.9494	1.0308	1.0334	1.0334	1.0334	0.4612	0.3960	0.2747	0.1104
1.00		0.9964	1.03134	1.0332	1.0338	1.0338	0.4834	0.4033	0.2824	0.1145
LMDJP-ARX(1) with $\hat{d}$, "LM First"										
0.10		0.6514	0.6718	0.6784	0.6803	0.2918	0.9654	0.9853	0.5088	0.2570
0.50	0	0.9809	1.0093	1.0124	1.0122	0.2918	1.0000	1.0000	0.5945	0.2576
1.00		0.9991	1.0309	1.0334	1.0335	0.2918	1.0000	1.0000	0.5133	0.2583
0.10		0.6532	0.5239	0.5303	0.5342	0.5347	0.3782	0.4134	0.5779	0.2230
0.50	0.5	0.9810	0.5243	0.5310	0.5481	0.5404	0.3043	0.4185	0.5240	0.2299
1.00		0.9991	0.5283	0.5331	0.5376	0.5376	0.3243	0.4385	0.5240	0.2599
0.10		0.6507	1.0324	1.0347	1.0347	1.0347	0.4725	0.4338	0.5883	0.2444
0.50	1	0.9810	1.0324	1.0343	1.0343	1.0342	0.4325	0.4755	0.5476	0.2224
1.00		0.9991	1.0334	1.0369	1.0373	1.0370	0.4720	05048	0.5755	0.2295

NOTES: This table reports the MonteCarlo average estimates of the strategic judgmental parameter (column 3) and 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 3.

Table 3: Application of the LMDJP-AR model to Nelson-Plosser's data.

Series	T	$\hat{p}$	DJP-AR(p), "Judgment First" Mean $\hat{q}$	Mean $\hat{d}$ LW	ELW	ELW2	2SFELW	DJP-AR(p), "LM First" Mean $\hat{q}$	Mean $\hat{d}$ LW	ELW	ELW2	2SFELW
Real GNP	62	1	0.0000	1.1132	1.2469	1.1993	1.9923	0.0000	1.6374	1.3469	1.2784	1.1931
Nominal GNP	62	1	0.0000	1.1872	1.6768	1.6973	1.6979	0.0000	0.9972	1.7769	1.8091	1.6984
Real per capita GNP	62	1	0.0000	0.8357	0.9399	0.9151	0.9157	0.0000	0.6432	1.0345	0.9924	0.9151
Industrial production	111	2	0.0262	1.0093	1.2547	1.1892	1.1897	0.0262	1.7245	1.3821	1.3136	1.2137
Employment	81	2	0.0000	0.5712	0.6622	0.6547	0.6692	0.0000	0.6679	0.7622	0.7682	0.7693
Unemployment rate	81	2	0.0650	0.7397	0.7661	0.8013	0.8014	0.0650	1.0067	0.8013	0.6413	0.6417
GNP deflator	82	2	0.0000	1.1708	1.1678	1.3199	1.3198	0.0000	1.4080	1.2673	1.3199	1.3198
CPI	111	2	0.0000	0.9655	1.3194	1.2395	1.2345	0.0000	1.2696	1.4194	1.3261	1.2345
Nominal wage	71	2	0.0000	1.2684	1.3805	1.5372	1.5372	0.0000	1.6825	1.4805	1.6338	1.5372
Real wage	71	2	0.0000	0.9769	1.0162	1.0352	1.0354	0.0000	1.5004	1.1162	1.1118	1.0357
Money stock	71	2	0.0000	1.2025	1.7048	1.7839	1.7839	0.0000	1.4553	1.8048	1.9233	1.7839
Velocity of money	102	2	0.0000	0.9926	1.0184	1.0552	1.0552	0.0000	0.0692	1.1182	1.1191	1.0552
Bond yield	71	2	0.0000	0.7724	1.6559	1.6605	1.6605	0.0000	0.0520	1.7559	1.7676	1.6608
Stock prices	100	2	0.0292	1.1489	1.2170	1.1167	1.1165	0.0262	1.4698	1.3451	1.2453	1.4250
				Test for no Long Memory (p-values)								
Real GNP				0.0583	0.0666	0.0636	0.0123		0.0651	0.0782	0.0685	0.0685
Nominal GNP				0.0629	0.0929	0.0942	0.0123		0.0623	0.0989	0.1003	0.0942
Real per capita GNP				0.0415	0.0477	0.0626	0.0123		0.0460	0.0538	0.0509	0.0462
Industrial production				0.0538	0.0643	0.0601	0.0123		0.0564	0.0732	0.0680	0.0622
Employment				0.0293	0.0357	0.0356	0.0123		0.0340	0.0415	0.0412	0.0356
Unemployment rate				0.0368	0.0392	0.0392	0.0123		0.0396	0.0423	0.0422	0.0363
GNP deflator				0.0609	0.0638	0.0698	0.0123		0.0652	0.0679	0.0711	0.0762
CPI				0.0483	0.0712	0.0631	0.0123		0.0512	0.0772	0.0715	0.0658
Nominal wage				0.0679	0.0749	0.0842	0.0123		0.0691	0.0811	0.0904	0.0845
Real wage				0.0499	0.0524	0.0535	0.0123		0.0556	0.0585	0.0582	0.0535
Money stock				0.0603	0.0941	0.0989	0.0123		0.0590	0.0999	0.1068	0.0982
Velocity of money				0.0502	0.0524	0.0547	0.0123		0.0562	0.0586	0.0587	0.0547
Bond yield				0.0378	0.0917	0.0923	0.0123		0.0412	0.0972	0.0920	0.0983
Stock prices				0.0605	0.0642	0.0582	0.0123		0.0608	0.0724	0.0662	0.0601

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" modelling strategy for the four estimator here considered (and exposed in Appendix), while the equivalent estimates for "LM First" is reported in the remaining columns. The lower panel displays the p-values of the test corresponding to (3), still for each of the four estimators. In both the modelling strategies we assume $q_0 = 0.1$

Table 4: Application of the ARX-LMDJP model to SPF data.

Series	DJS-AR(p), "Judgment First" Mean $\hat{q}$	Mean $\hat{d}$ LW	ELW	ELW2	2SFELW	DJS-AR(p), "LM First" Mean $\hat{d}$ LW	ELW	ELW2	2SFELW
Real data	. 0.8747	0.8105	0.8292	0.8595	0.2918	1.1953	0.0544	0.0544	0.4522
Nowcasts	0.0831	0.1665	0.2052	0.2056	0.2918	0.2222	0.2226	0.1225	0.2245
1-q-ahead	0.7519	1.1051	1.11450	1.1415	0.2918	0.4533	0.4932	0.3932	0.5027
2-q-ahead	0.7245	1.1214	1.1581	1.1608	0.2918	0.4495	0.5336	0.5359	0.4336
3-q-ahead	0.6951	1.1570	1.1700	1.1806	0.2918	0.4767	0.5749	0.6055	0.5742
1-y-ahead	0.5652	1.2157	1.0442	1.0823	0.6951	0.2889	0.5791	0.7470	0.7017
		Test for no Long Memory (p-values)							
Real data	.	0.0401	0.0415	0.0429	0.0123	0.0020	0.0021	0.0021	0.0021
Nowcasts		0.0067	0.0084	0.0084	0.0123	0.0074	0.0091	0.0092	0.0048
1-q-ahead		0.0578	0.0603	0.0601	0.0123	0.0186	0.0223	0.0224	0.0172
2-q-ahead		0.0590	0.0611	0.0612	0.0123	0.0214	0.0244	0.0246	0.0193
3-q-ahead		0.0610	0.0618	0.0625	0.0123	0.0249	0.0267	0.0283	0.0213
1-y-ahead		0.0643	0.0542	0.0563	0.0123	0.0361	0.0264	0.0364	0.0360

NOTES: This table reports the result of the application of the LMDJP-ARX model to SPF data on US Real GDP for several forecasting horizons. 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" modelling strategy for the four estimator here considered (exposed in Appendix), while the equivalent estimates for "LM First" is reported in the remaining columns. The lower panel displays the p-values of the test corresponding to (3), still for each of the four estimators. In both the modelling strategies we assume $q_0 = 0.1$ and an autoregressive order $p = 2$.

Table 5: Application of the LMDJP-AR model to Shiller's annual data.

			DJS-AR(p), "Judgment First"					DJS-AR(p), "LM First"				
Series	T	$\hat{p}$	Mean $\hat{q}$	Mean $\hat{d}$				Mean $\hat{q}$	Mean $\hat{d}$			
				LW	ELW	ELW2	2SFELW		LW	ELW	ELW2	2SFELW
P	146	3	0.0068	0.7535	1.1358	1.0591	0.2918	0.0068	1.0522	1.2289	1.1564	1.0520
D	145	3	0.0200	0.7326	1.1776	1.1861	0.2918	0.0200	1.3029	1.3083	1.2082	1.3083
E	143	3	0.0070	0.6107	0.7908	0.8217	0.2918	0.0070	0.8838	0.9081	0.8153	0.9081
R	141	3	0.0286	0.5652	0.6489	0.6654	0.2918	0.0286	0.7203	0.7255	0.6453	0.7245
RLong	142	3	1.0000	0.8627	0.9995	1.0657	0.2918	0.0000	-6.5407	1.1745	1.0861	1.1888
CPI	142	3	1.0000	1.4005	1.3178	1.6206	0.2918	0.0000	1.4005	1.3178	1.6206	0.2918
RealR	142	3	1.0000	1.1692	1.1885	1.1919	0.2918	1.0000	2.3178	2.2752	1.6206	1.6206
C	121	2	0.0095	1.3180	1.3833	1.2737	0.2918	0.0099	2.4622	2.4660	1.5564	2.4660
RealP	144	3	0.0144	0.8543	0.7492	0.8004	0.2918	0.0144	1.7512	0.7872	1.7511	0.7872
P∗	139	3	0.0084	1.1498	1.1498	1.9789	0.2918	0.0084	2.5538	2.5007	1.9789	2.5007
P × r	139	3	0.0075	1.2015	1.3062	1.3187	0.2918	0.0075	2.3062	2.3116	1.3187	2.3116
P × C	120	2	0.0075	1.1708	1.3618	1.3641	0.2918	0.0075	2.1741	2.1388	1.0925	2.1388
Real D	142	3	0.0146	0.6388	0.5684	0.6702	0.2918	0.0146	1.5537	1.5530	0.6617	1.5530
Return	142	3	0.9994	0.8178	0.8410	0.8448	0.2918	0.9994	2.0041	0.8417	0.8454	0.8582
$\ln(1+Return)$	142	3	1.0000	0.8180	0.8423	0.8461	0.2918	1.0000	1.6107	0.8423	0.8461	0.8461
RealE	141	3	0.9749	1.3742	1.2898	1.2958	0.2918	0.9749	1.3149	1.3206	0.3149	1.3205
P/E	141	3	0.9878	1.3754	1.3870	1.3852	0.2918	0.9878	1.3992	1.3974	0.3992	1.3974
E10	133	3	0.0077	1.3119	1.0905	1.2254	0.2918	0.0077	2.0905	2.1020	1.2254	2.1020
P/E10	132	3	0.9629	1.7005	1.7452	1.7426	0.2918	0.9629	1.7823	1.7792	0.7978	1.7792
				Test for no Long Memory (p-values)								
P				0.0341	0.0225	0.0563	0.0123		0.0631	0.0561	0.0569	0.0581
D				0.0355	0.0384	0.0382	0.0123		0.0563	0.0552	0.0540	0.0653
E				0.0483	0.0478	0.0514	0.0123		0.0662	0.0679	0.0585	0.0664
R				0.0471	0.0456	0.0450	0.0123		0.0593	0.0635	0.0644	0.0531
RLong				0.0256	0.0359	0.0395	0.0123		0.0483	0.0512	0.0532	0.0485
CPI				0.0338	0.0377	0.0490	0.0123		0.0753	0.0740	0.0734	0.0738
RealR				0.0509	0.0599	0.0480	0.0123		0.0807	0.0773	0.0822	0.0834
C				0.0518	0.0530	0.0500	0.0123		0.0774	0.0730	0.0981	0.0567
RealP				0.0496	0.0583	0.0579	0.0123		0.0692	0.0802	0.0775	0.0882
P∗				0.0591	0.0694	0.0734	0.0123		0.0840	0.0899	0.0924	0.0740
P × r				0.0389	0.0325	0.0422	0.0123		0.0504	0.0562	0.0559	0.0485
P × C				0.0560	0.0604	0.0628	0.0123		0.0863	0.0935	0.0999	0.0705
Real D				0.0503	0.0562	0.0566	0.0123		0.0615	0.0638	0.0664	0.0662
Return				0.0488	0.0563	0.0542	0.0123		0.0690	0.0722	0.0713	0.0583
$\ln(1+Return)$				0.0267	0.0308	0.0402	0.0123		0.0482	0.0499	0.0553	0.0425
RealE				0.0372	0.0432	0.0562	0.0123		0.0553	0.0602	0.0645	0.0483
P/E				0.0456	0.0553	0.0593	0.0123		0.0643	0.0705	0.0739	0.0760
E10				0.0600	0.0663	0.0634	0.0123		0.0853	0.0886	0.0734	0.0604
P/E10				0.0583	0.0666	0.0605	0.0123		0.0793	0.0814	0.0853	0.0555

NOTES: This table reports the result of the application of the LMDJP-AR model to Shiller's annual data on US economic variables. In the upper panel, the first three columns describe the variable labels, 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" modelling strategy for the four estimator here considered (and exposed in Appendix), while the equivalent estimates for "LM First" is reported in the remaining columns. The lower panel displays the p-values of the test corresponding to (3), still for each of the four estimators. In both the modelling strategies we assume $q_0 = 0.1$

APPENDIX

A Estimation

A.1 Estimating Long Memory

The spectral density of the process (2) in Main Text can be represented as

$$f_{\Phi}(\lambda) \sim G|\lambda|^{-2d} \quad \text{as} \quad \lambda \to 0 \tag{6}$$

where G corresponds to the spectral density of an AR(p) process and can be estimated via Local Whittle Likelihood (LW):

$$\begin{gathered}\sum_{j=1}^{m} \log f_u(\lambda) + \sum_{j=1}^{m} \frac{I_u(\lambda_j)}{f_u(\lambda_j)}, \\ I(\lambda) = |w_(\lambda)^2|, \quad w(\lambda) = (2\pi n)^{-1/2} \sum_{x=1}^{n} y_t e^{it\lambda}, \quad \lambda_j = 2\pi_j/n\end{gathered} \tag{7}$$

where $w()$, $I()$ and $\lambda()$are the discrete Fourier transform, the periodogram, and the frequency of y_t, respectively. We consider four different estimators of d: the Local Whittle Likelihood estimator (LW); the Exact Whittle Likelihood (ELW) the Modified Whittle Likelihood (ELW2) and Two-Step Feasible Whittle Likelihood (2SFEWL); for details, see Robinson (1995); Shimotsu and Phillips (2005, 2006); Shimotsu (2010).

A.2 The Deformed Likelihood

Let $y_1, \ldots, y_T$ be an i.i.d. sample from $p(y_i, \theta_0)$, $\theta_0 \in \Theta$, where all the elements of that vector Θ are part of the LMDJP defined in Section 2 of Main Document. Then the maximum Lq-estimator (MLqE) of θ_0 is

$$\hat{\theta}_T \doteq \max_{\tilde{\theta} \in \Theta} \sum_{t=1}^{T} L_q[p(y_t; \theta)], \quad q > 0, \tag{8}$$

where $Lq(x)$ is a Lq-transform. Moreover, if $p(\cdot)$ is Gaussian, its Deformed Likelihood is:

$$\mathcal{L}^q(\theta; y_t) = -0.5 * [q_0 T \log(2\pi) + (\log p(y_t) + u_t^2)^q], \tag{9}$$

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. Moreover, (8) is the result of the maximization of:

$$\sum_{t=1}^{T} w_t U(y_t, \theta) = 0 \tag{10}$$

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.

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. Typically, outliers are associated to very small weights. According to Figure 4, a small deviation from 1 is sufficient to modify drastically the log-likelihood. However, such a contamination is not uniform. For example, according to the analytic results by Ferrari and Yang (2010), if $p(\cdot, \cdot)$ is Gaussian, then the estimated mean of $\hat{\theta}$ does not depends of q.

The parameter q is a point measure of judgmental bias in the estimated model due to FP (or FU) *singularly.*

Finally, it is possible to prove that the q-Entropy coincides with the H-function corresponding to the DJP.

B Proof of Proposition 1

Part (a)

Let assume the state-space form for y_t as the following system of equations, named *Dynamic Judgmental System* (DJP, henceforth):

$$
\begin{aligned}
y_t \quad &= Z_t\alpha + X_t\beta + G_t\epsilon_t, \qquad \epsilon_t \sim iid(0, \sigma^2_\epsilon) \\
\alpha_{t+1} &= T_t\alpha_t + W\beta + H_t\eta_t, \quad \eta_t \sim iid(0, \sigma^2_\eta); \\
Z &= [1,\ \mathbf{0}_{m-1}]; \qquad X = [1,\ \ \mathbf{0}_{k-1}]; \qquad H = \boldsymbol{\phi}' \\
G &= \begin{bmatrix} \boldsymbol{I}_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix}; \qquad T = \begin{bmatrix} \boldsymbol{\phi}' \\ G \end{bmatrix}; \qquad W = [\boldsymbol{\phi}',\ G]',
\end{aligned} \tag{11}
$$

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:

$$
\begin{aligned}
\alpha_0 = [\mathbf{0}_m]; \quad \beta_0 = [\mathbf{0}_k]; \quad \boldsymbol{I}_{m^2} = \boldsymbol{I} \otimes [T,\ T]; \\
H^2 = HH'; \qquad\qquad vec(P) = \boldsymbol{I}_m^{-1} H^2.
\end{aligned} \tag{12}
$$

Then, we invoke the following standard assumptions:

Assumption 2. *(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* $t = 1 \dots n$.

Lemma 1. Suppose $\{y_t\}$ is a stationary ARFIMA(p,d,q) process. Under assumption A1 (i)—(iv), the finite state-space system (11) is stationary.

Proof. This is a re-proposition of Lemma 2.1 in Chan and Palma (1998) to which we refers for proof.

□

Remark 3. The system (11) and A1 (i)—(iv) are the same of Zanetti Chini (2023a)—that is, the presence of long memory does not affect the the state-space autoregressive form.

Now, we need to prove that our DJP is effectively an object that nests the q-Entropy and H-function defined in the Main Document. This is possible by the following.

Lemma 2. The q-Entropy is the H-function corresponding to the DJP.

Proof. Trivial if (i) remembering the definition of $H_q := E_f L_q\{g(X)\}$, where g and f are two density functions and (ii) that system (11) is invariant to the fact that long-memory parameter $d \neq 0$. See Ferrari and Yang (2010) and Zanetti Chini (2023a). □

Lemma 3. Consider the Deformed Logarithm in equation (1) in Main Document. Then, (i) if q approaches to 1, $H_q(\cdot, \cdot)$ nests the Shannon Entropy and $D_q(\cdot, \cdot)$ the Kullback-Leibler 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. □

After these preliminary results we can link the equation (11) to JP:

Lemma 4. (i) The DJP is never isomorphic to JP;
(ii) The DJP is isomorphic to (11).

Proof. It descends directly from (a) the part (ii) of Lemma 3 and (b) Lemma 4. Lets consider each of them separately.
(a) the part (ii) of Lemma 3 is necessary but not sufficient. If DJP were not isomorphic to the state-space form in equation 11, 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 4 is a sufficient, non-necessary condition for the next Proposition 1 to hold. Part (i) of Lemma 4 ensures a one-to-one relation among q-Entropy, a known D-function and the DJP; part (ii) of Lemma 3 ensures the one-to-one relationship among the true parameter θ_0 and the amount of the aggregate quote of judgment represented by q. □

In summary, the existence of a state-space form of a long-memory model is ensured by Lemma1; the feasibility of entropy-based method—and, specifically, the Lq-Likelihood—to measure JP is ensured by Lemmas 2 and 3, respectively. Finally by Lemma 4 proves the one-to-one relationship among (11) to be estimated. Hence, the combination of Lemmas 1–4 implies part (a) of Proposition 1 in Main Document.

Part (b)

The proof is divided in two cases:

CASE 1 (q=1). This nests the the classical Log-Likelihood, hence the proof is identical to the one in Chan and Palma (1998), Theorem 2.2.

CASE 2 (0<q<1). The proof of this case requires to verify if the mentioned results of (Chan and Palma, 1998) hold also for LqLikelihood.

Lemma 5. Under A1 (i)–(iv) and the MLqE in (8)–(10), we have that: (i) the DKF recursive equations are:

for 1,. . . , t, . . . , n,

$$v = y_t - Z\alpha_t - X\beta; \qquad F = ZPZ' + GG'; \tag{13}$$

$$C = PZ'/F; \qquad t = v/\sqrt{q_0 F^{(q_0-1)}}; \tag{14}$$

$$v_t = \alpha + C\sqrt{F}h; \qquad P_t = \alpha_t + Cq_0FC'(h/t); \qquad Q = HG'/F; \tag{15}$$

$$P_{t+1|t} = TV_tT' + HH' - (QFQ' + QFC'T' + TCFQ')(h/t); \tag{16}$$

$$\alpha_{t+1|t} = Tv_t + W\beta + Q\sqrt{(q_0F)}(h/t); \tag{17}$$

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

$$h = \begin{cases} t & if \quad (t < t^*) \\ h(t,a,b) & otherwise, \end{cases} \tag{18}$$

and

$$h(t,a,b) = \begin{cases} t; & if \quad |t| \le a \\ \frac{a}{b-a}(b-t) & if \quad a < t \le b, \\ \frac{a}{b-a}(b+t) & if \quad -b < t \le t-a, \\ 0 & if \quad |t| \ge b. \end{cases} \tag{19}$$

(ii) The Deformed Likelihood and filtered y_t are:

$$L = log(F) + log(q_0); \qquad S = v^2/F; \tag{20}$$

$$y_f = Zv_t + GG'\sqrt{q_0 F}h/(q_0 F); \qquad h_t = h/t; \tag{21}$$

$$\alpha_{t+1} = a; \qquad \Sigma_{t+1} = diag(P) \tag{22}$$

where $I = v$ and $\sigma_I^2 = F$.

(iii) The Deformed Likelihood's weights and the averaged LogLikelihood, concentrated LogLikelihood, LqLikelihood and estimated Lq-Likelihood measures are:

$$w_t = 1/(L)^{q_0}\mathcal{G}_t \tag{23}$$

$$\mathcal{L} = -0.5[T\log(2\pi) + L + S]; \qquad \mathcal{L}c = -0.5[T(\log(2\pi\mathcal{S}) + 1] + L; \tag{24}$$

$$\mathcal{L}_q = 0.5((q_0 T\log(2\pi) + (L+S)^{q_0}); \qquad \hat{\mathcal{L}}_q = 0.5((\hat{q}T\log(2\pi) + (L+S)^{\hat{q}}). \tag{25}$$

Proof. CASE 1: q=1. By Lemma 3, the Lq-Likelihood function coincides with the standard Likelihood; thus, the DKF coincides with the (robust) KF (see also corollary 2 of Zanetti Chini (2023a)), to which the proof is delegated.

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) $t = v/\sqrt{q_0 F^{q_0 - 1}}$ in eq. (10)–(12); (b) q_0 in the second addend of (11).

Let start from (b). By Lemma 1 in Zanetti Chini (2023a), 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, as instead required in case of higher dimentions. 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 in Zanetti Chini (2023a), 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 equation (10), the optimal solution is the first derivative of $\sqrt{F^q}$.

(ii) Direct consequence of (i).

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

After having verified that MLqE is feasible in our state-space modelling, we need to demonstrate that (11) is valid in finite samples to ensure that the Chan and Palma (1998) results holds also in our case. This is possible by the following

Lemma 6. Let $\{y_n\}$ be a finite sample of an ARFIMA (p,d,q) process. If F_1 is the variance of the initial state α_1. Then, the computation of the likelihood depends only on the first n components of the Deformed Kalman equations.

Proof. Trivial if noticing how $\mathcal{L}$ in (19) works. Namely, it is formed by three parts: the first depends on the sample size; the second on L, in turn depending on F and q_0 (that is, a constant); the third depends on estimated residuals v and the same F. Thus, the issue is demonstrating that F and ν depends on the first n components. In turn, this is immediate by noticing that v and F are linked by the object t in (9) and that matrix P (necessary to have F) is a recursion on n. □

Remark 4. This last result points that DKF is feasible at least as the classical KF in Chan and Palma (1998).

C Proof of Proposition 2

Trivial if considering the equation of $\mathcal{L}_q$ in (20) is function of q via L and that the last element can be re-written by using spectral density as $L = \int_{-\pi}^{\pi} G(\lambda)e^{i\omega}, \omega = [-\pi, \pi]$.

References

Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Statistics for long-memory processes: Probabilistic properties and statistical methods*. Springer.

Bullard, J., Evans, G. and Honkapohja, S. (2008). Monetary policy, judgment, and near-rational exuberance. *American Economic Review*, **98** (3), 1163–77.

Chan, N. and Palma, W. (1998). State space modeling of long-memory processes. *The Annals of Statistics*, **26** (2), 719–740.

Chevillon, G. and Mavroeidis, S. (2017). Learning can generate long memory. *Journal of econometrics*, **198** (1), 1–9.

Ferrari, D. and Yang, Y. (2010). Maximum Lq-likelihood. *The Annals of Statistics*, **38** (2), 753–783.

Harvey, A. and Phillips, G. D. (1979). Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika*, **66** (1), 49–58.

Hassler, U. (2018). *Time series analysis with long memory in view*. John Wiley & Sons.

Manganelli, S. (2009). Forecasting with judgment. *Journal of Business & Economic Statistics*, **27** (4), 553–563.

Nelson, C. R. and Plosser, C. R. (1982). Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of monetary economics*, **10** (2), 139–162.

Robinson, P. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, pp. 1630–1661.

Robinson, P. M. (2003). *Time series with long memory*. Advanced Texts in Econometrics.

Schennach, S. M. (2018). Long memory via networking. *Econometrica*, **86** (6), 2221–2248.

Shiller, R. J. (1992). *Market volatility*. MIT press.

Shimotsu, K. (2010). Exact local whittle estimation of fractional integration with unknown mean and time trend. *Econometric Theory*, **26** (2), 501–540.

— and Phillips, P. C. (2005). Exact local whittle estimation of fractional integration. *The Annals of statistics*, **33** (4), 1890–1933.

— and — (2006). Local whittle estimation of fractional integration and some of its variants. *Journal of Econometrics*, **130** (2), 209–233.

Svensson, L. E. (2005). Monetary policy with judgment: Forecast targeting. *International Journal of Central Banking*, **1** (1), 1–54.

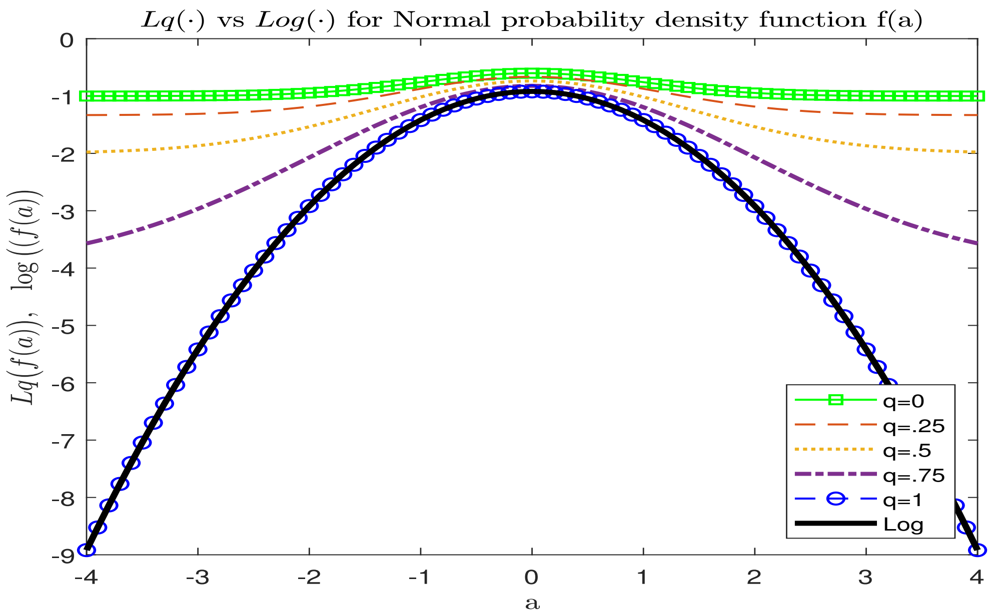
Vovk, V. and Shafer, G. (2005). Good randomized sequential probability forecasting is always possible. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67** (5), 747–763.

Zanetti Chini, E. (2023a). Can we estimate macroforecasters' misbehavior? *Journal of Economic Dynamics and Control*, **accepted**.

— (2023b). The strategy of professional forecasters. Department of Economics WP. 19, University of Bergamo.

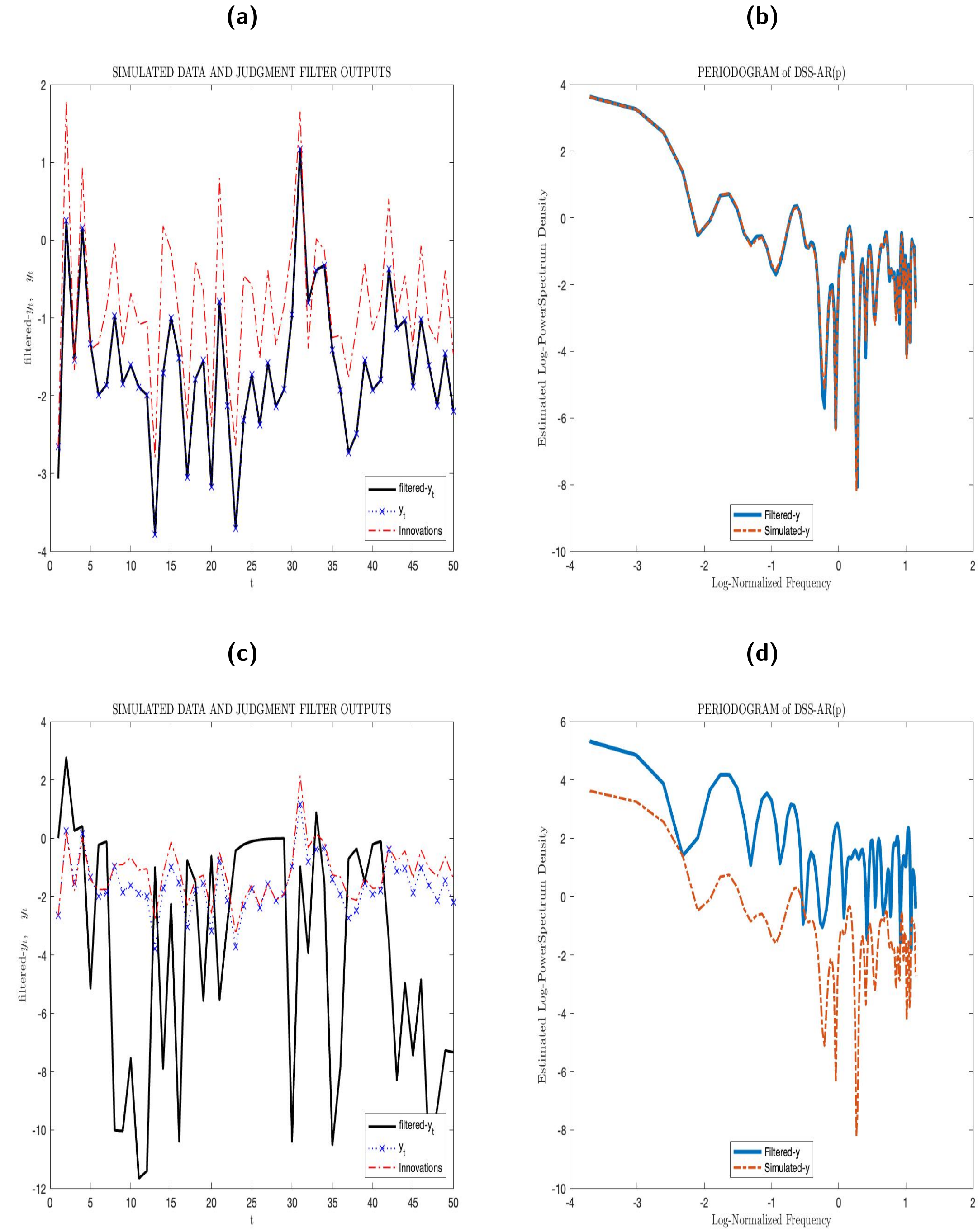
Additional Tables and Graphs

Figure 4: The Deformed Logarithm of a Standard Normal density



NOTE: This figure displays the behavior of the Deformed Logarithm applied to a standard normal probability density function f over the sequence $[-4; 4]$ given a set of values of q.

Figure 5: Examples of DJP-AR(1) and associated power-spectra for several q_0 and d_0= 0.



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.